



Instituto de Física Teórica
Universidade Estadual Paulista

DISSERTAÇÃO DE MESTRADO

IFT-D.009/05

**Generalizações do Movimento Browniano e suas Aplicações
à Física e a Finanças**

Dennis Fernandes Alves Bessada

Orientador

Prof. Dr. Gerson Francisco

Abril de 2005

*Ex Deo Nascimur
In Jesu Morimur
Per Sanctum Spiritum
Reviviscimus*

À Glória de Deus Todo Poderoso e de Jesus Cristo, o Salvador.

*“A virtude não tem padrão:
conforme cada um a honre ou despreze,
dela terá mais ou menos.”*

Platão

*“Não se deve dar ouvidos
àqueles que aconselham ao homem, por ser mortal,
que se limite a pensar coisas humanas e mortais;
ao contrário, porém, à medida do possível,
precisamos nos comportar como imortais
e tudo fazer para viver segundo a parte mais nobre
que há em nós.”*

Aristóteles

Dedicatória

Dedico este trabalho à Aparecida, minha mãe, à Elaine, minha esposa, à memória de meu pai, Antonio, à memória de meus avós, Luzia, Juvenal, Manuel e Anita, e aos meus irmãos, Pedro e Tânia.

AGRADECIMENTOS

- Agradeço ao Professor Gerson Francisco, não somente pela esmerada orientação, mas também pela amizade e pela confiança em mim depositada;
- Agradeço ao saudoso amigo de todas as horas, Antonio Peres Rubbi, pelo constante apoio;
- Agradeço aos grandes amigos, Victo dos Santos Filho, Walter Miguel e Eliezer Batista, aos quais tenho em profunda estima;
- Agradeço ao Professor Zimerman, meu primeiro orientador, a quem devo muito de meu conhecimento e aprendizado nas veredas da pesquisa científica;
- Agradeço ao Professor Nathan Berkovits, por sua orientação e grandes lições que me proporcionara;
- Agradeço aos Professores Bruto Pimentel, Diógenes Galetti, Gerhard Bund, Maria Cristina Abdalla, Paulo Leal Ferreira, Rogério Rosenfeld, José Francisco Gomes, Luiz Agostinho Ferreira, Antonio Aciolly, Hélio Fagundes, Alexandr Dorokhov, com os quais muito aprendi em agradáveis conversas, fossem no campo da Física ou da cultura geral;
- Agradeço também a todos os funcionários do IFT, do presente e do passado, em especial Sussumu Futaki, Luiz Carlos, Sr. Jorge, Édina, Zezé, Rosane, Marina, Alexandre, José Geraldo e Marcelo;
- Agradeço aos Professores Fernando Manuel Ramos e Reinaldo Rosas, do INPE, não somente pelas discussões sobre os temas aqui apresentados, mas fundamentalmente pela grande amizade.

ÍNDICE

INTRODUÇÃO.....	<i>i</i>
CAPÍTULO I.....	1
A TEORIA DA PROBABILIDADE	1
I.1 A Teoria da Probabilidade em Espaços Finitos	1
I.1.1 Conceitos Iniciais.....	1
I.1.2 Propriedades das Medidas de Probabilidade	3
I.1.3 Variáveis Aleatórias e Funções de Distribuição	7
I.1.4 Probabilidade Condicional e Independência.....	10
I.1.5 Valor Esperado e Variância de uma Variável Aleatória.....	15
I.1.6 O Passeio Aleatório	18
I.2 A Teoria Axiomática da Probabilidade.....	23
I.2.1 O Espaço de Probabilidade	23
I.2.2 Variáveis Aleatórias – Funções de Distribuição.....	27
I.2.3 Momentos	33
I.2.4 Convergência de Seqüências de Variáveis Aleatórias.....	38
I.2.5 Funções Características.....	41
I.2.6 Teoremas de Limite	46
CAPÍTULO II	53
A TEORIA DOS PROCESSOS ESTOCÁSTICOS.....	53
II.1 Conceitos Iniciais.....	53
II.2 Processos de Poisson e Processos de Wiener	60
II.3 Martingales	64
II.4 Processos de Markov	66
CAPÍTULO III.....	69
A TEORIA DE LÉVY	69
III.1 Introdução	69
III.2 Distribuições Infinitamente Divisíveis – Conceitos Iniciais.....	70
III.3 A Representação Canônica de Kolmogorov	76
III.4 As Representações Canônicas de Lévy-Khinchine e de Lévy.....	80
III.5 Distribuições Estáveis.....	83
CAPÍTULO IV.....	91
O MOVIMENTO BROWNIANO	91
IV.1 Introdução	91
IV.2 A Teoria de Einstein do Movimento Browniano.....	92
IV.2.1 A Teoria da Difusão.....	92
IV.2.2 O Conceito de Pressão Osmótica.....	95
IV.2.3 O Movimento de Partículas suspensas em um Líquido.....	98
IV.2.4 O Movimento Browniano	105
IV.3 A Teoria de Langevin	106

CAPÍTULO V	109
A TEORIA MATEMÁTICA DO MOVIMENTO BROWNIANO.....	109
V.1 O Movimento Browniano como um Processo Estocástico.....	109
V.1.1 O Passeio Aleatório e o Movimento Browniano	109
V.1.2 O Movimento Browniano Contínuo a Tempos Discretos	112
V.1.3 A Definição Geral do Movimento Browniano – Processos de Wiener	113
V.2 A Integral de Itô.....	116
V.3 Equações Diferenciais Estocásticas	122
V.3.1 O Processo de Wiener.....	124
V.3.2 O Movimento Browniano Geométrico	124
CAPÍTULO VI.....	127
APLICAÇÕES DO MOVIMENTO BROWNIANO: A TEORIA DE BLACK & SCHOLES.....	127
VI.1 Introdução	127
VI.2 Finanças – Conceitos Básicos.....	128
VI.2.1 Taxas de Juros.....	128
VI.2.2 Arbitragem e a Hipótese do Mercado Eficiente.....	131
VI.3 O Conceito de Derivativos.....	132
VI.4 Opções	133
VI.4.1 Definições	133
VI.4.2 Precificando Opções	135
VI.4.3 O Modelo Binomial de Precificação de Opções.....	137
VI.5 A Teoria de Black & Scholes	142
VI.5.1 A Equação de Black & Scholes	142
VI.5.2 A Solução da Equação de Black & Scholes para uma Opção Europeia... ..	145
CAPÍTULO VII	153
GENERALIZAÇÕES DO MOVIMENTO BROWNIANO E SUAS APLICAÇÕES ÀS SÉRIES FINANCEIRAS E À FÍSICA	153
VII.1 Processos de Lévy.....	153
VII.2 Cálculo do Expoente de Estabilidade para Processos Estocásticos Simulados	155
VII.2.1 Cálculo do Expoente de Estabilidade para um Movimento Browniano Simulado	156
VII.2.2 Cálculo do Expoente de Estabilidade para um Processo de Cauchy	162
VII.3 Cálculo do Expoente de Estabilidade para Séries Financeiras	166
CONCLUSÕES.....	171
BIBLIOGRAFIA.....	173
APÊNDICE.....	175
Aplicações da Teoria de Lévy à Precificação de Opções.....	175

RESUMO

Realizamos neste trabalho uma exposição geral da Teoria do Movimento Browniano, desde suas primeiras observações, feitas no âmbito da Biologia, até sua completa descrição segundo as leis da Mecânica Estatística, formulação esta efetuada por Einstein em 1905. Com base nestes princípios físicos analisamos a Teoria do Movimento Browniano de Einstein como sendo um processo estocástico, o que permite sua generalização para um processo de Lévy.

Fazemos uma exposição da Teoria de Lévy, e aplicamo-la em seguida na análise de dados provenientes do índice IBOVESPA. Comparamos os resultados com as distribuições empírica e a modelada via distribuição gaussiana, demonstrando efetivamente que a série financeira analisada apresenta um comportamento não-gaussiano.

Palavras-Chave:

Movimento Browniano, Processos Estocásticos não-Gaussianos

Áreas do Conhecimento:

Mecânica Estatística, Teoria dos Processos Estocásticos

ABSTRACT

We review in this work the foundations of the Theory of Brownian Motion, from the first observations made in Biology to its complete description according to the laws of Statistical Mechanics performed by Einstein in 1905. Afterwards we discuss the Einstein's Theory of Brownian Motion as a stochastic process, since this connection allows its generalization to a Lévy process.

After a brief review of Lévy Theory we analyze IBOVESPA data within this framework. We compare the outcomes with the empirical and gaussian distributions, showing effectively that the analyzed financial series behaves exactly as a non-gaussian stochastic process.

Keywords:

Brownian Motion, non-Gaussian Processes

Research Fields :

Statistical Mechanics, Theory of Stochastic Processes

INTRODUÇÃO

O Movimento Browniano surgiu no âmbito da Biologia como fruto da observação do movimento intermitente de partículas orgânicas e inorgânicas em suspensão em líquidos. Foi Albert Einstein [8] quem primeiramente analisou este fenômeno do ponto de vista físico, e nos legou a Teoria do Movimento Browniano tal qual conhecemos. A principal característica desta teoria se deve à solução obtida para a densidade de distribuição de partículas dispersas no líquido, cuja forma matemática segue uma função de densidade de probabilidade gaussiana.

A Teoria do Movimento Browniano tornou-se então a mais utilizada para a modelagem de fenômenos que envolvem um processo de difusão. Com isso, tornou-se sólida base para a análise de muitos fenômenos nos mais diversos ramos do conhecimento, dentre os quais a Biologia, a Química, e até mesmo Finanças, sem contar, é claro, a Física.

Na área de Finanças a Teoria do Movimento Browniano encontrou uma grande aplicação através da famosa teoria de Black & Scholes ([17] – [22] – [31] – [35]). Estes dois pesquisadores assumiram que a dinâmica de preços de um determinado ativo segue um Movimento Browniano, e obtiveram a partir desta proposição uma equação estocástica cuja solução permite a precificação de um objeto financeiro denominado *opção* (para uma introdução elementar, ver [1]). Esta teoria alcançou um estrondoso sucesso nos anos 70 e 80, tendo tornado-se uma das principais ferramentas de análise de investimentos desde então.

Porém, com os avassaladores avanços da tecnologia de informação, principalmente na última década, as medições da evolução temporal de preços atingiu uma alta resolução, o que significa dizer que se é possível obter dados da variação do preço de um ativo financeiro a cada minuto, por exemplo. Com isso, a complexidade do sistema aumentou significativamente devido à grande profusão de dados, o que permitiu efetuar análises mais precisas quanto à real dinâmica de preços do mercado. Em função disso, pode-se hoje perguntar: a dinâmica de preços do mercado segue realmente um Movimento Browniano?

A análise detalhada das distribuições de probabilidade empíricas de dados do mercado financeiro mostrou que a distribuição gaussiana não é a mais adequada para a modelagem dos chamados “*eventos raros*”, ou seja, eventos cuja probabilidade de ocorrência, apesar de ser muito pequena, é não nula. Para tais eventos, a probabilidade de ocorrência, segundo uma distribuição gaussiana, é nula. Com isso, vemos uma discrepância entre a distribuição empírica e a distribuição teórica, que assume a proposição de Black & Scholes, a de que a dinâmica de preços de um dado ativo segue um Movimento Browniano.

Esta inconsistência entre o empírico e o teórico levou muitos pesquisadores na década passada, principalmente físicos e matemáticos, a questionar o uso do Movimento Browniano como modelo de dinâmica de preços ([3] – [24]). Pesquisaram, então, qual seria a melhor alternativa para tentar reconciliar a observação com a teoria. Encontraram uma possível resposta em uma teoria formulada por um brilhante matemático francês da década de 20: a teoria de Paul Lévy ([13] – [21] – [23] – [29]).

A Teoria de Lévy apresenta uma fórmula para a densidade de distribuição de uma variável aleatória estável e infinitamente divisível que depende de dois parâmetros básicos: o *expoente de estabilidade*, α ($0 < \alpha \leq 2$), e o *expoente de assimetria*, δ ($-1 \leq \delta \leq 1$). Estes dois parâmetros dão origem a uma família infinita de funções densidade de probabilidade, dentre as quais figura a densidade de probabilidade gaussiana, cujo expoente de estabilidade é $\alpha = 2$. Com isso, pode-se dizer que a Teoria de Lévy *generaliza* a Teoria do Movimento Browniano, já que esta última é um caso particular da primeira.

A aplicação da Teoria de Lévy para a análise do comportamento temporal de séries financeiras tem dado muito bons resultados nos últimos anos [24].

O presente trabalho tem como principal objetivo trazer à tona estas questões e analisar a Teoria do Movimento Browniano, discutindo os conceitos subjacentes e indicando os pontos a partir dos quais uma teoria mais geral é necessária. Discutimos a Teoria de Lévy do ponto de vista matemático, e aplicamos a mesma para analisar o comportamento de séries financeiras do mercado brasileiro. Analisamos a distribuição de probabilidades associada aos dados do índice financeiro denominado IBOVESPA, usando principalmente algoritmos escritos em MATLAB para calcular seu expoente de estabilidade. Com isso, comparamos as curvas empírica e a teórica, calculada segundo a expressão de uma função densidade de probabilidade de Lévy.

A presente dissertação de mestrado está estruturada da seguinte maneira:

No Capítulo I discutimos fundamentalmente a Teoria da Probabilidade, que é a base sobre a qual toda a teoria do Movimento Browniano e a de Lévy são construídas.

No Capítulo II discutimos a Teoria dos Processos Estocásticos, onde analisamos a estrutura temporal de diversas variáveis aleatórias.

No Capítulo III introduzimos a Teoria de Lévy, onde também discutimos a questão do expoente de estabilidade, o ponto-chave para se diferenciar as soluções possíveis enquanto funções de distribuição de probabilidade.

No Capítulo IV apresentamos a Teoria do Movimento Browniano conforme discutida por Einstein em seu famoso artigo de 1905.

No Capítulo V discutimos a Teoria do Movimento Browniano como sendo parte da Teoria dos Processos Estocásticos.

No Capítulo VI introduzimos a Teoria de Black & Scholes como sendo uma aplicação da Teoria do Movimento Browniano à área de Finanças.

Finalmente, no Capítulo VII efetuamos as aplicações da teoria discutida para dados do índice IBOVESPA.

CAPÍTULO I

A TEORIA DA PROBABILIDADE

I.1 A Teoria da Probabilidade em Espaços Finitos

I.1.1 Conceitos Iniciais

Para introduzirmos as idéias básicas da Teoria da Probabilidade vamos recorrer aos seus primórdios na Idade Média ([12], [9], [32], [33]). Remontemos aos campos de batalha, onde os soldados medievais, nos intervalos entre combates, tinham como principal passatempo o jogo de dados. Jogavam da seguinte maneira: após lançar dois dados, somava-se os pontos obtidos em suas faces voltadas para cima. Venceria o jogo quem adivinhasse qual seria o resultado após um dado número de jogadas. Após inúmeras jogadas, em anos e anos de observações os soldados perceberam que, dentre todos os pontos possíveis, o número 7 era o que mais freqüentemente ocorria. Além disso, a observação mostrara que o número 6, apesar de ocorrer em menor freqüência que o número 7, ocorria com maior freqüência em relação ao número 5. Assim, através da *observação* os soldados perceberam que, mesmo ser poder prever *com certeza* qual seria o resultado do jogo, sabiam que teriam mais *chances* de ganhar caso apostassem nos números que mais *freqüentemente* ocorriam.

Vamos analisar agora as observações dos soldados em termos mais matemáticos. Os pontos possíveis de se obter são dados pelo conjunto discreto

$$\Omega = \{2,3,4,5,6,7,8,9,10,11,12\}. \quad \text{I.1}$$

Denotemos os resultados possíveis de cada lançamento de dados como o par ordenado $(,)$, onde na primeira posição colocaremos o resultado obtido do primeiro dado e na segunda posição colocaremos o resultado obtido do segundo. Com isso, os resultados possíveis são:

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Tabela I.1 – Resultados possíveis de se obter após o lançamento de dois dados

A partir da tabela I.1 podemos verificar as ocorrências dos pontos possíveis do conjunto I.1:

Pontos	2	3	4	5	6	7	8	9	10	11	12
Ocorrências	1	2	3	4	5	6	5	4	3	2	1

Tabela I.2 – Ocorrências dos pontos possíveis

Com isso, confirmam-se as observações dos soldados medievais: o número 7 é o que ocorre com maior frequência, com 6 ocorrências dentre as 36 possíveis. O número 6 ocorre mais frequentemente que o 5, conforme se observa da tabela (o número 6 ocorre cinco vezes, ao passo que o 5 ocorre apenas quatro). Deste modo, tem maior *chance de acertar* quem apostar nos números 6, 7 e 8, pois são eles que ocorrem com *maior frequência*.

Deste simples e histórico exemplo observamos que é de fato possível avaliarmos quais são, dentro de um conjunto de resultados possíveis, quais terão maior *chance* de ocorrer. E é neste contexto que surge a idéia de *probabilidade*. Antes de discutirmos seu significado, introduzamos alguns conceitos iniciais motivados pelo exemplo acima:

- ✘ **Experimento** : É o fenômeno que será objeto de uma análise matemática, devendo ser repetido inúmeras vezes em idênticas condições;
- ✘ **Espaço Amostral** : É o conjunto não-vazio de todos os resultados possíveis a serem obtidos de um experimento; denotaremos-lo pela letra Ω . Os elementos $\omega \in \Omega$ são denominados pontos amostrais.
- ✘ **Evento** : É o conjunto de todos os resultados que disponham de alguma semelhança entre si; são portanto subconjuntos do espaço amostral.

Como exemplos de experimentos podemos citar o movimento de partículas em um gás, o movimento de subida e descida de ações no mercado financeiro, a turbulência de um fluido, o cruzamento de genes, entre muitos outros. O espaço amostral pode ser qualquer conjunto, até mesmo o dos números reais; no entanto, nesta seção trataremos apenas de conjuntos finitos, como o I.1. . Já como exemplo de evento podemos tomar os resultados da tabela I.2.: lá encontramos o evento de ocorrência de números pares, ímpares, maiores que seis, etc. Para formarmos um evento basta apenas estabelecer sua lei de formação, como por exemplo o conjunto dos resultados pares.

Com estes conceitos em mente podemos agora introduzir o conceito de *probabilidade*. Devemos salientar que há outras definições possíveis, mas nos restringiremos aqui à mais usual. Tomemos como fundamento nosso exemplo: vimos que os resultados que ocorriam com *maior frequência* eram os números 6, 7 e 8; ou seja, podemos dizer, intuitivamente, que estes são os *mais prováveis* de ocorrer. Podemos apresentar inúmeros exemplos onde a maior frequência de ocorrência de um dado evento está intimamente relacionada com a idéia de este ser o mais *provável*. Esta conexão da idéia de probabilidade com a de frequência, no caso relativa ao espaço amostral é a base de toda a teoria em espaços finitos. Então, em face de tais colocações, introduzimos o seguinte conceito:

- ✘ **Medida de Probabilidade** : É a medida associada à frequência relativa de um evento A a seu espaço amostral discreto Ω ; denotando-se a probabilidade do evento A como $P[A]$, segue-se que

$$P[A] := \frac{\#(A)}{\#(\Omega)}, \quad \text{I.2}$$

onde $\#(\bullet)$ denota a cardinalidade do respectivo conjunto.

A identificação da probabilidade de um evento A com sua frequência relativa tem como premissa básica o fato experimental de que para um valor muito grande de $\#(\Omega)$ o raio $\#(A)/\#(\Omega)$ tende a ser constante.

Voltando ao exemplo histórico, vemos que, em termos de probabilidades a tabela I.2 realmente demonstra que o evento $\{7\}$ é o mais provável de ocorrer. Com isso, podemos concluir que a Teoria da Probabilidade é a realização matemática do processo de medição experimental.

I.1.2 Propriedades das Medidas de Probabilidade

Vamos explorar um pouco mais a noção de evento apresentada anteriormente. Da Teoria de Conjuntos segue-se que, se $A, B \subseteq \Omega$ então

$$A \cup B \subseteq \Omega, \quad \text{I.3a}$$

$$A \cap B \subseteq \Omega, \quad \text{I.3b}$$

$$A^c := \Omega \setminus A \subseteq \Omega. \quad \text{I.3c}$$

ou seja, as operações de união, intersecção e diferença entre eventos resulta também em um evento. No caso do exemplo histórico, consideremos o evento $A = \{2,4,6,8,10,12\}$ dos resultados pares e o evento $B = \{3,5,7,9,11\}$ dos resultados ímpares. Das operações acima segue-se que

$$A \cup B = \Omega, \quad \text{I.4a}$$

$$A \cap B = \emptyset, \quad \text{I.4b}$$

$$A^c = B. \quad \text{I.4c}$$

isto é, o conjunto \emptyset também é um evento. Com isso podemos definir

✘ **Evento Complementar :** Se A é um evento, então A^c é o evento que ocorre se e somente se o evento A não ocorrer;

✘ **Evento Impossível :** \emptyset é o evento que não pode ocorrer;

✘ **Eventos mutuamente Exclusivos :** Se $A, B \subset \Omega$ tal que $A \cap B = \emptyset$ dizemos que A e B são eventos mutuamente exclusivos, ou seja, a ocorrência de um *impede* a ocorrência do outro.

Conforme o exposto acima, uniões, intersecções, complementares e o conjunto vazio também são eventos. Com estes conjuntos esgotamos todas as possibilidades de eventos possíveis para um dado espaço amostral discreto; podemos formar uma *classe* com estes, que consistiria a *classe* dos eventos de Ω . Esta classe recebe um nome especial, *sigma-álgebra*, definida como

I. 1. Definição : Seja Ω um conjunto não-vazio. Uma σ - álgebra é uma classe \mathcal{F} de subconjuntos de Ω com as seguintes propriedades:

- ✱ $\emptyset \in \mathcal{F}$;
- ✱ Se $A \in \mathcal{F}$, então $A^c \in \mathcal{F}$;
- ✱ Se A_1, A_2, \dots é uma seqüência de conjuntos em \mathcal{F} , então $\bigcup_{n=1}^{\infty} A_n$ também está em \mathcal{F} .

Com este conceito podemos enunciar a seguinte propriedade, válida para espaços amostrais finitos:

I.2. Proposição [15] : Seja Ω um conjunto finito não-vazio e $\mathcal{P}(\Omega)$ a classe formada por todos os seus subconjuntos (que denominaremos de 'power set'). Neste caso

$$\mathcal{F} = \mathcal{P}(\Omega).$$

Uma sub-sigma-álgebra é uma sigma-álgebra \mathcal{G} tal que $\mathcal{G} \subseteq \mathcal{F}$.

Uma vez estabelecida esta propriedade, vamos analisar no que segue as probabilidades associadas a uma σ - álgebra finita. Da definição de probabilidade, equação I.2., vemos que esta é uma função definida sobre $\mathcal{P}(\Omega)$, cujo contradomínio é um conjunto $X \subset \mathcal{R}$. Tal subconjunto é encontrado através da seguinte propriedade: como para qualquer evento A vale a relação $\emptyset \subseteq A \subseteq \Omega$, concluímos que X é limitado, pois é válida a relação

$$\emptyset \subseteq A \subseteq \Omega \rightarrow P[\emptyset] \leq P[A] \leq P[\Omega]. \quad \text{I.5}$$

Da equação I.2 seguem imediatamente as seguintes propriedades:

$$P[\Omega] = 1; \quad \text{I.6}$$

$$P[\emptyset] = 0; \quad \text{I.7}$$

de onde concluímos que

$$P: \mathcal{P}(\Omega) \rightarrow [0,1]. \quad \text{I.8}$$

Pela relação acima vemos portanto que a sigma-álgebra associada ao espaço amostral de um dado experimento consiste no *domínio* da medida de probabilidade. E é por este fator que as sigma-álgebras desempenham um papel tão fundamental na teoria.

Para obtermos o valor da função P sobre uniões e intersecções entre elementos pertencentes à σ - álgebra $\mathcal{P}(\Omega)$ podemos recorrer às relações I.7 e I.8 da seguinte maneira: à partir das relações

$$\Omega \cup \emptyset = \Omega, \quad \text{I.9a}$$

$$\Omega \cap \emptyset = \emptyset, \quad \text{I.9b}$$

obtemos

$$P[\Omega \cup \emptyset] = 1, \quad \text{I.10a}$$

$$P[\Omega \cap \emptyset] = 0. \quad \text{I.10b}$$

As relações acima sugerem que

$$P[\Omega \cup \emptyset] = P[\Omega] + P[\emptyset], \quad P[\Omega \cap \emptyset] = P[\Omega] \cdot P[\emptyset]. \quad \text{I.11}$$

Com isso, motivados pela definição acima podemos *definir* a probabilidade da união de dois eventos em termos da medida I.2 da seguinte forma:

I.3. Definição : Sejam A e B eventos mutuamente exclusivos; então, a probabilidade da união destes eventos é igual à soma das probabilidades dos eventos individuais:

$$P[A \cup B] = P[A] + P[B]. \quad \text{I.12}$$

A equação acima é estendida facilmente para a medida de probabilidade (I.2) através da relação

$$P[A \cup B] := \frac{\#(A \cup B)}{\#(\Omega)} \quad \text{I.13}$$

para quaisquer eventos $A, B \subset \Omega$ tal que $A \cap B = \emptyset$.

A equação I.12 é essencial para que possamos enquadrar a Teoria da Probabilidade em espaços finitos no escopo da Teoria da Medida, o que resultará em um ganho formidável do ponto de vista matemático. Isso porque a Teoria da Medida tem bases lógicas extremamente bem fundamentadas, o que não acontecia com a Teoria da Probabilidade até o advento do trabalho de eminentes matemáticos do século passado. Na realidade, muitos dos grandes matemáticos anteriores ao século XVIII não creditavam a Teoria da Probabilidade como uma disciplina da Matemática, e sim como uma teoria à margem da mesma. A construção da Teoria da Probabilidade como uma Teoria da Medida, conforme ilustraremos na seção (I.2), representa a formalização necessária para sua inclusão no escopo das teorias matemáticas mais fundamentais para a descrição dos fenômenos naturais.

A equação (I.12) nos permite determinar imediatamente todas as operações básicas da medida de probabilidade sobre $\mathcal{P}(\Omega)$: ([4], [5], [15])

I.4. Proposição [4] – [5] - [15] : Para os eventos discriminados abaixo, todos pertencentes a $\mathcal{P}(\Omega)$, valem as propriedades

- I.** Se $\{A_k\}_{k=1}^n$ for uma seqüência de eventos mutuamente exclusivos (*i. e.*, para $k \neq l$ segue-se que $A_k \cap A_l = \emptyset$) é válida a seguinte relação:

$$P\left[\bigcup_{k=1}^n A_k\right] = \sum_{k=1}^n P[A_k];$$

- II.** $P[A^c] = 1 - P[A];$

III. Se $A \cap B \neq \emptyset$, então

$$P[A \cup B] = P[A] + P[B] - P[A \cap B];$$

IV. Para quaisquer eventos $\{A_k\}_{k=1}^n$ vale

$$P\left[\bigcup_{k=1}^n A_k\right] \leq \sum_{k=1}^n P[A_k].$$

Para ilustrar as propriedades acima consideremos o exemplo histórico. Tomemos dois eventos quaisquer, como, por exemplo, o evento $A = \{2,4,6,8,10,12\}$ dos resultados pares e o evento $B = \{7,8,9,10,11,12\}$ dos resultados com valor acima de 7. As probabilidades associadas a estes eventos são, de acordo com a equação (I.2), iguais a

$$P[A] = P[B] = \frac{6}{11} \quad \text{I.14}$$

Tomemos agora os complementares desses eventos, $A^c = \{3,5,7,9,11\}$ e $B^c = \{2,3,4,5,6\}$. Suas probabilidades de ocorrência são dadas respectivamente por

$$P[A^c] = P[B^c] = \frac{5}{11}. \quad \text{I.15}$$

Aplicando-se a propriedade **II.** ao resultado (I.14) obtemos o mesmo valor que o obtido em (I.15).

Calculemos agora as probabilidades associadas à união e à intersecção de eventos em $\mathcal{P}(\Omega)$. Pela equação (I.2) as probabilidades associadas à união dos eventos A e B , $A \cup B = \{2,4,6,7,8,9,10,11,12\}$, e associadas à sua intersecção, $A \cap B = \{8,10,12\}$, são dadas por

$$P[A \cup B] = \frac{9}{11} \quad \text{I.16a}$$

$$P[A \cap B] = \frac{3}{11}. \quad \text{I.16b}$$

Aplicando-se a propriedade **III.** obtemos, a partir de (I.14) e (I.16b) o mesmo valor que obtido em (I.16a).

Notemos que ainda falta a equação que relaciona a probabilidade da intersecção entre dois eventos; para melhor discutirmo-la com maior clareza faz-se necessária a introdução de alguns conceitos adicionais como *probabilidade condicional* e *independência*, que serão objetos da subseção I.

Com os conceitos discutidos até então podemos concluir vários fatos de capital importância. Vimos que, para analisarmos um experimento em termos probabilísticos, é fundamental termos três elementos:

- ✧ um *espaço amostral*, que nos dá os possíveis resultados de um dado experimento aleatório;
- ✧ uma *sigma-álgebra*, que nos fornece todas as relações possíveis entre os eventos do espaço amostral,
- ✧ e uma *medida de probabilidade*, que nos indica a probabilidade de ocorrência de qualquer evento da sigma-álgebra.

Com estes três elementos temos um sistema fechado, pois podemos assim calcular as probabilidades de *todos* os eventos de Ω visto que estamos considerando sua sigma-álgebra associada. Este fato somente é possível se introduzirmos a medida de probabilidade P dotada da propriedade advinda da Definição I.3 (conseqüentemente traduzida pela equação (I.13)). Assim, vemos que a estrutura matemática mais elementar para se construir um modelo probabilístico sobre um espaço finito consiste em uma tripla $(\Omega, \mathcal{P}(\Omega), P)$, introduzida pela seguinte

I.5. Definição : Seja Ω um espaço amostral finito e $\mathcal{P}(\Omega)$ sua sigma-álgebra associada. Sendo a função

$$P : \mathcal{P}(\Omega) \rightarrow [0,1]$$

uma medida de probabilidade dotada da propriedade

$$P[A \cup B] = P[A] + P[B],$$

denominamos a tripla $(\Omega, \mathcal{P}(\Omega), P)$ de *espaço de probabilidade finito*.

Será a partir da Definição I.5 que lançaremos as bases da Teoria Axiomática na seção I.2.

I.1.3 Variáveis Aleatórias e Funções de Distribuição

Após a estruturação matemática apresentada na subseção anterior, retornemos ao exemplo histórico da subseção I.1.1., a fim de elucidar alguns conceitos adicionais. Tomamos como eventos os resultados obtidos a partir da soma dos pontos presentes nas faces superiores de dois dados lançados, cujos valores possíveis foram coligidos na tabela I.1.. Considerando os dados desta tabela, verificamos que a cada dois pares ordenados simétricos corresponde um mesmo número, que é exatamente igual à soma das coordenadas de cada par. Matematicamente esta relação se traduz como

$$X[(x, y)] = x + y, \tag{I.17}$$

ou seja, o conjunto dos resultados $\mathcal{R}_x = \{2,3,4,5,6,7,8,9,10,11,12\}$ dos eventos é gerado pela função X definida acima. Em outras palavras, a informação sobre o resultado obtido decorre da função (I.17). Analisando sob este prisma, surge a seguinte questão: dado o resultado, é possível inferirmos matematicamente os eventos que lhe dão origem?

Para respondermos esta questão, vamos analisar o mapeamento inverso de (I.17): tomando-se

$$X^{-1}(x) = \{(a, b) \in \Omega : 2 \leq a + b \leq x\} \quad \text{I.18}$$

vemos que $X^{-1} \subset \mathcal{P}(\Omega)$, ou seja, o mapeamento inverso de X *preserva* a estrutura de sigma-álgebra do espaço de probabilidades $(\Omega, \mathcal{P}(\Omega), P)$. Com isso, podemos, a partir do espaço de resultados e da função (I.18), *reconstruir* o espaço dos eventos do jogo de dados. Esta importante propriedade nos leva a introduzir a seguinte

I.6. Definição : Seja $(\Omega, \mathcal{P}(\Omega), P)$ um espaço de probabilidades finito. Uma *variável aleatória simples* é uma função

$$X : \Omega \longrightarrow \mathcal{R}_x \quad \text{I.19}$$

tal que $A_x = \{\omega \in \Omega : X(\omega) = x\} \in \mathcal{P}(\Omega)$ para todo $x \in \mathcal{R}_x$.

A função mais simples que pode ser considerada como variável aleatória é a função indicadora χ_A , que tem como forma

$$\chi_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$$

Então, para o caso de uma função discreta que toma apenas alguns poucos valores,

$$X(\omega_k) = x_k$$

segue-se que podemos representá-la em termos de funções indicadoras; para isto, dividamos o espaço amostral Ω em uma *partição* $\{A_n\}_{n=1}^N$ de subconjuntos disjuntos, de tal modo que

$$\Omega = \bigcup_{n=1}^N A_n .$$

Deste modo, podemos construir uma variável aleatória como se fosse um vetor expandido em uma base $\{\chi_{A_n}\}$

$$X(\omega) = x_1 \chi_{A_1} + x_2 \chi_{A_2} + \dots + x_N \chi_{A_N} .$$

No caso específico $X(\omega_k) = x_k$, a expansão acima somente é possível se escolhermos a partição $A_k = \{\omega_k\}$.

O termo variável aleatória pode às vezes dar margem a confusões, uma vez que estas não são variáveis, mas simplesmente funções a valores reais. Ademais, estas não são aleatórias, visto que seu domínio de definição é um espaço amostral, e *não* um espaço gerado pela função de conjuntos $P[\otimes]$. Apesar destes fatores, preservou-se o termo variável aleatória para designar funções como as da Definição I.6.

A introdução do conceito de variáveis aleatórias representa um grande ganho do ponto de vista matemático, uma vez que a estrutura de espaço de medida é invariante sob tais funções. Para completarmos ainda mais a abrangência do conceito de variáveis aleatórias, podemos nos perguntar se é possível criarmos um espaço de medida *induzido* pela aplicação da variável aleatória sobre o espaço de probabilidades discreto. A resposta é afirmativa, pois, conforme já fora exposto na Proposição I.2, a sigma-álgebra associada ao espaço dos resultados \mathcal{R}_x é simplesmente seu ‘power set’ $\mathcal{P}(\mathcal{R}_x)$. Lembrando-se que o espaço de resultados é gerado pela variável aleatória X , podemos então fazer a seguinte associação:

I.7 Definição : Seja X uma variável aleatória sobre $(\Omega, \mathcal{P}(\Omega), P)$; a *sigma-álgebra gerada por X* , $\sigma[X]$, é a classe de todos os conjuntos $\{\omega \in \Omega : X(\omega) \in A\}$, onde $A \in \mathcal{R}_x$.

Assim, a sigma-álgebra $\mathcal{P}(\mathcal{R}_x)$ é, na realidade, igual à $\sigma[X]$.

Resta-nos agora saber qual seria a medida associada à esta sigma-álgebra. Esta medida deve manter as mesmas propriedades da medida de probabilidade, pois assim garantiríamos a preservação da estrutura do espaço de probabilidades. A medida a ser considerada deve ter então como domínios

$$p: \mathcal{R}_x \longrightarrow [0,1]. \quad \text{I.20}$$

Deste modo, a medida em questão seria uma função a valores reais, e não uma função de conjuntos. Esta característica representa uma grande simplificação estrutural. Para encontrarmos sua forma explícita, consideremos as relações

$$X^{-1}: \mathcal{R}_x \longrightarrow \mathcal{P}(\Omega) \quad \text{e} \quad P: \mathcal{P}(\Omega) \rightarrow [0,1],$$

de onde se segue imediatamente a seguinte composição:

$$P \circ X^{-1}: \mathcal{R}_x \longrightarrow [0,1]. \quad \text{I.21}$$

Com isso, a função definida acima possui o domínio e o contradomínio requeridos em (I.20). Para encontrarmos a expressão desta nova função, basta usarmos o conjunto $A_x = \{\omega \in \Omega : X(\omega) = x\}$, onde X é a variável aleatória definida em (I.17) e $x \in \mathcal{R}_x$:

$$p(x) = P[\omega \in \Omega : X(\omega) = x]. \quad \text{I.22}$$

É evidente que a função real acima, somada em todos os valores de x , assume 1 como resultado:

$$\begin{aligned} \sum_{x \in \mathcal{R}_x} p(x) &= \sum_{x \in \mathcal{R}_x} P[X(\omega) = x] \\ &= P \left[\bigcup_{x \in \mathcal{R}_x} \{X(\omega) = x\} \right] \\ &= 1, \end{aligned}$$

pois $\bigcup_{x \in \mathfrak{R}_x} \{X(\omega) = x\} = \Omega$. Portanto, podemos concluir que as funções (I.22) preservam a estrutura das medidas de probabilidades, uma vez que a soma destas funções sobre todos os valores do espaço dos resultados é igual a 1. Introduzamos então a seguinte

I. 8. Definição : Seja $(\Omega, \mathcal{P}(\Omega), P)$ um espaço de probabilidades finito e X uma variável aleatória definida sobre este; a função à valores reais

$$p(x) := P[\omega \in \Omega : X(\omega) = x]$$

é denominada *função densidade de probabilidade discreta*. O conjunto $\{p(x_1), \dots, p(x_N)\}$ consiste na *distribuição de probabilidades* de Ω , e satisfaz

$$\sum_{k=1}^N p(x_k) = 1.$$

A partir do conceito de variáveis aleatórias e de densidades de probabilidade podemos então concluir que a relação

$$(\Omega, \mathcal{P}(\Omega), P) \xrightarrow{X} (\mathfrak{R}_x, \mathcal{P}(\mathfrak{R}_x), p) \quad \text{I.23}$$

é válida em espaços finitos. A vantagem deste mapeamento é que, em sendo as funções de densidade de probabilidade definidas em um subconjunto dos reais, é sempre possível estabelecermos uma fórmula simples para as mesmas. Veremos em seções e capítulos posteriores a imensa utilidade e importância das funções densidade de probabilidade.

I.1.4 Probabilidade Condicional e Independência

Na seção anterior discutimos as propriedades fundamentais da medida de probabilidade em espaços finitos, decorrentes do axioma I.3. Enunciamos tal axioma em função de duas propriedades da teoria de conjuntos relacionadas ao espaço amostral e ao evento impossível (I.9a) e (I.9b), que levou ao resultado I.11. Note-se que a relação

$$P[\Omega \cap \emptyset] = P[\Omega] \cdot P[\emptyset] \quad \text{I.24}$$

é sugerida em função do fato de Ω e \emptyset serem obviamente mutuamente exclusivos. A pergunta que surge em face disso é: qual seria a relação entre dois eventos que não são mutuamente exclusivos?

Para respondermos a esta pergunta vamos analisar detalhadamente o conceito de intersecção de eventos. Para tanto, vamos recorrer a um exemplo. Suponhamos que, dentro de uma empresa Ω com $\#(\Omega)$ funcionários hajam $\#(A)$ diretores e $\#(B)$ profissionais formados em Física. É evidente que, se houver diretores formados em física, estarão eles enquadrados em $A \cap B$; portanto, se houver uma escolha ao acaso de um funcionário da empresa existe uma probabilidade não-nula $P[A \cap B]$ deste ser um diretor e formado em Física.

Analisemos então os eventos de modo a calcularmos esta probabilidade. Dividindo-se o processo de escolha em duas etapas, onde a primeira se refere à escolha propriamente dita e a segunda à verificação do resultado, e estabelecendo-se as combinações possíveis, chegamos à seguinte tabela:

		Etapas	
		Escolher a pessoa ao acaso	Verificar se é formado em Física
Eventos	$\omega \in A$ (Diretor)	$\omega \in B$	$\omega \notin B$
		$\omega \notin A$ (Não é um Diretor)	$\omega \in B$

Tabela I.3 – Possibilidades de escolha ao acaso dos funcionários da empresa

Por esta tabela podemos observar que, se A ocorrer, então há $\#(A \cap B)$ possibilidades de escolha em $\#(A)$ possíveis; por outro lado, se A não ocorrer observamos que não há nenhuma possibilidade deste funcionário escolhido ser diretor, muito menos então ser um diretor formado em Física. Com isso, verificamos que o evento ‘escolha de um diretor formado em Física’ *depende*, ou, de outra maneira, está *condicionado* à *ocorrência do evento* ‘escolha de um diretor’. Então, podemos concluir que, neste caso, tais eventos apresentam uma *relação de dependência entre si*, implicando que *a ocorrência de um evento possibilita a ocorrência do outro*.

Antes de definir mais formalmente estes eventos dependentes, calculemos a probabilidade de ocorrência do evento A dado que o evento B ocorreu. Como vimos, se A ocorreu, então a probabilidade que B ocorra é

$$\frac{\#(A \cap B)}{\#(A)},$$

ou então

$$\frac{P[A \cap B]}{P[A]}.$$

Com isso, introduzimos a seguinte

I. 9. Definição : Seja A um evento com $P[A] > 0$; Então a *probabilidade condicional de um evento B dado que A ocorreu (ou probabilidade de B dado A)* é dada pela equação

$$P[B | A] = \frac{P[A \cap B]}{P[A]}. \quad \text{I.25}$$

A probabilidade condicional I.17 satisfaz todas as propriedades de uma medida de probabilidade ([2], [4], [5], [9], [12], [15], [32]) :

I. 10. Teorema : São válidas as seguintes relações:

- I. $P[A | A] = 1$;
- II. $P[\emptyset | A] = 0$;
- III. $P[A \cup B | C] = P[A | C] + P[B | C] - P[A \cap B | C]$;
- IV. Para os eventos $\{A_k\}_{k=1}^n \subset \Omega$ vale

$$P\left[\bigcap_{k=1}^n A_k\right] = P[A_1] \cdot P[A_2 | A_1] \cdot P[A_3 | A_1 \cap A_2] \cdots P\left[A_n | \bigcap_{k=1}^{n-1} A_k\right].$$

Observemos que na definição de probabilidade condicional mencionamos um exemplo onde um evento A precedia um outro evento B , guardando este último uma relação de dependência com o primeiro. Vimos que, neste caso, tais eventos são mutuamente exclusivos. No entanto, há inúmeros casos onde dois eventos, que não são mutuamente exclusivos, apresentam *independência* entre si, *i. e.*, a ocorrência de um não afeta o resultado do outro. Para ilustrarmos melhor este conceito vamos recorrer a um exemplo: suponhamos que um dado jogo é constituído por duas etapas simultâneas, onde a primeira etapa consiste na retirada de uma carta e a segunda consiste no lançamento de um dado. Seria vencedor o jogador que conseguisse retirar uma carta vermelha e obter a face com cinco pontos no dado. A pergunta que colocamos é: qual a probabilidade de ganho deste jogador?

Para analisarmos este jogo, consideremos que Ω_1 seja o espaço amostral dos resultados possíveis do evento ‘retirar uma carta’ e Ω_2 seja o espaço amostral do evento ‘lançar um dado’. Podemos considerar que Ω , o espaço amostral do jogo em questão, seja o produto cartesiano $\Omega_1 \times \Omega_2$, uma vez que tal jogo é resultado de duas etapas simultâneas. A partir disso, consideremos A o evento ‘carta vermelha’ (com qualquer pontuação obtida no dado) e B o evento ‘cinco pontos’ (independentemente da cor da carta), com cardinalidades $\#(A) = 156$ e $\#(B) = 52$. Usando o fato que $\#(\Omega) = 312$, obtemos

$$P[A] = \frac{1}{2}, \quad P[B] = \frac{1}{6}. \quad \text{I.26a}$$

A partir do resultado $\#(A \cap B) = 26$, segue-se que

$$P[A \cap B] = \frac{1}{12}. \quad \text{I.26b}$$

Usando-se a fórmula I.17, vemos que

$$P[B|A] = \frac{1}{6}, \quad \text{I.26c}$$

o que nos indica o seguinte resultado:

$$P[B|A] = P[B] \text{ e } P[A \cap B] = P[A] \cdot P[B]. \quad \text{I.27}$$

Vamos analisar melhor a informação contida na equação I.19. Pelo exemplo posto, podemos observar que a *ocorrência de A não interfere na ocorrência de B*; com, isso, o conceito de probabilidade condicional, pelos cálculos acima, se reduz simplesmente ao de probabilidade usual. Isso significa que se dois eventos forem *independentes*, a informação ‘evento A ocorreu’ *não altera* o resultado da informação ‘evento B ocorreu’.

O resultado acima é importantíssimo dentro do escopo da Teoria da Probabilidade, e consistirá na base de toda a discussão que faremos a respeito da Teoria do Movimento Browniano. Antes de definirmos então este conceito tão importante, vamos provar o resultado I.19 para um caso mais geral.

Consideremos, como no exemplo acima, que um experimento seja constituído por duas etapas, onde o espaço amostral de cada etapa é representado por Ω_1 e Ω_2 . Neste caso, $\Omega = \Omega_1 \times \Omega_2$ e $\#(\Omega) = \#(\Omega_1) \#(\Omega_2)$. Seja A um dado evento caracterizado por um conjunto $A_1 \subset \Omega_1$ de resultados (no exemplo acima, A_1 era simplesmente o conjunto das cartas vermelhas) e B outro evento caracterizado por um conjunto $A_2 \subset \Omega_2$ de resultados (pelo mesmo exemplo, A_2 era o conjunto ‘cinco pontos’); com isso, $A = \{(x, y) : x \in A_1, y \in \Omega_2\}$ e $B = \{(x, y) : x \in \Omega_1, y \in A_2\}$. As cardinalidades associadas a estes eventos são, respectivamente, dadas por $\#(A) = \#(A_1) \#(\Omega_2)$ e $\#(B) = \#(\Omega_1) \#(A_2)$. A intersecção dos eventos A e B é dada por $A \cap B = \{(x, y) : x \in A_1, y \in A_2\}$, com cardinalidade $\#(A \cap B) = \#(A_1) \#(A_2)$.

Com os resultados acima, podemos calcular a probabilidade condicional $P[B|A]$ a partir da equação I.17:

$$\begin{aligned} P[B|A] &= \frac{\#(A \cap B)}{\#(A)} \\ &= \frac{\#(A_1) \#(A_2)}{\#(A)} \\ &= \frac{1}{\#(A)} \frac{\#(A)}{\#(\Omega_1)} \frac{\#(B)}{\#(\Omega_2)} \\ &= P[B]. \end{aligned}$$

Com isso, provamos que o resultado I.19 é válido para casos mais gerais. Podemos então introduzir a seguinte

I. 11. Definição : Dois eventos são *estocasticamente independentes* se

$$P[B | A] = P[B] \quad \text{I.28}$$

ou seja, a informação relativa à ocorrência do evento A *não interfere* no resultado da informação relativa à ocorrência do evento B . Ademais,

$$P[A \cap B] = P[A] \cdot P[B]. \quad \text{I.29}$$

A propriedade I.21 pode ser estendida para um número arbitrário de eventos, conforme demonstra a proposição a seguir ([2], [4], [5], [9], [12], [15], [32]) :

I. 12. Proposição : Os eventos $\{A_k\}_{k=1}^n \subset \Omega$ são denominados *mutuamente independentes* se valerem as seguintes relações:

$$P[A_i \cap A_j] = P[A_i]P[A_j], \quad i, j \in K$$

$$P[A_i \cap A_j \cap A_k] = P[A_i]P[A_j]P[A_k], \quad i, j, k \in K$$

.....

$$P\left[\bigcap_{k=1}^n A_k\right] = \prod_{k=1}^n P[A_k],$$

onde $K = \{1, 2, \dots, n\}$.

Até o momento em nossa discussão acerca de probabilidade condicional e independência não mencionamos o uso de sigma-álgebras. Tampouco estendemos tal conceito para variáveis aleatórias. Faria sentido falarmos de sigma-álgebras e variáveis aleatórias independentes? A resposta é positiva para a primeira questão: como sigma-álgebras são compostas de eventos, e como já sabemos determinar se dois eventos são independentes, faz sentido então considerar que duas sigma-álgebras sejam independentes se eventos de uma sejam independentes em relação a eventos da outra; em outras palavras,

I.13. Definição : Sejam \mathcal{G} e \mathcal{H} sub-sigma-álgebras de \mathcal{F} . Dizemos que \mathcal{G} e \mathcal{H} são *independentes* se

$$P[A \cap B] = P[A]P[B] \quad \forall A \in \mathcal{G}, B \in \mathcal{H}$$

Um exemplo instrutivo de se considerar é o caso de variáveis aleatórias definidas sobre um par cartesiano. Sejam X e Y duas variáveis aleatórias sobre Ω ; consideremos agora o espaço amostral Ω^2 e o seguinte conjunto:

$$A \times B = \{(X, Y) : X \in A, Y \in B\}.$$

Pela definição acima vemos que o produto cartesiano $A \times B$ consiste simplesmente no seguinte evento:

$$\{(X, Y) \in A \times B\} = \{X \in A\} \cap \{Y \in B\}, \quad \text{I.30}$$

ou seja, para que X e Y sejam independentes basta que

$$\begin{aligned} P[(X, Y) \in A \times B] &= P[\{X \in A\} \cap \{Y \in B\}] \\ &= P[X \in A]P[Y \in B]. \end{aligned} \quad \text{I.31}$$

Para a segunda questão também é afirmativa a resposta: como podemos gerar sigma-álgebras a partir da ação de variáveis aleatórias, torna-se fácil pela definição acima decidir se tais sigma-álgebras geradas são independentes; assim, vale a seguinte

I.14. Definição : Dizemos que duas variáveis aleatórias X e Y são *independentes* se as sigma-álgebras geradas $\sigma[X]$ e $\sigma[Y]$ são independentes.

Nas seções que seguem discutiremos mais detalhadamente as importantes implicações que os conceitos de probabilidade condicional e independência trazem à Teoria da Probabilidade. Será com base no conceito de probabilidade condicional, por exemplo, que construiremos os importantes conceitos da Teoria dos Processos Estocásticos, como Processos de Markov e Martingales, fundamentais para uma completa discussão do Movimento Browniano e suas conseqüentes generalizações.

I.1.5 Valor Esperado e Variância de uma Variável Aleatória

Nas subseções que se sucederam até o momento, vimos que a base de qualquer teoria probabilística consiste em um espaço de probabilidades $(\Omega, \mathcal{P}(\Omega), P)$. Toda a estrutura de eventos de um dado experimento se encontra no bojo de sua sigma-álgebra associada $\mathcal{P}(\Omega)$; em outras palavras, podemos dizer que toda a *informação* respectiva a um dado experimento está contida em sua sigma-álgebra. Podemos então categorizar um dado sistema simplesmente em se conhecendo sua sigma-álgebra.

Para termos uma idéia da abrangência deste conceito, consideremos um exemplo bem conhecido: um censo demográfico. Neste, os pesquisadores colhem inúmeras informações relevantes para um melhor conhecimento da população de uma nação: idade, sexo, grau de escolaridade, atividade profissional, renda familiar, entre outras. Com todos estes dados, compomos o perfil de toda uma população. Pensando em termos de conjuntos fica fácil identificarmos uma sigma-álgebra associada ao experimento ‘censo demográfico’: $\mathcal{F} = \{\text{Toda a população } (\Omega), \emptyset, \{\text{conjunto dos jovens } (J)\}, \{\text{conjunto dos homens } (H)\}, \{\text{conjunto das mulheres } (M)\}, \dots, \{\text{conjunto dos homens jovens } (H \cap J)\}, \{\text{conjunto das mulheres jovens } (M \cap J)\}, \dots\}$. Por este exemplo notamos realmente que todas as informações relacionadas à uma dada população estão em sua sigma-álgebra.

No entanto, ainda nos falta algo para complementar estas informações. Temos subconjuntos das populações, mas ainda não dispomos de um elemento que represente a população *como um todo*. Tal elemento nos permitiria, por exemplo, *comparar* dois conjuntos de dados diferentes. Para melhor esclarecermos este ponto, tomemos o seguinte exemplo: suponhamos que o governo suíço tenha encomendado uma pesquisa que vise determinar alguns índices de desenvolvimento, como renda *per capita*. Isso significa que o governo deseja obter um número que *melhor represente* a renda da população total do país. Assim, um número representaria a população como um todo, que é, neste caso, simplesmente obtido

dividindo-se a renda total arrecadada em um dado exercício fiscal pelo número total da população. Um outro número que também poderia quantificar a renda de uma população seria sua *renda média*, ou seja, uma quantia que é *intermediária* em relação aos maiores e menores valores de renda desta população. Para calcularmos o valor desta renda média basta simplesmente tomarmos todas as rendas individuais, somá-las e finalmente dividi-las pelo número total de rendas consideradas. Assim, se r_i é o valor da renda do i -ésimo habitante, e N é o número de habitantes, então a renda média será dada por

$$\bar{r} = \frac{1}{N} \sum_{i=1}^N r_i .$$

Deste modo, em sendo a renda média um número, podemos fazer comparações entre diferentes conjuntos de dados; assim, se \$7000 for a renda média da Suíça e \$6500 for a da Alemanha, vemos que, *proporcionalmente*, um suíço ganha mais que um alemão. É claro que, dentro deste exemplo, não podemos afirmar que todo suíço ganha melhor que todo alemão; sem dúvida podem haver alemães que ganhem muito mais que alguns suíços, e pode até acontecer que a renda total alemã seja maior que a Suíça (como aliás o é). No entanto, o que a média reflete é a proporção entre a renda total e a população, por isso é que na Suíça *em média* (ou seja, proporcionalmente) se ganha mais do que na Alemanha.

Pela discussão acima notamos que a média aritmética é um bom critério para compararmos conjuntos de dados. Consideremos agora tais médias dentro do contexto da Teoria da Probabilidade. Sendo o espaço amostral $\Omega = \{\text{população suíça}\}$, introduzamos a variável aleatória

$$X(\omega_k) = r_k ,$$

onde ω_k representa um habitante da Suíça e r_k representa sua renda. É notável que podem haver pessoas obtendo rendas iguais; suponhamos que hajam $m < n$ rendas que se repitam. Portanto, sendo n_k ($k = \{1, 2, \dots, m\}$) o número de vezes que a renda r_k comparece na análise, segue-se que o valor médio da renda será dado pela equação

$$EX = \frac{1}{N} \sum_{k=1}^m n_k r_k .$$

No contexto da Teoria da Probabilidade em espaços finitos, sabemos que a razão n_k / N representa a probabilidade de que a renda de uma pessoa escolhida ao acaso seja igual a r_k ; com isso, rescrevendo-se a equação acima em termos de probabilidades obtemos a seguinte

I. 15. Definição : Seja $(\Omega, \mathcal{P}(\Omega), P)$ um espaço de probabilidades finito e X uma variável aleatória sobre Ω . O *valor esperado* de X é definido por

$$EX := \sum_{\omega \in \Omega} X(\omega) P[\{\omega\}] .$$

A equação acima pode ainda ser rescrita observando-se o seguinte fato: como o espaço amostral Ω é finito, a variável aleatória X pode assumir apenas alguns poucos valores,

$\{x_1, x_2, \dots, x_m\}$; particionando-se Ω nos subconjuntos $\{X(\omega_1) = x_1\}$, $\{X(\omega_2) = x_2\}, \dots,$
 $\{X(\omega_m) = x_m\}$ segue-se que

$$\begin{aligned}
 EX &:= \sum_{\omega \in \Omega} X(\omega)P[\{\omega\}] \\
 &= \sum_{k=1}^m \sum_{\omega \in \{X=x_k\}} X(\omega)P[\{\omega\}] \\
 &= \sum_{k=1}^m x_k \sum_{\omega \in \{X=x_k\}} P[\{\omega\}] \\
 &= \sum_{k=1}^m x_k P[X = x_k] \\
 &= \sum_{k=1}^m x_k p(x_k).
 \end{aligned} \tag{I.32}$$

Uma vez estabelecido o conceito de valores esperados, discutamos algumas de suas propriedades ([2], [4], [5], [9], [12], [15], [32]) :

I.16. Teorema : Sejam X e Y duas variáveis aleatórias independentes e α um número real. Então vale:

- I. $E[X + Y] = EX + EY$;
- II. $E[\alpha X] = \alpha EX$;
- III. $E[XY] = EX \cdot EY$.

I.33

Em nossa discussão anterior foi-se colocado que o valor esperado de uma variável aleatória *representaria* o conjunto formado por todos os seus valores numéricos possíveis; no entanto, podemos nos perguntar: quão boa é esta representação? Em outras palavras, em *quanto* os valores da variável aleatória se *desviam* do seu valor médio? Poderíamos pensar em termos da soma dos desvios individuais $X(\omega) - EX$; no entanto, tais desvios poderiam ser simétricos e assim cancelarem-se uns aos outros. Isto é, poderíamos ter um desvio total igual à zero mesmo que a variável aleatória apresentasse uma oscilação em torno de seu valor esperado. Poderíamos eliminar este fato se tomarmos $|X(\omega) - EX|$; no entanto, a expressão do desvio seria não-diferenciável, portanto impraticável. O melhor a se fazer então seria tomar os quadrados de cada termo e somá-los, para assim obtermos um valor para o desvio da variável aleatória em termos de seu valor médio. Com isso, introduzimos então a seguinte

I. 16 Definição : Seja $(\Omega, \mathcal{P}(\Omega), P)$ um espaço de probabilidades finito e X uma variável aleatória sobre Ω . A *variância de X* é definida como sendo a seguinte expressão:

$$\text{var } X := \sum_{\omega \in \Omega} \{X(\omega) - EX\}^2 P[\{\omega\}].$$

Ou seja,

$$\text{var } X = \sum_{k=1}^m (x_k - EX)^2 p(x_k),$$

ou ainda

$$\text{var } X = E[X - EX]^2.$$

Desenvolvendo-se o lado direito da equação acima obtemos a equação

$$\text{var } X = EX^2 - (EX)^2, \quad \text{I.34}$$

muito mais útil do ponto de vista operacional.

A variância possui propriedades semelhantes às dos valores esperados ([2], [4], [5], [9], [12], [15], [32]) :

I.17. Teorema : Sejam X e Y duas variáveis aleatórias independentes e α um número real. Então vale:

I. $\text{var}[X + Y] = \text{var}[X] + \text{var}[Y];$

II. $\text{var}[\alpha X] = \alpha^2 \text{var}[X].$

Pela definição de variância, podemos introduzir um conceito mais geral, o de *covariância*. Consideremos uma certa variável aleatória Y , com valor esperado EY ; definamos agora a seguinte expressão:

$$\text{cov}(X, Y) = E[(X - EX)(Y - EY)];$$

é fácil observar que, se X e Y forem independentes, $\text{cov}(X, Y) = 0$; portanto, a covariância serve como um indicador de independência de duas variáveis aleatórias. Ademais, também é fácil mostrar que

$$\text{cov}(X, X) = \text{var } X.$$

I.1.6 O Passeio Aleatório

Vimos no início deste capítulo que a teoria da probabilidade nasceu atrelada aos jogos de azar. Tanto isso é verdade que os primeiros problemas que envolviam o conceito de probabilidade tinham como fonte as possibilidades de resultados de jogos de dados, de moedas, de cartas, e assim por diante. As correspondências entre Pascal e Fermat, consideradas como a gênese da teoria da probabilidade como parte da matemática continham vários problemas associados a estes mesmos jogos. Nesta subseção nos centraremos em um jogo que, de tão simples, não mereceria, a princípio, atenção; entretanto, suas conseqüências são tão abrangentes, que encontraremos inúmeros exemplos em outros campos do conhecimento de seu uso. Isso porque o jogo que consideraremos, o jogo de moedas, tem como estrutura básica um conceito de aplicação universal: o *passeio aleatório*. Este conceito é

tão importante que é através dele que se deriva o principal teorema da teoria da probabilidade: o *teorema central do limite*.

Consideremos um jogo que consista no lançamento de uma moeda três vezes. Suponhamos que a tônica do jogo seja a seguinte: a cada cara obtida, ganha-se \$1, ao passo que a cada coroa obtida não se ganha nada. Qual seriam então as probabilidades de ganhos neste jogo, a cada passo?

Iniciemos nossa análise fazendo algumas considerações iniciais. Chamaremos cada lançamento realizado de ‘passo’. Denominemos o resultado ‘cara’ como H e o resultado ‘coroa’ como T . Assim, no primeiro passo, podemos obter ou H , ou T ; no segundo passo, o resultado associado ao lançamento da moeda também é H ou T , mas, em termos de *resultado acumulado*, obteríamos HH , HT , TH ou TT , dependendo do resultado do primeiro passo. Os eventos acumulados consistem, portanto, de seqüências dos resultados anteriores; assim, mais natural então representá-los como sendo n -uplas do tipo (H, T, \dots, H) , e assim por diante. Portanto, o espaço amostral associado a cada lançamento de moeda é $\Omega = \{H, T\}$, ao passo que o espaço amostral associado ao resultado acumulado no n -ésimo passo é dado pelo produto cartesiano $\Omega^n = \times_{k=1}^n \Omega$.

Os resultados obtidos a cada passo podem ser representados por meio de um diagrama muito útil, denominado *árvore binomial*:

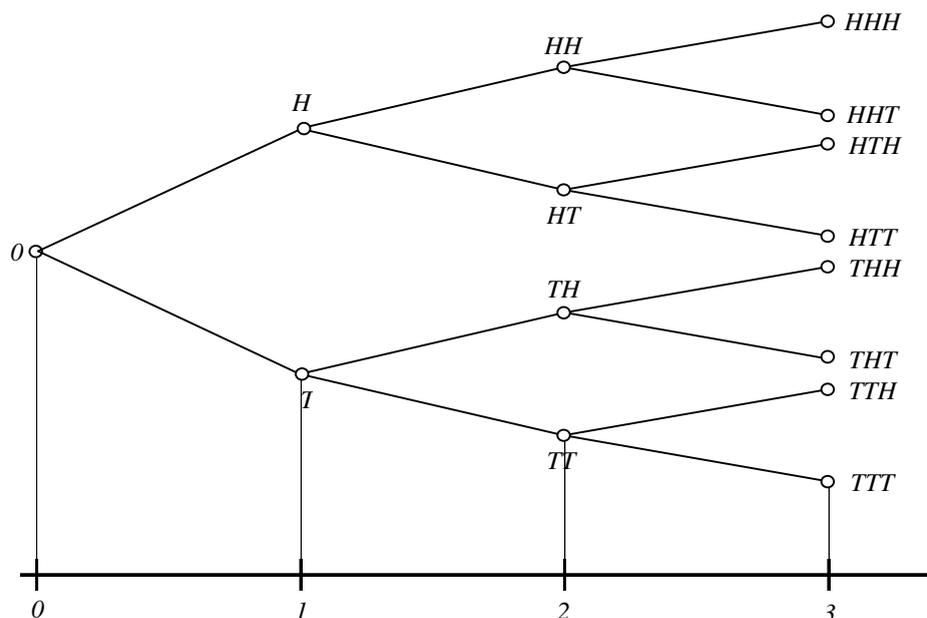


Figura I.1 – Representação gráfica dos resultados do lançamento de uma moeda três vezes

De acordo com estas duas situações distintas, teremos também dois tipos diferentes de ganhos. O primeiro tipo se relaciona ao ganho obtido a cada passo; já o segundo, se relaciona ao ganho acumulado até um determinado passo. Em relação à primeira situação, como a cada passo ou o jogador ganha \$1 ou não ganha nada, podemos representar este ganho por meio da variável aleatória $X_k : \Omega \rightarrow \mathcal{R}$,

$$X_k(H) = +1, \quad X_k(T) = 0,$$

onde k indica o passo. Já para a segunda situação, o resultado acumulado consiste no número de caras obtidas até o n -ésimo passo, que por sua vez é dado pela soma do número de caras obtidas a cada passo $k \leq n$; portanto, se $S_n : \Omega^n \rightarrow \mathcal{R}$ é a variável aleatória que nos fornece este número, esta assume a forma

$$S_n(\omega) = \sum_{k=1}^n X_k(\omega_k), \quad \text{I.35}$$

onde $\omega = (\omega_1, \omega_2, \dots, \omega_n)$.

Uma vez estabelecidas as variáveis aleatórias que consistem na base de nosso jogo, passemos ao cálculo da probabilidade de ocorrência de cada resultado. Consideremos que a probabilidade de se obter uma cara a cada passo seja igual a

$$P[X_k(\omega) = +1] = p, \quad \text{I.36a}$$

onde $0 < p < 1$, e que a probabilidade de obtermos coroa seja igual a

$$P[X_k(\omega) = 0] = q, \quad \text{I.36b}$$

onde naturalmente $q = 1 - p$. Com estes resultados em mão, podemos deduzir as probabilidades para cada resultado acumulado. Vimos na subseção I.1.4. como calcular a probabilidade de variáveis aleatórias definidas sobre um produto cartesiano; utilizemos este conceito então para calcularmos as probabilidades associadas a este jogo. Definamos inicialmente a nova variável

$$Z_n : \Omega^n \rightarrow \mathcal{R}$$

$$(\omega_1, \dots, \omega_n) \mapsto (X_1(\omega_1), \dots, X_n(\omega_n));$$

por esta definição torna-se claro que cada n -upla $\omega = (\omega_1, \omega_2, \dots, \omega_n)$ fica univocamente determinada por $Z_n(\omega)$. Portanto,

$$\begin{aligned} P[(\omega_1, \omega_2, \dots, \omega_n)] &= P[(X_1(\omega_1), X_2(\omega_2), \dots, X_n(\omega_n))] \\ &= P[\{X_1(\omega_1)\} \cap \{X_2(\omega_2)\} \cap \dots \cap \{X_n(\omega_n)\}], \end{aligned} \quad \text{I.37}$$

onde usamos a definição (I.30). Os eventos em intersecção no lado direito da última igualdade acima são independentes entre si, uma vez que o resultado do lançamento da moeda em um passo não influi na obtenção da informação do resultado do passo seguinte. Devemos notar que somente o *resultado acumulado depende dos resultados anteriores*. Com isso, a equação (I.37) torna-se

$$P[(\omega_1, \omega_2, \dots, \omega_n)] = P[X_1(\omega_1)]P[X_2(\omega_2)] \cdots P[X_n(\omega_n)]. \quad \text{I.38}$$

Tomemos, como exemplo, os pontos amostrais (H, H) e (T, H) em Ω^2 ; suas probabilidades de ocorrência são iguais a $P[(H, H)] = p^2$ e $P[(T, H)] = pq$, respectivamente.

Com este aparato desenvolvido até aqui estamos aptos a calcular a probabilidade de ocorrência de um determinado valor para S_n . Consideremos, por exemplo, que um jogador queira saber qual a probabilidade de se ganhar $\$ \kappa$ após n passos, ou seja, $P[\omega \in \Omega^n : S^n(\omega) = \kappa]$. Notemos que um dado ponto amostral $\omega \in \Omega^n$ pode levar a idênticos valores para S_n ; assim, por exemplo, $S_3(H, H, T)$ é igual a $S_3(H, T, H)$ e a $S_3(T, H, H)$, uma vez que a soma das variáveis aleatórias X_k correspondentes a cada ponto amostral $\omega_k \in \Omega$ é igual a 2. Portanto, devemos considerar estes valores iguais no cômputo de cada probabilidade. Supondo-se que na n -upla $\omega = (\omega_1, \omega_2, \dots, \omega_n)$ hajam κ caras, existem $n!/\kappa!(n-\kappa)!$ n -uplas $\omega \in \Omega^n$ diferentes que resultarão no mesmo valor para S_n ; deste modo, como a probabilidade de ocorrência de cada n -upla $\omega = (\omega_1, \omega_2, \dots, \omega_n)$ é igual a $p^\kappa q^{n-\kappa}$, segue-se que a probabilidade de ganharmos $\$ \kappa$ caras em n jogadas será igual a

$$P[\omega \in \Omega^n : S^n(\omega) = \kappa] = \binom{n}{\kappa} p^\kappa q^{n-\kappa}. \quad \text{I.39}$$

De acordo com a definição I.8, a expressão acima corresponde à função densidade de probabilidade de um passeio aleatório determinado pela variável S_n ; denominamos esta por *função densidade de probabilidade de Bernoulli*.

Calculemos agora a valor esperado e a variância deste processo. De acordo com a definição I.15, o valor esperado de S_n será igual a

$$\begin{aligned} ES_n &= \sum_{\omega \in \Omega^n} S_n(\omega) P[\omega \in \Omega^n : S^n(\omega) = \kappa] \\ &= \sum_{\kappa=0}^n \kappa P[\omega \in \Omega^n : S^n(\omega) = \kappa] \\ &= \sum_{\kappa=1}^n \frac{n!}{(\kappa-1)!(n-\kappa)!} p^\kappa q^{n-\kappa} \\ &= np, \end{aligned} \quad \text{I.40}$$

onde efetuamos algumas manipulações triviais com os somatórios.

A variância de um passeio aleatório pode ser determinada por meio das equações (I.32) e (I.40), resultando em

$$\text{var } S_n = npq. \quad \text{I.41}$$

Resta-nos agora obter o valor da probabilidade de ocorrência de uma cara, p . Experimentalmente, ao realizarmos uma série de n lançamentos de uma moeda, observamos que quanto maior for o valor de n mais próximo de 0.5 será o valor de p . Este fato experimental motivou Bernoulli a desenvolver o *primeiro teorema de limite*, o ‘*teorema áureo*’ (que posteriormente viria a ser chamado por Poisson de *Lei dos Grandes Números*),

que valida teoricamente o valor $p = 0.5$. Voltaremos a este tema na subseção I.2.6., onde trataremos deste assunto mais pormenorizadamente.

Por meio da equação (I.39) podemos finalmente responder a questão colocada no início desta subseção, assumindo-se então $p = 0.5$: um jogador que entre neste jogo de moedas terá uma probabilidade de 0.125 de ganhar \$3, uma probabilidade 0.375 de ganhar \$2, uma probabilidade 0.375 de ganhar \$1 e uma probabilidade 0.125 de não ganhar nada.

O passeio aleatório considerado até agora não admite valores negativos em sua formulação. Para tornarmos-lo mais completo, analisemos sua versão simétrica; como anteriormente, consideremos que o experimento-base consista em um jogo de moedas, com as seguintes regras: para cada cara obtida, ganha-se \$1; só que, desta vez, para cada coroa obtida, *perde-se* \$1. Perguntamos, então: quais são as probabilidades de perdas e de ganhos deste jogo ao final do terceiro lançamento?

Os espaços amostrais são idênticos em relação ao jogo anterior. Em termos de variáveis aleatórias, consideraremos

$$X_k(H) = +1, \quad X_k(T) = -1, \quad \text{I.42}$$

sendo a soma destas variáveis no n -ésimo passo igual à variável aleatória $S_n : \Omega^n \rightarrow \mathcal{R}$:

$$S_n(\omega) = \sum_{k=1}^n X_k(\omega_k), \quad \text{I.43}$$

onde $\omega = (\omega_1, \omega_2, \dots, \omega_n)$. Para fins de simplicidade, podemos separar a variável aleatória S_n em termos de suas partes positiva e negativa,

$$X_n^\pm(\omega) = \sum_{k=1}^n X_k^\pm(\omega_k), \quad \text{I.44a}$$

onde

$$X_k^+(\omega_k) = \begin{cases} +1, & \omega_k = H, \\ 0, & \omega_k = T \end{cases}, \quad X_k^-(\omega_k) = \begin{cases} -1, & \omega_k = T, \\ 0, & \omega_k = H \end{cases}. \quad \text{I.44b}$$

É fácil observar que as variáveis X_n^+ e X_n^- são *independentes*; ademais, apresentam a mesma *distribuição*, a de Bernoulli. O ganho obtido no n -ésimo passo é dado pela soma

$$S_n(\omega) = X_n^+(\omega) + X_n^-(\omega). \quad \text{I.45}$$

Se em n lançamentos obtivermos k caras, teremos $n - k$ coroas; com isso, S_n assume a expressão

$$S_n = 2k - n. \quad \text{I.46}$$

Suponhamos que ao final do n -ésimo passo tenhamos obtido \$ ξ ; calculemos agora probabilidade de se conseguir este valor, $P[\omega \in \Omega^n : S_n(\omega) = \xi]$. Usando a equação (I.46), é

fácil deduzir que o número de caras é igual a $(n+\xi)/2$ e o número de coroas é igual a $(n-\xi)/2$; como as variáveis aleatórias X^+_n e X^-_n apresentam a mesma distribuição, obtemos sua forma usando-se a expressão (I.39):

$$P\left[X^+_n = \frac{n+\xi}{2}\right] = P\left[X^-_n = \frac{n-\xi}{2}\right] = \binom{n}{\frac{n+\xi}{2}} p^{\frac{n+\xi}{2}} q^{\frac{n-\xi}{2}},$$

Portanto, a probabilidade de se obter um ganho (ou perda) de ξ é igual a

$$P[\omega \in \Omega^n : S^n(\omega) = \kappa] = \binom{n}{\frac{n+\xi}{2}} p^{\frac{n+\xi}{2}} q^{\frac{n-\xi}{2}}. \quad \text{I.47}$$

O passeio aleatório acima foi aplicado a um simples jogo de azar; entretanto, suas implicações em outros ramos do conhecimento são profundas. Em Física, por exemplo, podemos interpretar as variáveis aleatórias (I.42) como sendo deslocamentos unitários de uma partícula em um líquido; sua soma, equação (I.43), representaria o deslocamento parcial decorridos n passos. Este modelo discreto é a base do Movimento Browniano!

No caso de Finanças, cada passo (I.42) pode representar um movimento de baixa ou de alta ação; com isso, a modelagem de preços também pode resultar de um passeio aleatório!

No próximo capítulo voltaremos a abordar com maior detalhe os conceitos do passeio aleatório de modo a introduzir um importante conceito: o de *processos estocásticos*. Discutiremos, também, um outro importante limite do passeio aleatório: o limite de Poisson.

I.2 A Teoria Axiomática da Probabilidade

I.2.1 O Espaço de Probabilidade

Na seção anterior vimos que para caracterizarmos completamente um dado experimento necessitamos de três elementos: um espaço amostral, uma sigma-álgebra e uma medida de probabilidade. Para espaços arbitrários a definição destes elementos é análoga à introduzida na seção anterior:

I.18. Definição : Seja Ω um conjunto não-vazio, o qual denotaremos de *espaço amostral*. Uma *sigma-álgebra* é uma classe \mathcal{F} de subconjuntos (os *eventos*) de Ω que satisfazem as seguintes propriedades:

- I. $\Omega, \emptyset \in \mathcal{F}$;
- II. Se $A \in \mathcal{F}$, então $A^c \in \mathcal{F}$;
- III. Se $\{A_1, A_2, \dots\}$ é uma seqüência de eventos em \mathcal{F} , então

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}.$$

A dupla (Ω, \mathcal{F}) denomina-se *espaço mensurável*.

No caso de Ω ser finito, a definição acima coincide com o *power set* deste espaço amostral finito, conforme já discutido na proposição I.2.

Uma vez introduzido o conceito de sigma-álgebra em um espaço amostral abstrato, definamos agora a medida de probabilidade sobre este espaço:

I.19. Definição : Seja (Ω, \mathcal{F}) um espaço mensurável. Uma *medida de probabilidade* sobre este espaço é uma função

$$P: \mathcal{F} \longrightarrow [0,1],$$

de tal modo que

I. $P[\emptyset] = 0, \quad P[\Omega] = 1;$

II. Se $\{A_1, A_2, \dots\} \in \mathcal{F}$ é uma seqüência de eventos mutuamente exclusivos, então

$$P\left[\bigcup_{n=1}^{\infty} A_n\right] = \sum_{n=1}^{\infty} P[A_n].$$

A definição I.19 se reduz ao caso finito simplesmente tomando-se no item *ii* acima a seqüência $\{A_1, A_2, \dots, A_n, \dots\}$ com $A_{k \leq n} \neq \emptyset$ e $A_{k > n} = \emptyset$. Com isso obtemos novamente o item I. da Proposição I.4.

I.20 Definição : Seja Ω um espaço amostral não-vazio, \mathcal{F} uma sigma-álgebra em Ω e P uma medida sobre (Ω, \mathcal{F}) . A tripla (Ω, \mathcal{F}, P) é denominada *espaço de probabilidade*.

As propriedades fundamentais das medidas de probabilidade são análogas às da Proposição I.4., com exceção das duas últimas abaixo ([2], [4], [5], [9], [12], [15], [32]) :

I.21. Proposição : A medida P sobre um espaço mensurável (Ω, \mathcal{F}) satisfaz as seguintes propriedades:

I. Se $A, B \in \mathcal{F}$, então

$$P[A \cup B] = P[A] + P[B] - P[A \cap B];$$

II. Se $\{A_1, A_2, \dots\} \in \mathcal{F}$, então

$$P\left[\bigcup_{n=1}^{\infty} A_n\right] \geq \sum_{n=1}^{\infty} P[A_n];$$

III. Se $A, B \in \mathcal{F}$ e $A \subseteq B$, então

$$P[B] = P[A] + P[B/A], \quad \text{ou}$$

$$P[A] \leq P[B].$$

IV. (Continuidade por baixo) Se $\{A_1, A_2, \dots\}$ é uma seqüência de eventos em \mathcal{F} tal que $A_1 \subseteq A_2 \subseteq \dots$, então

$$P\left[\bigcup_{n=1}^{\infty} A_n\right] = \lim_{n \rightarrow \infty} P[A_n].$$

V. (Continuidade por cima) Se $\{A_1, A_2, \dots\}$ é uma seqüência de eventos em \mathcal{F} tal que $A_1 \supseteq A_2 \supseteq \dots$, então

$$P\left[\bigcap_{n=1}^{\infty} A_n\right] = \lim_{n \rightarrow \infty} P[A_n].$$

O primeiro exemplo de um espaço de probabilidade abstrato que poderíamos considerar é o caso da reta real, pois é o conjunto mais importante em termos de aplicações práticas. Mas, antes de prosseguirmos com este caso, necessitamos de ferramental adicional para tal digressão. O teorema abaixo é fundamental para a discussão que se segue ([2], [4], [5], [9], [12], [15], [32]):

I.22. Teorema : Seja \mathcal{G} qualquer classe não-vazia de eventos de Ω . Então existe uma e somente uma sigma-álgebra \mathcal{F} de Ω que satisfaça

I. $\mathcal{F} \supseteq \mathcal{G}$;

II. Se \mathcal{H} for qualquer sigma-álgebra em Ω tal que $\mathcal{H} \supseteq \mathcal{G}$, então $\mathcal{H} \supseteq \mathcal{F}$.

Neste caso, \mathcal{F} é a *menor* sigma-álgebra em Ω que contém todos os eventos de \mathcal{G} . Dizemos que \mathcal{F} é a sigma-álgebra *gerada* por \mathcal{G} e denotamo-la por $\sigma[\mathcal{G}]$.

O teorema acima nos diz que podemos gerar sigma-álgebras a partir de uma dada classe de eventos de um espaço amostral. Então, para extrapolarmos este teorema para o caso da reta real, basta considerarmos então $\Omega = \mathcal{R}$ e escolhermos uma classe de eventos de \mathcal{R} conveniente. Observemos que a partir da definição da reta real, $\mathcal{R} = (-\infty, +\infty)$, todo conjunto da forma

$$I(x) = (-\infty, x], \quad -\infty < x < \infty$$

pertence à \mathcal{R} ; assim, considerando-se a classe I de todos os intervalos da forma $I(x)$ acima, podemos observar que

$$\mathcal{R}, \emptyset \in I.$$

Com isso, o primeiro requisito para se construir uma sigma-álgebra é satisfeito. Isso nos indica que a classe dos intervalos semi-abertos I definida acima pode ser a classe desejada para então usarmos o Teorema I. 22. De fato, com estes elementos podemos provar o seguinte [15]:

I.23 Teorema : Seja I a classe composta por todos os intervalos semi-abertos da reta real $(-\infty, x]$ com $-\infty < x < \infty$; então, pelo Teorema I.22 existe uma mínima sigma-álgebra que contém a classe I ,

$$\mathcal{B}[\mathcal{R}] = \sigma[\mathcal{R}].$$

Com isso, introduzimos a seguinte

I.24. Definição : A sigma-álgebra deduzida no Teorema I.23 é denominada *sigma-álgebra de Borel*. Os eventos pertencentes à $\mathcal{B}[\mathcal{R}]$ recebem o nome de *eventos de Borel ou borelianos*.

Uma vez introduzida a sigma-álgebra de Borel, analisemos agora os eventos borelianos. Sendo $(-\infty, a] \in \mathcal{B}[\mathcal{R}]$, segue-se que também que $(a, \infty) = (-\infty, a]^c \in \mathcal{B}[\mathcal{R}]$; portanto, sendo $(-\infty, b]$ e $(a, \infty]$ borelianos, segue-se que sua intersecção

$$(a, b] = (a, \infty) \cap (-\infty, b] \quad \text{I.48}$$

também é um boreliano. Com isso, $(a, b] \in \mathcal{B}[\mathcal{R}]$. Intervalos semi-fechados à direita também são borelianos, pois $[a, b) = (a, b]^c$. Então, a partir da relação

$$[a, b) = (a, b] \cup [a, b), \quad \text{I.49}$$

concluimos que $[a, b) \in \mathcal{B}[\mathcal{R}]$. Da relação

$$(a, b) = [a, b)^c, \quad \text{I.50}$$

concluimos também que $(a, b) \in \mathcal{B}[\mathcal{R}]$.

Consideremos agora o evento unitário

$$\{a\} = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, a + \frac{1}{n} \right) \quad \text{I.51}$$

como a intersecção de intervalos abertos, que são borelianos, resulta em um evento boreliano, concluimos que também pontos (eventos unitários) são borelianos. Com isso, $\{a\} \in \mathcal{B}[\mathcal{R}]$. Por intermédio desta propriedade observamos também que todos os conjuntos que contém um número finito de elementos também são borelianos: se $A = \{a_1, a_2, \dots, a_n\}$, então

$$A = \bigcup_{n=1}^N \{a_n\}; \quad \text{I.52}$$

ademais, a propriedade acima pode ser estendida para um qualquer evento com um número contável de elementos: então, se $A = \{a_1, a_2, \dots\}$ segue-se que

$$A = \bigcup_{n=1}^{\infty} \{a_n\} \quad \text{I.53}$$

é um evento boreliano. Com isso, concluímos que o conjunto dos números racionais, \mathcal{Q} e seu complemento, o conjunto dos números irracionais I são também borelianos.

Pela discussão acima notamos claramente que a sigma-álgebra de Borel é a mais importante de todas, pois engloba todos os eventos que encontramos mais frequentemente: intervalos abertos, semi-abertos, fechados, semi-fechados, os números racionais, irracionais e assim por diante. Observemos que também o caso finito, objeto de nossa primeira seção, pertence à sigma-álgebra de Borel. No entanto, a sigma-álgebra de Borel não engloba *todos* os eventos de \mathcal{R} ; ou seja, $\mathcal{B}[\mathcal{R}] \subset \mathcal{P}[\mathcal{R}]$ e $\mathcal{B}[\mathcal{R}] \neq \mathcal{P}[\mathcal{R}]$. Exemplos de eventos que não são borelianos são os conjuntos de Cantor, que não discutiremos aqui.

I.2.2 Variáveis Aleatórias – Funções de Distribuição

Em nossa discussão anterior acerca de variáveis aleatórias em espaços finitos, vimos que sua principal característica era a de preservar o espaço de probabilidade. Para generalizarmos este conceito para um espaço abstrato, devemos contemplar esta importante propriedade. Para tanto, tomaremos como espaço alvo o espaço mensurável $(\mathcal{R}, \mathcal{B}[\mathcal{R}])$, que é, como vimos, a generalização do espaço finito de resultados mencionados na seção I.1. Tal generalização resulta no conceito de *funções mensuráveis em espaços de medida*:

I.26. Definição : Seja (Ω, \mathcal{F}, P) um espaço de probabilidade e $X : \Omega \longrightarrow \mathcal{R}$ uma função real. X é uma *variável aleatória* se for *mensurável* em relação à \mathcal{F} , ou seja, se satisfizer

$$X^{-1}(I) \subseteq \mathcal{F}$$

onde I é a classe definida no Teorema I.23.

A Definição I.26 admite outras formas, a saber ([2], [4], [5], [15]):

I.27. Proposição : As afirmações seguintes são equivalentes:

- I. X é mensurável;
- II. $X[(-\infty, x)] \in \mathcal{F}$;
- III. $X[(x, \infty)] \in \mathcal{F}$;

$$\text{IV. } X[[x, \infty)) \in \mathcal{F}.$$

O próximo teorema nos proporciona as condições necessárias e suficientes para a mensurabilidade de uma função:

I.28. Teorema : Seja (Ω, \mathcal{F}, P) um espaço de probabilidade e $X : \Omega \longrightarrow \mathcal{R}$. Então X é mensurável se e somente se $X^{-1}(B) \in \mathcal{F}$ para cada evento boreliano, ou seja

$$X^{-1}(\mathcal{B}) \subseteq \mathcal{F}.$$

As variáveis aleatórias apresentam inúmeras propriedades operacionais importantes([2], [4], [5], [15]):

I.29. Proposição : Sejam X e $Y : \Omega \longrightarrow \mathcal{R}$ variáveis aleatórias e α um número real. Então as seguintes expressões são também variáveis aleatórias:

- I. αX ;
- II. $\frac{1}{X}$ para $X \neq 0$;
- III. \sqrt{X} para $X \geq 0$;
- IV. $X + Y$;
- V. $X \circ Y$.

Conforme vimos na subseção I.1.3, podemos sempre aproximar variáveis aleatórias simples por funções indicadoras, analogamente ao que ocorre com vetores expandidos em uma dada base. Sabemos de vários ramos da Análise Matemática que é conveniente estabelecermos também representações em termos de bases para funções contínuas, pois, em muitos casos, a expressão de tal função é complexa a ponto de não nos permitir manipulá-las de uma forma adequada. Por isso, é importante que estabeleçamos também a expansão de variáveis aleatórias contínuas em termos de funções mais elementares. A primeira pergunta que nos surge é: combinando-se variáveis aleatórias obtemos uma variável aleatória? E o inverso, também é válido? Para um número finito de variáveis aleatórias vimos, pela proposição I.29., que sua soma resulta também em uma variável aleatória. E no caso de um quantidade enumerável de variáveis?

Para respondermos esta questão, consideremos uma dada seqüência $\{X_n(\omega)\}_{n=1}^{\infty}$ de variáveis aleatórias. Fixemos um ponto amostral ω qualquer em Ω ; tomemos agora as imagens $X(\omega)$ e $X_{m'}(\omega)$ deste ponto, onde m' é um dado número natural. Tracemos agora uma bola $B_{\varepsilon'}(\omega)$ ao redor de $X(\omega)$, de modo que esta contenha também o ponto $X_{m'}(\omega)$ (vide Figura I.1). Escolhamos agora um número natural m'' tal que $m'' > m'$; se $X_{m''}(\omega)$ não estiver contida na bola de raio ε' escolhamos uma nova bola de raio ε'' , centrada em $X(\omega)$, que contenha $X_{m''}(\omega)$. Procedamos deste modo sucessivamente.

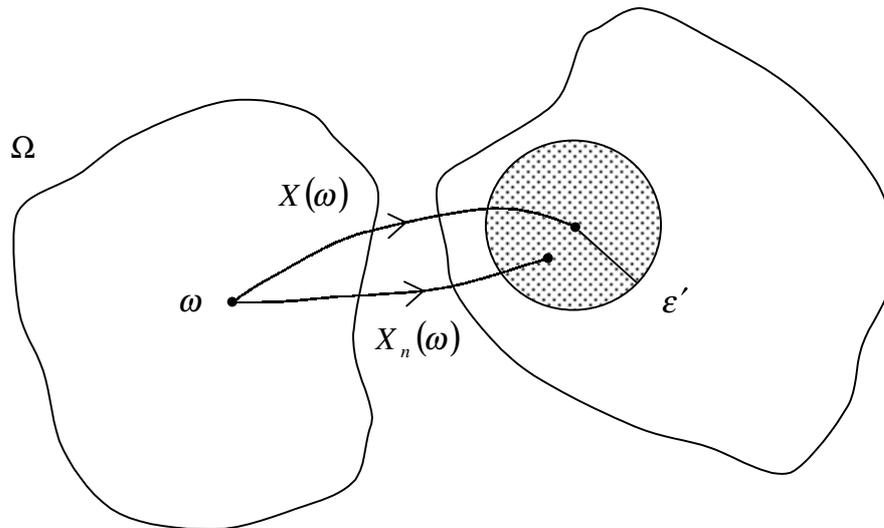


Figura I.2 – Diagrama para análise de convergência de variáveis aleatórias.

Se, após inúmeras escolhas, encontrarmos um número natural m de tal modo que para qualquer $n > m$ exista uma bola de raio ε , com ε infinitesimal, que contenha o ponto $X_n(\omega)$, vemos que $X_n(\omega)$ possui um valor numérico muito próximo do de $X(\omega)$; em outras palavras, dizemos que $X_n(\omega)$ converge para $X(\omega)$ no ponto ω fixado. Notemos que a convergência neste ponto depende de forma crucial da escolha do raio ε da bola, o que influencia, conseqüentemente, o valor de m .

Discutimos acima o conceito de convergência em um dado ponto; resta-nos estender este conceito para todos os pontos do espaço amostral Ω . Construamos de modo análogo para todos os valores de $\omega \in \Omega$, as bolas $B_\varepsilon(\omega)$ correspondentes. Notemos que, dentre a classe de conjuntos definida pelas bolas, há duas possibilidades: ou há uma bola de raio fixo ε que englobe todas as outras, ou cada seqüência definida por ω possui uma bola distinta dentro da qual se dá a convergência. No primeiro caso, há tão somente um número natural m a partir do qual ocorre a convergência; ou seja, para todo $n > m$. Nesta situação, m será determinado tão somente em função do raio da bola ε . Já no segundo caso, como há diferentes bolas que determinam a convergência de cada seqüência, m será definido também em função de cada valor de ω .

Temos aí definidos, portanto, dois modos de convergência diferentes, definidos da seguinte maneira:

I. 30. Definição : Seja $\{X_n(\omega)\}_{n=1}^{\infty} : \Omega \rightarrow \mathcal{R}$ uma seqüência de variáveis aleatórias. Dizemos que X_n converge uniformemente a um dado valor X se, para cada $\varepsilon > 0$ real e para cada $\omega \in \Omega$ existir um inteiro positivo $m = m(\varepsilon)$ tal que

$$|X_n(\omega) - X(\omega)| < \varepsilon, \quad \text{para } n > m.$$

I. 31. Definição : Seja $\{X_n(\omega)\}_{n=1}^{\infty} : \Omega \rightarrow \mathcal{R}$ uma seqüência de variáveis aleatórias. Dizemos que X_n converge a um dado valor X se, para cada $\varepsilon > 0$ real e para cada $\omega \in \Omega$ existir um inteiro positivo $m = m(\varepsilon, \omega)$ tal que

$$|X_n(\omega) - X(\omega)| < \varepsilon, \quad \text{para } n > m.$$

Neste caso, dizemos que X é o limite da seqüência $\{X_n(\omega)\}_{n=1}^{\infty}$, e o representamos por

$$\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)$$

para todo $\omega \in \Omega$.

Uma vez estabelecida esta definição, podemos agora responder à questão colocada anteriormente por meio do seguinte ([2], [4], [5], [15])

I. 32. Teorema : Seja X_1, X_2, \dots uma seqüência limitada de variáveis aleatórias para cada $\omega \in \Omega$. Então,

- I. $\sup X_n$ e $\inf X_n$ são variáveis aleatórias;
- II. $\limsup X_n$ e $\liminf X_n$ são variáveis aleatórias;
- III. Se $\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)$ para todo $\omega \in \Omega$, então X também é uma variável aleatória.

Uma vez discutidas as propriedades das variáveis aleatórias contínuas, voltemos nossa discussão à generalização do conceito de funções de distribuição. Na subseção I.1.3 vimos que a definição de uma função densidade de probabilidade é igual à expressão $p(x) = P[\omega : X(\omega) = x]$; entretanto, no caso contínuo, não podemos empregar esta definição, pois

$$P[\omega : X(\omega) = x] = 0$$

para todo $x \in \mathcal{R}$, conforme podemos facilmente verificar. Para podermos investigar a expressão matemática adequada das funções densidade de probabilidade para o caso contínuo, suponhamos que X seja uma variável aleatória real definida sobre um espaço de probabilidade arbitrário (Ω, \mathcal{F}, P) . De acordo com a proposição I.27, $X[(-\infty, x]] \in \mathcal{F}$; portanto,

$$X^{-1}[(-\infty, x]] = \{\omega \in \Omega : X(\omega) \leq x\}.$$

Deste modo, a composição $P \circ X^{-1}[(-\infty, x]]$ é uma função a valores reais em x ; é através desta que iremos generalizar o conceito de densidade de probabilidade para o caso contínuo, conforme veremos. Introduzamos então a seguinte

I. 33. Definição : Seja (Ω, \mathcal{F}, P) um espaço de probabilidades e X uma variável aleatória $X : \Omega \longrightarrow \mathcal{R}$; a função $F : \mathcal{R} \longrightarrow [0,1]$, definida por

$$F(x) := P \circ X^{-1} [(-\infty, x]],$$

é chamada de *função de distribuição da variável aleatória X* .

Pela definição acima inferimos imediatamente as seguintes propriedades das funções de distribuição ([2], [4]):

I. 34. Teorema : Seja $F(x)$ a função densidade de probabilidade de uma dada variável aleatória X ; então F será uma função não-decrescente em \mathcal{R} e contínua pela direita em toda a parte. Ademais, satisfaz as seguintes propriedades:

$$\lim_{x \rightarrow -\infty} F(x) = 0,$$

$$\lim_{x \rightarrow +\infty} F(x) = 1.$$

Vejamos agora como poderemos introduzir o conceito de função densidade de probabilidade no caso contínuo. Como $P[X = x] = 0$ para todo x , consideremos então de forma alternativa a probabilidade de ocorrência de um dado ponto em uma vizinhança $(x - \varepsilon, x + \varepsilon]$ de x , onde ε é um número real infinitesimal. Usando-se a relação

$$(x - \varepsilon, x + \varepsilon] = (-\infty, x + \varepsilon] \cap (-\infty, x - \varepsilon]^c,$$

e a propriedade das funções inversas

$$X^{-1} \{(x - \varepsilon, x + \varepsilon]\} = X^{-1} \{(-\infty, x + \varepsilon]\} \cap X^{-1} \{(-\infty, x - \varepsilon]\}^c,$$

segue-se que

$$\begin{aligned} P[X^{-1} \{(x - \varepsilon, x + \varepsilon]\}] &= P[X^{-1} \{(-\infty, x + \varepsilon]\} \cap X^{-1} \{(-\infty, x - \varepsilon]\}^c] \\ &= P[X^{-1} \{(-\infty, x + \varepsilon]\}] + P[X^{-1} \{(-\infty, x - \varepsilon]\}^c] \\ &\quad - P[X^{-1} \{(-\infty, x + \varepsilon]\} \cup [x - \varepsilon, \infty))\}] \\ &= F(x + \varepsilon) - F(x - \varepsilon). \end{aligned}$$

Portanto,

$$P[x - \varepsilon < X \leq x + \varepsilon] = F(x + \varepsilon) - F(x - \varepsilon).$$

Nesta dedução usamos a propriedade III. da proposição I.4, além da definição I.33.; expandindo-se o lado direito da expressão acima ao redor de ε , obtemos

$$P[x - \varepsilon < X \leq x + \varepsilon] = 2\varepsilon \frac{d}{dx} F(x),$$

a menos de termos de ordem $O(\varepsilon^3)$.

A equação assim obtida expressa a probabilidade de ocorrência de um determinado valor dentro de um intervalo infinitesimal, o que nos dá imediatamente a idéia de *densidade*; portanto, podemos associar a densidade de probabilidade de uma variável aleatória como sendo a derivada primeira de sua função de distribuição:

I. 35. Definição : Seja $F(x)$ a função densidade de probabilidade de uma dada variável aleatória X ; definimos a *função densidade de probabilidade* de X como sendo

$$f(x) := \frac{d}{dx} F(x).$$

Portanto,

$$F(x) = \int_{-\infty}^x f(z) dz,$$

e, pelo teorema I.34.,

$$\int_{-\infty}^{+\infty} f(x) dx = 1,$$

conforme esperamos de uma função densidade de probabilidade.

Seguem abaixo algumas funções densidade de probabilidade importantes do ponto de vista de aplicações:

✧ **Distribuição Unitária :**

$$f(x) = \frac{1}{b-a}, \quad a \leq x \leq b;$$

✧ **Distribuição de Poisson :**

$$f(x) = \frac{1}{x!} e^{-\lambda} \lambda^x, \quad \lambda > 0 \quad \text{e} \quad x \in \mathcal{N};$$

✧ **Distribuição de Cauchy :**

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \mathcal{R};$$

✧ **Distribuição Gama :**

$$f(x) = \frac{1}{\beta \Gamma(\alpha)} \left(\frac{x}{\beta} \right)^{\alpha-1} e^{-\frac{x}{\beta}}.$$

I.2.3 Momentos

Passemos agora à generalização do conceito de momentos para um espaço de probabilidades arbitrário. Partamos, para tanto, do conceito desenvolvido para o caso finito. Introduzimos na subseção I.1.5 a expressão do valor esperado para uma variável aleatória finita $\{X(\omega_k) = x_k\}$, $k \leq m$, como sendo igual a

$$EX = \sum_{k=1}^m x_k p(x_k),$$

onde $p(\omega_k)$ denota a probabilidade de ocorrência de cada x_k . Como $p(\omega_k) = P[X(\omega_k) = x_k]$, e P é uma medida de probabilidade, a equação acima nos sugere que, para um caso geral, o somatório acima deva ser substituído por uma integral. De imediato, podemos imaginar que esta integral deva ser de Stieltjes, em relação à medida P . Para uma dada função F , a integral de Stieltjes satisfaz dois preceitos básicos que aproveitaremos aqui em nossa generalização:

$$\int_{[a,b]} dF = F(b) - F(a);$$

$$\int_{[A,B]} \chi_{[a,b]} dF = F(b) - F(a), \quad [a,b] \subseteq [A,B].$$

Definiremos então, com base nesta analogia, uma integral que satisfaça os seguintes preceitos:

$$\int_{\Omega} dP = 1,$$

$$\int_{\Omega} \chi_A dP = P[A].$$

Estabelecidas então estas propriedades básicas, podemos investigar os casos mais complexos. Consideremos inicialmente o caso de uma variável aleatória simples,

$$X(\omega) = \sum_{n=1}^N x_n \chi_{A_n}(\omega).$$

Então,

$$\begin{aligned}\int_{\Omega} X dP &= \sum_{n=1}^N x_n \int_{\Omega} \chi_{A_n} dP \\ &= \sum_{n=1}^N x_n P[A_n].\end{aligned}$$

Portanto, se $A_n = \{\omega \in \Omega : X(\omega) = x_n\}$, recuperamos a expressão do valor esperado para uma variável aleatória discreta.

Consideremos agora uma variável aleatória arbitrária não-negativa X , que não possa se expressa por meio de variáveis aleatórias simples; introduzamos agora o conjunto $Z = \{\Phi : \Omega \rightarrow \mathcal{R} : \Phi \text{ simples e } \Phi(\omega) \leq X(\omega), \forall \omega \in \Omega\}$. Notemos que o conjunto Z contém todas as funções simples que aproximam X por baixo. Portanto, dentre estas funções, há certamente uma que melhor aproxima X . Portanto, nada mais natural que definirmos a integral de X como sendo a integral da variável aleatória simples que mais se aproxima desta; portanto, ajuntando-se as definições dadas acima, podemos definir integrais em um espaço de probabilidades da seguinte maneira:

I.36. Definição : Seja (Ω, \mathcal{F}, P) um espaço de probabilidades e X uma variável aleatória $X : \Omega \longrightarrow \mathcal{R}$, não-negativa. Considerando-se o conjunto

$$Z = \{\Phi : \Omega \rightarrow \mathcal{R} : \Phi \text{ simples e } \Phi(\omega) \leq X(\omega), \forall \omega \in \Omega\},$$

definimos a *integral de X* como sendo a expressão

$$\int_{\Omega} X dP = \sup_Z \int_{\Omega} \Phi dP.$$

Uma vez introduzida a expressão da integral de uma variável aleatória, podemos retomar o conceito de valor esperado de uma variável aleatória sobre um espaço de probabilidades arbitrário. A definição acima de integral aplica-se somente a variáveis aleatórias não-negativas; portanto, de modo a generalizarmos por completo o conceito, precisamos introduzir as integrais definidas para as variáveis que apresentem valores negativos. Para tanto, consideremos uma variável aleatória X arbitrária, podendo assumir tanto valores negativos quanto positivos. Separemos agora X em suas partes positiva e negativa:

$$X^+(\omega) = \max\{X(\omega), 0\}, \quad X^-(\omega) = \max\{-X(\omega), 0\}.$$

$X^-(\omega)$, apesar de representar a parte negativa de X , é uma variável positiva, pela definição. Portanto, podemos calcular sua integral por meio da definição I.36.. Para recuperarmos o fato de esta ser a integral da parte negativa de uma variável aleatória, basta multiplicarmos a integral obtida por (-1) . O resultado da integral da variável aleatória X é, com isso, igual a

$$\int_{\Omega} X dP = \int_{\Omega} X^+ dP - \int_{\Omega} X^- dP,$$

onde cada integral do lado direito da equação acima está definida de acordo com I.36.. Entretanto, a expressão acima somente fará sentido *se ambas* as integrais não forem infinitas ao mesmo tempo. Portanto, com esta observação, podemos introduzir então o valor esperado de uma variável aleatória arbitrária através da seguinte

I. 37. Definição : Seja (Ω, \mathcal{F}, P) um espaço de probabilidades e X uma variável aleatória $X : \Omega \longrightarrow \mathcal{R}$. Definimos o *valor esperado* da partes positiva e negativa de X como sendo

$$EX^+ = \int_{\Omega} X^+ dP, \quad EX^- = \int_{\Omega} X^- dP,$$

onde $X^+(\omega) = \max\{X(\omega), 0\}$ e $X^-(\omega) = \max\{-X(\omega), 0\}$. O *valor esperado* de X será dado pela expressão

$$EX = EX^+ - EX^-$$

se e somente se no mínimo um dos termos do lado direito da equação acima for *finito*.

O valor esperado de uma variável aleatória satisfaz as seguintes propriedades triviais ([2], [4], [5], [9], [12], [15], [32]):

I. 38. Proposição :

I. Seja X uma variável aleatória. Supondo ser EX finito e κ um número real, então

$$E[\kappa X] = \kappa EX;$$

II. Sejam X e Y variáveis aleatórias. Supondo-se que EX e EY são finitos, então vale

$$E[X + Y] = EX + EY;$$

III. Sejam X e Y duas variáveis aleatórias independentes. Então,

$$E[XY] = EX \cdot EY;$$

IV. Sejam X e Y variáveis aleatórias tais que $X \geq Y$ sobre Ω . Então

$$EX \geq EY;$$

V. Seja X uma variável aleatória. Então EX é finito se e somente se $E|X|$ for finito.

VI. Suponhamos que X seja limitada. Então EX é finito.

A definição do valor esperado de uma variável aleatória, conforme introduzida acima, traz um inconveniente: a complexidade de se lidar com as integrais definidas em relação à

medidas arbitrárias. É muito mais cômodo, indubitavelmente, trabalharmos com integrais de Riemann ou de Stieltjes. É possível, conforme veremos, expressar o valor esperado de uma dada variável aleatória em termos de uma integral de Stieltjes. A chave desta identificação reside no conceito das funções de distribuição, que são funções a valores reais que expressam a estrutura probabilística de um dado modelo.

Consideremos, para efetuamos a identificação mencionada, uma variável aleatória limitada: $Z = g(X)$, se $a < X \leq b$, e $Z = 0$, caso contrário. g é uma função a valores reais. Pela proposição I.37. V., sabemos que EZ é finito. Particionemos o intervalo $(a, b]$ em n subintervalos $I_k = a + k \frac{b-a}{n}$; sejam m_k e M_k os limites mínimo e máximo de $g(X)$ para cada intervalo I_k . Então, se $x_{k-1} < X \leq x_k$ para um dado ω , segue-se que

$$m_k \leq g(X) \leq M_k.$$

Integrando-se a expressão acima em relação à P no intervalo I_k , obtemos, usando-se $\int_{\Omega} \chi_A dP = P[A]$,

$$m_k P[x_{k-1} < X \leq x_k] \leq \int_{\Omega} \chi_{I_k} g(X) \leq M_k P[x_{k-1} < X \leq x_k].$$

Na subseção anterior vimos que $P[x_{k-1} < X \leq x_k] = F(x_k) - F(x_{k-1})$; portanto, da expressão acima obtemos imediatamente a relação

$$\sum_k m_k \{F(x_k) - F(x_{k-1})\} \leq \int_{\Omega} \chi_{(a,b]} g(X) \leq \sum_k M_k \{F(x_k) - F(x_{k-1})\},$$

onde usamos o fato que, em sendo $(a, b] = \coprod_k I_k$ (ou seja, a união disjunta de todos os conjuntos I_k), vale $\chi_{(a,b]} = \sum_k \chi_{I_k}$. Portanto,

$$s_n \leq EZ \leq S_n,$$

onde $s_n = \sum_k m_k \{F(x_k) - F(x_{k-1})\}$ e $S_n = \sum_k M_k \{F(x_k) - F(x_{k-1})\}$. O momento da variável aleatória Z é igual a $EZ = \int_{\Omega} \chi_{(a,b]} g(X)$ pois, pela definição de Z , esta é nula fora do intervalo $(a, b]$. Por um raciocínio análogo, é fácil de ver que

$$s_n \leq \sum_k g(x_k) \{F(x_k) - F(x_{k-1})\} \leq S_n.$$

Denominando-se $G_n = \sum_k g(x_k) \{F(x_k) - F(x_{k-1})\}$, segue-se das duas desigualdades que

$$|G_n - EZ| \leq S_n - s_n,$$

ou

$$|G_n - EZ| \leq \sum_k^n (M_k - m_k) \{F(x_k) - F(x_{k-1})\}.$$

Como

$$\lim_{n \rightarrow \infty} \sum_k^n g(x_k) \{F(x_k) - F(x_{k-1})\} = \int_a^b g dF,$$

e

$$\lim_{n \rightarrow \infty} \max\{M_k - m_k, k = 1, \dots, n\} = 0,$$

obtemos então

$$EZ = \int_a^b g dF.$$

A expressão acima pode ser facilmente generalizada para todo o conjunto dos reais através do seguinte ([2], [4], [10]):

I.39. Teorema : Seja X uma variável aleatória tendo como função de distribuição F , e seja $g: \mathcal{R} \rightarrow \mathcal{R}$ contínua em \mathcal{R} . Então $Eg(X)$ é finita se e somente se $\int_{-\infty}^{\infty} g dF$ é absolutamente convergente; neste caso,

$$Eg(X) = \int_{-\infty}^{\infty} g dF.$$

Por meio do teorema I.39. obtemos então a expressão do valor esperado de uma variável aleatória em termos de uma integral de Stieltjes. Podemos simplificar ainda mais esta expressão em uma integral de Riemann, usando-se para tanto a função densidade de probabilidade. Esta simplificação é permitida graças ao seguinte ([2], [4], [10]):

I. 40. Teorema : Suponhamos que $f: \mathcal{R} \rightarrow [0, \infty)$ seja Borel-mensurável. A variável aleatória X terá f como função densidade de probabilidade se e somente se

$$F(x) = \int_{-\infty}^x f(\xi) d\xi. \quad x \in \mathcal{R}.$$

É fácil mostrar que o valor esperado de X é igual a

$$EX = \int_{-\infty}^{\infty} x dF .$$

Para finalizar esta subseção, consideremos agora a generalização do conceito de variância de uma variável aleatória. Vimos pela definição I.16. que a variância de uma variável aleatória é dada pela expressão

$$\text{var } X = \sum_{k=1}^m (x_k - EX)^2 p(x_k) .$$

A generalização desta expressão é imediata para um espaço de probabilidades arbitrário; basta substituímos o somatório por uma integral, obtendo-se assim a expressão

$$\text{var } X = \int_{\Omega} (X - EX)^2 dP$$

ou

$$\text{var } X = E[X - EX]^2 .$$

Se EX^2 for finito, então

$$\text{var } X = EX^2 - [EX]^2 .$$

A covariância pode ser definida do mesmo modo que fizemos para o caso de espaços finitos:

$$\text{cov}(X, Y) = E[(X - EX)(Y - EY)] .$$

Podemos definir momentos de ordem superior, que desempenham um importante papel na teoria estatística:

$$EX^p = \int_{\Omega} X^p dP .$$

É através destes momentos de ordem superior que definimos, por exemplo, as medidas de assimetria e curtose de uma dada distribuição. Apesar de sua importância capital na análise estatística, não as discutiremos neste trabalho.

I.2.4 Convergência de Sequências de Variáveis Aleatórias

Retomemos agora a discussão iniciada na subseção I.2.2. a respeito da convergência de séries de variáveis aleatórias. Vimos que uma dada seqüência de variáveis aleatórias pode convergir ou em toda a parte, ou uniformemente; entretanto, encontramos vários exemplos não somente dentro da teoria da probabilidade, mas também na teoria da medida em geral, onde nem sempre a convergência ocorre em todos os pontos.

Consideremos, portanto, que uma dada seqüência de variáveis aleatórias, $\{X_n(\omega)\}_{n=1}^{\infty}$, convirja para uma variável aleatória X , exceto em determinados pontos amostrais; denominemos por A o conjunto de tais pontos. Se a probabilidade de ocorrência do evento A for nula, garantimos que a convergência desta seqüência ocorre com certeza em A^c , mesmo apesar desta falhar em A . Este é o caso, portanto, de um modo de convergência fraco, mas que desempenha um papel importante no escopo da teoria da probabilidade, conforme veremos.

Em toda esta subseção consideraremos que tanto uma seqüência $\{X_n(\omega)\}_{n=1}^{\infty}$, quanto X estão definidas sobre um espaço de probabilidade (Ω, \mathcal{F}, P) . Isto posto, introduzimos nossa primeira

I. 41. Definição : Seja $\{X_n(\omega)\}_{n=1}^{\infty}$ uma seqüência de variáveis aleatórias. Dizemos que X_n converge quase que certamente a um dado valor X , denominada por $X_n \xrightarrow{q.c.} X$, se existe um evento $A \in \mathcal{F}$ tal que $P[A] = 0$ e

$$X_n \rightarrow X \quad \text{em } A^c, \quad n \rightarrow \infty.$$

Os modos de convergência tratados nas definições I.30. e I.31. empregam diretamente as propriedades das seqüências de variáveis aleatórias e da variável-limite X , sem fazer menção à probabilidade de ocorrência destes. Para que seja válida a convergência da seqüência, é necessário que a probabilidade de que $|X_n(\omega) - X(\omega)| < \varepsilon$ seja não-nula para todo $\varepsilon > 0$, o que nos define um novo modo; portanto,

I. 42. Definição : Seja $\{X_n(\omega)\}_{n=1}^{\infty}$ uma seqüência de variáveis aleatórias. Dizemos que X_n converge em probabilidade a X , denotado por $X_n \xrightarrow{P} X$ se, para cada $\varepsilon > 0$ real valer a seguinte relação

$$P[|X_n - X| \geq \varepsilon] \rightarrow 0, \quad n \rightarrow \infty.$$

Suponhamos agora que seqüência de variáveis aleatórias $\{X_n(\omega)\}_{n=1}^{\infty}$ e X possuam os q primeiros momentos finitos, e que cada componente X_n tenha F_n e X tenha F como função de distribuição. Estes elementos formam as seqüências $\{EX_n^q(\omega)\}_{n=1}^{\infty}$ e $\{F_n(x)\}_{n=1}^{\infty}$; como tais, é interessante investigar suas propriedades de convergência, uma vez que são funções a valores reais. Uma vez analisadas tais propriedades, estudaremos suas conexões com os modos definidos anteriormente. Introduzamos primeiramente as definições propriamente ditas:

I. 43. Definição : Seja $\{X_n(\omega)\}_{n=1}^{\infty}$ uma seqüência de variáveis aleatórias, cujas respectivas funções de distribuição estão agrupadas na seqüência $\{F_n(x)\}_{n=1}^{\infty}$. Dizemos que X_n converge em distribuição a X , denotado por $X_n \xrightarrow{D} X$, se

$$F_n(x) \rightarrow F(x), \quad n \rightarrow \infty,$$

onde x é um ponto de continuidade de F .

I. 44. Definição : Seja $\{X_n(\omega)\}_{n=1}^{\infty}$ uma seqüência de variáveis aleatórias, com os q primeiros momentos finitos. Dizemos que X_n converge no q -ésimo momento a X , denotado por $X_n \xrightarrow{q} X$, se

$$E|X_n - X|^q \rightarrow 0, \quad n \rightarrow \infty.$$

Dizemos que X_n converge na média a X , tomando-se $q=1$, e que X_n converge na média quadrática a X , tomando-se $q=2$ e denotamos estes casos especiais de convergência por $X_n \xrightarrow{m} X$ e $X_n \xrightarrow{m.s.} X$, respectivamente.

Estes são, portanto, os modos de convergência mais importantes que podemos estabelecer para uma seqüência de variáveis aleatórias. Tais modos guardam uma restrita relação entre si, o que nos propiciará importantes ferramentas para futuros desenvolvimentos. O teoremas seguintes nos mostram as relações existentes entre estes diferentes modos; em todos estes estamos considerando que $\{X_n(\omega)\}_{n=1}^{\infty}$ é uma seqüência de variáveis aleatórias, bem como X , com os q primeiros momentos finitos. São válidas as seguintes afirmações ([2], [4], [5], [10]) :

I. 47. Teorema : Se $X_n \rightarrow X$ então $X_n \xrightarrow{q.c.} X$.

I. 46. Teorema : Se $X_n \xrightarrow{q.c.} X$ ou $X_n \xrightarrow{q} X$, então $X_n \xrightarrow{P} X$.

I. 48. Teorema : Se $X_n \xrightarrow{P} X$, então $X_n \xrightarrow{D} X$.

O inverso do teorema I.48 não é válido exceto no caso de X ser constante em Ω ([2], [4], [5], [10]) :

I. 49. Teorema : Se $X_n \xrightarrow{D} X$, onde $X(\omega) = \kappa$ para qualquer $\omega \in \Omega$, $\kappa \in \mathcal{R}$, então $X_n \xrightarrow{P} \kappa$.

Há outro teorema, por sinal muito importante, que trata das condições necessárias e suficientes para que uma dada seqüência convirja em distribuição ([2], [4], [5], [10]) :

I. 50. Teorema : Seja $\{X_n(\omega)\}_{n=1}^{\infty}$ uma seqüência de variáveis aleatórias. Então $X_n \xrightarrow{D} X$ se e somente se

$$Ef(X_n) \rightarrow Ef(X)$$

para toda função $f : \mathcal{R} \rightarrow \mathcal{R}$ contínua e limitada em \mathcal{R} .

No diagrama abaixo resumimos as relações existentes entre os diversos modos de convergência assegurados pelos teoremas acima:

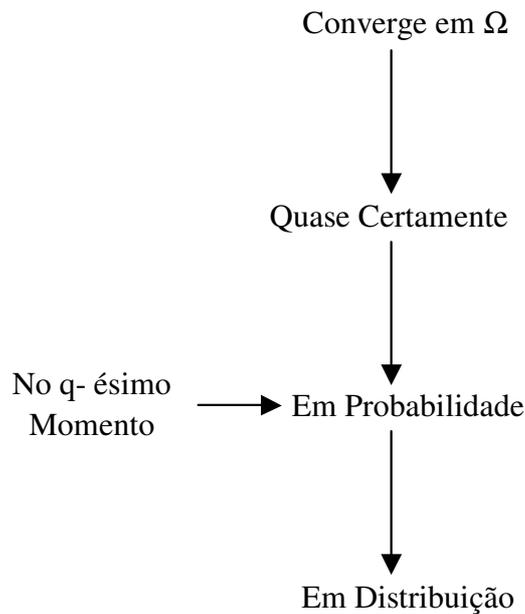


Figura I.3 – Relações existentes entre os diversos modos de convergência.

Estes são, portanto, os principais modos de convergência encontrados na teoria da probabilidade. Tais modos desempenharão um papel fundamental nos capítulos que seguem, principalmente na elaboração dos conceitos que levarão aos processos de Lévy e da integral estocástica de Itô.

Com este teorema, concluímos nossa discussão acerca dos modos de convergência de seqüências de variáveis aleatórias.

I.2.5 Funções Características

Depois de introduzidos os elementos mais importantes da teoria da probabilidade, podemos agora retomar o passeio aleatório desenvolvido na subseção I.1.6.. Discutimos nesta que o passeio aleatório constitui um modelo chave em toda a teoria da probabilidade, sendo até mesmo a base do Movimento Browniano e do Modelo de Black & Scholes, conforme veremos. O princípio fundamental do passeio aleatório consiste na introdução de uma variável aleatória S_N que é a soma de passos simples de magnitude 1 ou -1 . Isto significa dizer que estamos decompondo um dado experimento em porções mais simples, mais facilmente analisáveis. Portanto, a lógica por trás deste raciocínio é extremamente útil, pois nos permite representar um dado modelo complexo em termos de partes mais simples.

Nesta subseção estenderemos o princípio-base do passeio aleatório para um caso geral: ou seja, analisaremos como uma dada variável aleatória pode ser decomposta em partes mais simples. Isso compreende uma análise de como as funções de distribuição, momentos, etc, se comportam.

Observemos que em um passeio aleatório dois fatores estão implícitos: primeiramente, as variáveis que assumem valor ± 1 são *independentes* entre si; e, segundo, apresentam a mesma *distribuição*. Estes dois fatores são essenciais na análise que pretendemos fazer; portanto, formalizemo-lo na seguinte

I. 51. Definição : Seja $\{X_n(\omega)\}_{n=1}^N$ uma seqüência de variáveis aleatórias *independentes* entre si e que apresentem a mesma distribuição. Dizemos que tal seqüência é *independente e identicamente distribuída, iid*.

Consideraremos por toda esta subseção que as funções de distribuição de $\{X_n(\omega)\}_{n=1}^N$, definidas sobre um espaço de probabilidades (Ω, \mathcal{F}, P) , são dadas pela seqüência $\{F_n(x)\}_{n=1}^N$. Também fixamos a variável aleatória $S_N : \Omega^N \rightarrow \mathcal{R}$ como sendo a soma destas variáveis, dada por

$$S_N(\omega) = \sum_{k=1}^N X_k(\omega_k), \quad \text{I.54}$$

onde $\omega = (\omega_1, \omega_2, \dots, \omega_N)$ e $\omega_n \in \Omega$, $n \in \{1, 2, \dots, N\}$.

Por toda esta subseção investigaremos a decomposição (I.54), onde a seqüência $\{X_n(\omega)\}_{n=1}^N$ é *iid*. Antes de analisarmos as propriedades de convergência de S_N , verifiquemos as relações existentes entre $F_N(x)$, sua função de distribuição, e $\{F_n(x)\}_{n=1}^N$. Consideremos, por questões de simplicidade, $S = X_1 + X_2$; é possível demonstrar neste caso que

$$F(x) = \int_{-\infty}^{+\infty} dF_2(x_2) F_1(x - x_2).$$

Vemos então que a expressão de $F(x)$ em termos das funções de distribuição de X_1 e X_2 assume uma forma complexa, consistindo assim em uma dificuldade operacional. Esta dificuldade torna-se ainda maior em escolhendo-se $N > 2$. O uso de funções de distribuição para a análise da convergência de somas de variáveis *iid* mostra-se portanto inconveniente devido à esta complexidade. Precisamos desenvolver então uma nova ferramenta que nos permita analisar a convergência destas somas sem o uso direto das funções de distribuição. Em outras palavras, precisamos encontrar uma função ϕ de tal modo que, se ϕ estiver associada a S_N e $\{\phi_n\}_{n=1}^N$ à seqüência $\{X_n(\omega)\}_{n=1}^N$, a expressão matemática de ϕ seja simples, e que a convergência de $\{\phi_n\}_{n=1}^N$ em ϕ implique $F_n(x) \rightarrow F(x)$.

Para construirmos esta nova ferramenta utilizemos o teorema I.49., que nos diz que $X_n \xrightarrow{D} X$ se e somente se $Ef(X_n) \rightarrow Ef(X)$ para toda função $f : \mathcal{R} \rightarrow \mathcal{R}$ contínua e limitada em \mathcal{R} . Tomando-se como base a variável-soma $S = X_1 + X_2$ segue-se imediatamente que, em escolhendo-se $f(S) = e^S$ e $f(X_i) = e^{X_i}$, $i = 1, 2$, obtemos $Ef(S) = Ee^{X_1+X_2} = Ee^{X_1} Ee^{X_2} = Ef(X_1)Ef(X_2)$, uma vez que X_1 e X_2 são independentes. A escolha $f(S) = e^S$ não somente atende ao teorema I.50 como também torna a simples a relação entre a variável-soma e suas componentes. Será esta função que, a menos de algumas alterações em seu argumento, adotaremos em nossa análise que segue.

Introduzamos então a seguinte

I. 52. Definição : Seja X uma variável aleatória com função de distribuição F . Definimos a *função característica* de X como sendo a função a valores reais

$$\begin{aligned}\phi(\xi) &:= Ee^{i\xi X} \\ &= \int_{-\infty}^{+\infty} e^{i\xi x} dF.\end{aligned}$$

Como uma conseqüência do teorema I.50, temos o seguinte resultado ([2], [4], [10]):

I. 53. Teorema : Sejam $\{X_n(\omega)\}_{n=1}^N$ e variáveis aleatórias independentes dotadas de funções características $\{\phi_n(\xi)\}_{n=1}^N$. Se

$$X_n \xrightarrow{D} X,$$

então

$$\phi_n(\xi) \rightarrow \phi(\xi).$$

Apresentamos abaixo a expressão das funções características dos principais tipos de variáveis aleatórias:

⊗ **Distribuição Uniforme :**

$$\phi(\xi) = \frac{e^{i\xi} - 1}{i\xi}$$

⊗ **Distribuição de Bernoulli :**

$$\phi(\xi) = pe^{i\xi} + (1-p)e^{-i\xi}$$

⊗ **Distribuição Unitária :**

$$\phi(\xi) = \exp[i\xi x]$$

⊗ **Distribuição de Poisson :**

$$\phi(\xi) = \exp[\lambda(e^{i\xi} - 1)]$$

⊗ **Distribuição de Cauchy :**

$$\phi(\xi) = \exp[i\alpha\xi - \beta|x|]$$

✧ **Distribuição Gama :**

$$\phi(\xi) = \left[1 - i \frac{\xi}{\sigma} \right]^{-\lambda}.$$

Antes de discutirmos a questão da convergência de funções características, introduzamos algumas de suas importantes propriedades que serão úteis mais à frente em nossa discussão. A primeira delas, e a mais importante para nossas considerações iniciais, está diz respeito diretamente às somas de variáveis aleatórias ([2], [4], [10]) :

I. 54. Teorema : Sejam $\{X_n(\omega)\}_{n=1}^N$ e variáveis aleatórias independentes dotadas de funções características $\{\phi_n(\xi)\}_{n=1}^N$. Sendo S_N a soma destas variáveis, dotada de $\psi_N(\xi)$ como função característica, é válida a seguinte relação:

$$\psi_N(\xi) = \phi_1(\xi) \phi_2(\xi) \cdots \phi_N(\xi).$$

A propriedade acima decorre imediatamente do teorema I.50, e consiste no principal fator atrativo para o uso de funções características para a análise de convergência de somas de variáveis aleatórias independentes. Do teorema acima também depreendemos um importante corolário ([2], [4], [10]) :

I. 55. Corolário : Se a seqüência $\{X_n(\omega)\}_{n=1}^N$ for *iid*, dotada de funções características $\{\phi(\xi)\}_{n=1}^N$, então a função característica $\psi_N(\xi)$ da soma S_N será por sua vez igual a

$$\psi_N(\xi) = \phi(\xi)^N.$$

Pelo teorema I.54. e o corolário I.55. percebemos claramente a grande simplicidade trazida pelas funções características à teoria da probabilidade. Além desta vantagem, as funções características, conforme podemos facilmente provar, são contínuas, ao passo que as funções de distribuições podem ter um número finito de pontos de descontinuidade.

As funções características também são muito úteis no cômputo dos momentos de variáveis aleatórias ([2], [4], [10]) :

I. 56. Proposição : Sendo X uma variável aleatória com função característica $\phi(\xi)$ e tendo o momento EX^n finito, vale

$$EX^n = (-i)^n \left. \frac{\partial^n}{\partial \xi^n} \phi(\xi) \right|_{\xi=0}.$$

Nestas condições, $\phi(\xi)$ é também representável sob a forma

$$\phi(\xi) = \sum_{k=0}^n \frac{1}{k!} i^k \xi^k EX^k + g(\xi),$$

onde $\xi^{-n} g(\xi) \rightarrow 0$ quando $\xi \rightarrow 0$.

É plenamente possível invertermos a equação que define uma função característica em I.53. ([2], [4], [10]) :

I. 57. Teorema : Seja X uma variável aleatória com função característica $\phi(\xi)$ e função de distribuição F ; definamos inicialmente

$$\tilde{F}(x) := \frac{1}{2} \{F(x-) + F(x)\},$$

onde $x \in \mathcal{R}$. Então, para quaisquer números reais x e y vale

$$\tilde{F}(x) - \tilde{F}(y) = \lim_{\alpha \rightarrow \infty} \frac{1}{2\pi} \int_{-\alpha}^{\alpha} \frac{e^{-iy\xi} - e^{-ix\xi}}{i\xi} \phi(\xi) d\xi .$$

Esta equação é chamada de *fórmula de inversão*.

Decorre imediatamente deste teorema um importante resultado ([2], [4], [10]) :

I. 58. Teorema : Sejam X e Y variáveis aleatórias com a mesma função característica. Então X e Y também apresentarão a mesma função de distribuição. De modo inverso, se X e Y tem a mesma função de distribuição, então também apresentam a mesma função característica.

A partir destas propriedades, podemos finalmente discutir as propriedades de convergência das funções características, através dos seguintes teoremas ([2], [4], [10]) :

I. 59. Teorema : Seja $\{X_n(\omega)\}_{n=1}^N$ uma seqüência de variáveis aleatórias dotadas de funções características $\{\phi_n(\xi)\}_{n=1}^N$. Se $\phi_n(\xi)$ converge para um dado limite $\phi(\xi)$ para cada $\xi \in \mathcal{R}$, e supondo-se que ϕ seja contínua no ponto 0, então ϕ é uma função característica, e

$$X_n \xrightarrow{D} X .$$

Por meio dos teoremas I.53 e I.59 podemos provar o importantíssimo resultado ([2], [4], [10]) :

I. 60. Teorema : Seja $\{X_n(\omega)\}_{n=1}^N$ uma seqüência de variáveis aleatórias dotadas de funções características $\{\phi_n(\xi)\}_{n=1}^N$. Então,

$$X_n \xrightarrow{D} X ,$$

se e somente se

$$\phi_n(\xi) \rightarrow \phi(\xi) .$$

O teorema I.60 consiste no resultado mais importante da teoria das funções características. Conforme nos mostra o teorema I.58, vemos que funções características e funções de distribuição são absolutamente intercambiáveis entre si no que tange às propriedades de convergência de variáveis aleatórias. É extremamente vantajoso o uso das

funções características em vez das de distribuição, por ser sua forma, conforme vimos, muito mais simples do ponto de vista operacional em se tratando somas de variáveis aleatórias independentes. Com estes resultados em mãos, podemos analisar o teorema mais importante da teoria da probabilidade: o teorema central do limite.

I.2.6 Teoremas de Limite

O passeio aleatório, conforme discutimos na subseção I.1.6., foi estudado pela primeira vez por James Bernoulli que, analisando o clássico problema das probabilidades de ganhos em jogos de moedas, inferiu, através de seu ‘teorema áureo’, que a probabilidade de obtenção de k caras em n jogadas se aproxima de 0.5 para um valor grande de n . Em termos matemáticos, considerando-se a variável aleatória I.35, o teorema áureo de Bernoulli nos diz que

$$\lim_{N \rightarrow \infty} P \left[\left| \frac{1}{N} S_N - \frac{1}{2} \right| < \varepsilon \right] = 1,$$

para todo $\varepsilon > 0$. Em termos da notação empregada para a convergência em probabilidade, o teorema áureo de Bernoulli se torna

$$\frac{1}{N} S_N \xrightarrow{P} \frac{1}{2}. \quad \text{I.55}$$

Este teorema, cujo nome se popularizou nos anais probabilísticos como *Lei dos Grandes Números* após o trabalho de Simeon Denis Poisson, válida, do ponto de vista teórico, o conhecido resultado experimental: em lançando-se um grande número de vezes uma moeda a frequência relativa de ocorrência de uma cara é igual a 0.5.

Antes de prosseguirmos adiante façamos algumas considerações para o restante da subseção. Assumiremos uma seqüência $\{X_n(\omega)\}_{n=1}^N$ de variáveis aleatórias *iid* definidas sobre um espaço de probabilidade (Ω, \mathcal{F}, P) , e dotadas de uma função de distribuição F e função característica ϕ . A variável aleatória $S_N : \Omega^N \rightarrow \mathcal{R}$ consiste na soma das variáveis X_n ,

$$S_N(\omega) = \sum_{k=1}^N X_k(\omega_k),$$

onde $\omega = (\omega_1, \omega_2, \dots, \omega_N)$ e $\omega_n \in \Omega$, $n \in \{1, 2, \dots, N\}$. A função característica de S_N é dada por ϕ_N .

Feitas estas considerações, analisemos uma versão mais abrangente da Lei dos Grandes Números, que engloba o teorema áureo de Bernoulli como caso particular. Investigaremos as propriedades de convergência da variável aleatória S_N/N , conforme consta no próprio teorema áureo. A função característica de S_N/N é igual a $\phi_N(\xi/N)$, conforme podemos facilmente demonstrar. Já a função característica de cada parcela X_n/N de S_N/N é dada por $\phi(\xi/N)$, cuja forma aproximada pode ser estimada por meio da proposição I.56,

$$\phi\left(\frac{\xi}{N}\right) = 1 + \frac{i\xi}{N} \{\mu + \delta_N\},$$

onde consideramos $\mu = EX^n$ e $\delta_N \rightarrow 0$ quando $N \rightarrow \infty$. Pelo corolário I.55, sabemos que

$$\begin{aligned} \phi_N\left(\frac{\xi}{N}\right) &= \phi^N\left(\frac{\xi}{N}\right) \\ &= \left\{1 + \frac{i\xi}{N} \{\mu + \delta_N\}\right\}^N; \end{aligned}$$

portanto,

$$\lim_{N \rightarrow \infty} \phi_N\left(\frac{\xi}{N}\right) = e^{i\mu\xi}. \quad \text{I.56}$$

A função característica de S_N / N converge então para a função característica de uma variável aleatória constante $X(\omega) = \mu$, $\omega \in \Omega$. Por meio do teorema I.59, a relação (I.56) implica em

$$\frac{1}{N} S_N \xrightarrow{D} \mu,$$

o que, por sua vez, implica na relação

$$\frac{1}{N} S_N \xrightarrow{P} \mu,$$

onde usamos o teorema I.49. O teorema áureo de Bernoulli, dado pela equação I.55., consiste no caso particular da lei acima quando as variáveis aleatórias $\{X_n(\omega)\}_{n=1}^N$ assumem a forma $X_n(H) = +1$, $X_n(T) = 0$.

Demonstramos, na discussão acima, o importante ([2], [4], [9], [10], [32], [33])

I. 61. Teorema (Lei Fraca dos Grandes Números) : Seja $\{X_n(\omega)\}_{n=1}^N$ uma seqüência de variáveis aleatórias *iid* com valor esperado igual a μ . Sendo

$$S_N(\omega) = \sum_{k=1}^N X_k(\omega_k),$$

onde $\omega = (\omega_1, \omega_2, \dots, \omega_N)$ e $\omega_n \in \Omega$, $n \in \{1, 2, \dots, N\}$, é válida a relação

$$\frac{1}{N} S_N \xrightarrow{P} \mu.$$

Após introduzido a lei fraca dos grandes números conforme fizera Bernoulli (a demonstração que empregamos não fora a original adotada por ele), um passo importante que foi dado na teoria dos limites de probabilidades foi dado por Abraham de Moivre. Posteriormente, como complemento ao teorema desenvolvido por de Moivre, Pierre-Simon de Laplace consolidou o teorema que ficou historicamente conhecido por *teorema de de Moivre-Laplace*. Demonstraremos aqui este importante teorema, de modo a nos fornecer um importante *insight* na introdução posterior do teorema central do limite.

Consideremos, para tanto, a função densidade de probabilidade de um passeio aleatório (I.39)

$$P[\omega \in \Omega^n : S_n(\omega) = \kappa] = \binom{n}{\kappa} p^\kappa q^{n-\kappa}.$$

Estudemos o comportamento desta função quando tomamos o limite contínuo desta expressão, $n \rightarrow \infty$; para tanto, usemos a aproximação de Stirling

$$n! \sim \sqrt{2\pi n} n^{n+1/2} e^{-n},$$

que nos permite investigar o comportamento dos fatoriais para valores cada vez maiores de n . Inserindo-se a aproximação de Stirling na função densidade de probabilidade de Bernoulli, obtemos

$$P[S_n(\omega) = \kappa] = \left[\frac{n}{2\pi\kappa(n-\kappa)} \right]^{1/2} \left(\frac{np}{\kappa} \right)^\kappa \left(\frac{nq}{n-\kappa} \right)^{n-\kappa}. \quad \text{I.57}$$

Os dois últimos termos do lado direito da equação acima são iguais a 1 quando $\kappa = np$. Reescrevendo-se a equação (I.57) em termos da variável

$$\delta_\kappa = \kappa - np,$$

obtemos

$$P[S_n(\omega) = \kappa] = \left[\frac{n}{2\pi(np + \delta_\kappa)(nq - \delta_\kappa)} \right]^{1/2} \left(1 + \frac{\delta_\kappa}{np} \right)^{np + \delta_\kappa} \left(1 - \frac{\delta_\kappa}{nq} \right)^{nq - \delta_\kappa}. \quad \text{I.58}$$

Tomemos os logaritmos dos dois últimos termos do lado direito da equação (I.58), expandamo-la em uma série de Taylor, e agrupemos as potências de δ_κ :

$$\log \left[\left(1 + \frac{\delta_\kappa}{np} \right)^{np + \delta_\kappa} \left(1 - \frac{\delta_\kappa}{nq} \right)^{nq - \delta_\kappa} \right] = \frac{\delta_\kappa^2}{2n} \left(\frac{1}{p} + \frac{1}{q} \right) - \frac{\delta_\kappa^3}{6n^2} \left(\frac{1}{p^2} - \frac{1}{q^2} \right) + \dots \quad \text{I.59}$$

Assumindo-se

$$\lim_{\kappa \rightarrow \infty} \frac{\delta_\kappa^3}{n^2} = 0, \quad \text{I.60}$$

segue-se que a expansão (I.59) resulta na expressão

$$\left(1 + \frac{\delta_\kappa}{np}\right)^{np+\delta_\kappa} \left(1 - \frac{\delta_\kappa}{nq}\right)^{nq-\delta_\kappa} \sim \exp\left\{\frac{\delta_\kappa^2}{2npq}\right\}. \quad \text{I.61}$$

Por outro lado, em considerando-se a equação que define δ_κ dividida por n em ambos os lados,

$$\frac{\delta_\kappa}{n} = \frac{\kappa}{n} - p,$$

vemos que $|\delta_\kappa/n| \rightarrow 0$ pelo teorema áureo de Bernoulli. Deste modo, o denominador do primeiro termo da equação (I.58) fornece, assintoticamente, $(2\pi npq)^{-1/2}$. Portanto, a expressão (I.58) torna-se assintoticamente igual a

$$P[S_n(\omega) = \kappa] \sim \frac{1}{\sqrt{2\pi npq}} \exp\left\{-\frac{1}{2} \frac{\delta_\kappa^2}{npq}\right\}. \quad \text{I.62}$$

Introduzindo-se a nova variável

$$x_\kappa = h\delta_\kappa, \quad h = \frac{1}{\sqrt{npq}} \quad \text{I.63}$$

obtemos, no limite $n \rightarrow \infty$ e em assumindo-se a condição (I.54), a seguinte expressão

$$P[S_n(\omega) = \kappa] \sim \frac{h}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2} x_\kappa^2\right\}. \quad \text{I.64}$$

O lado direito da equação (I.64) corresponde, a menos da constante h , a uma função densidade de probabilidade gaussiana. Assim, vemos que, no limite $n \rightarrow \infty$ um passeio aleatório converte-se em um processo gaussiano!

Consideremos agora o limite obtido acima em termos de funções de distribuição. Em termos de um passeio aleatório,

$$\begin{aligned} P[S_n(\omega) \leq \kappa] &= \sum_{n=0}^{\kappa} P[S_n(\omega) = n] \\ &\sim hf(x_0) + hf(x_1) + \dots + hf(x_\kappa), \end{aligned} \quad \text{I.65}$$

onde

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\}.$$

O produto $hf(x_n)$ corresponde à área do retângulo de lado h sob a curva $y = f(x)$ entre os pontos $x_n - h/2$ e $x_n + h/2$; podemos então escrever para h pequeno

$$hf(x_n) \sim \int_{x_n - h/2}^{x_n + h/2} f(x) dx = F(x_{n+h/2}) - F(x_{n-h/2}). \quad \text{I.66}$$

Reescrevendo-se a equação (I.65) usando a prescrição (I.66) nos leva ao seguinte resultado:

$$P[S_n \leq \kappa] \sim F(x_\kappa) - F(x_0),$$

onde $h \rightarrow 0$. Tomando-se $x_0 \rightarrow -\infty$ e usando-se o teorema I.34, obtemos finalmente

$$P[S_n \leq \kappa] \sim F(x_\kappa). \quad \text{I.67}$$

O valor $x_\kappa = h(\kappa - Np)$ comparece como limite superior na integral do lado direito de (I.66); podemos absorver esta expressão no argumento da medida de probabilidade no lado esquerdo de (I.66):

$$P[S_n \leq x_\kappa \sqrt{npq} + np] \sim F(x_\kappa),$$

o que fornece

$$P\left[\frac{S_n - np}{\sqrt{npq}} \leq x_\kappa\right] \sim F(x_\kappa). \quad \text{I.68}$$

Provamos assim o importante teorema de limite denominado ([2], [4], [9], [10], [32], [33]):

I. 62. Teorema (de Moivre-Laplace): Seja $\{X_n(\omega)\}_{n=1}^N$ um passeio aleatório e

$$S_N(\omega) = \sum_{k=1}^N X_k(\omega_k),$$

onde $\omega = (\omega_1, \omega_2, \dots, \omega_N)$ e $\omega_n \in \Omega$, $n \in \{1, 2, \dots, N\}$. Seja ξ uma variável tal que

$$\xi = \frac{x - Np}{\sqrt{Npq}};$$

se ξ satisfizer a relação $h\xi^3 \rightarrow 0$ para $N \rightarrow \infty$, então

$$P\left[\frac{S_N(\omega) - Np}{\sqrt{Npq}} \leq \xi\right] \sim \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\xi} e^{-\frac{1}{2}x^2} dx.$$

Podemos estender imediatamente o teorema de de Moivre-Laplace para o caso de uma seqüência de variáveis aleatórias *iid*, conforme procedemos na extensão do teorema áureo de Bernoulli para a Lei Fraca dos Grandes Números. Para tanto, montemos a seqüência de variáveis aleatórias que convergirão a um dado limite. Tomemos como ponto de partida a variável aleatória que comparece no argumento da medida de probabilidade (I.68). Observemos que, em sendo a soma S_n composta pelas variáveis aleatórias (I.34), seus valores esperados e variâncias são dados por $EX_k = p$ e $\text{var } X_k = pq$ respectivamente; portanto, a variável que comparece no argumento de (I.68) é igual a

$$\frac{S_n - nEX}{\sqrt{n \text{ var } X}}, \quad \text{I.69}$$

onde X representa qualquer uma das variáveis X_k . A variável aleatória (I.69) será a base da nossa generalização. Definamos portanto

$$W_k = \frac{X_k - EX}{\sqrt{N \text{ var } X}}, \quad \text{I.70}$$

onde agora $\{X_n(\omega)\}_{n=1}^N$ é uma seqüência de variáveis aleatórias *iid*. É fácil demonstrar que as variáveis W_k são *iid* também; portanto, se $\phi(\xi)$ é a função característica de X_k , a função característica de W_k será dada por $\phi(\xi)e^{-i\mu\xi}$, onde $\mu = EX$. Já a função característica $\phi_N(\xi)$ da variável aleatória

$$S_N(\omega) = \sum_{k=1}^N W_k(\omega_k)$$

será então igual a $\phi(\xi)^N e^{-iN\mu\xi}$, conforme o corolário I.55. Expandindo-se esta função característica de acordo com a proposição I.56, obtemos

$$= \left\{ \left[1 - \frac{\xi^2}{2N} \{1 + \delta_N\} \right] e^{-i\mu\xi} \right\}^N,$$

onde $\delta_N \rightarrow 0$ quando $N \rightarrow \infty$. Tomando-se o limite, observamos que

$$\lim_{N \rightarrow \infty} \phi_N(\xi) = e^{-\frac{1}{2}\xi^2}.$$

Portanto, o teorema de de Moivre-Laplace é também válido para uma soma de variáveis aleatórias *iid* quaisquer. Demonstramos assim o teorema mais importante da teoria da probabilidade, o *teorema do limite central* ([2], [4], [9], [10], [32], [33]) :

I. 63. Teorema do Limite Central : Seja $\{X_n(\omega)\}_{n=1}^N$ uma seqüência de variáveis aleatórias *iid*. Para cada número real ξ é válida a seguinte relação:

$$P\left[\frac{\sum_{k=1}^N X_k - NEX}{\sqrt{N \text{ var } X}} \leq \xi\right] \sim \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\xi} e^{-\frac{1}{2}x^2} dx, \quad \text{I.71}$$

onde $EX = EX_k$ e $\text{var } X = \text{var } X_k$, $\forall k \in \mathcal{N}$.

CAPÍTULO II

A TEORIA DOS PROCESSOS ESTOCÁSTICOS

II.1 Conceitos Iniciais

No capítulo anterior discutimos um modelo extremamente importante dentro da Teoria da Probabilidade, o modelo binomial. Vimos que a partir dele obtemos em seu limite contínuo a distribuição normal, distribuição esta que forma a base da modelagem estatística de sistemas físicos. O modelo binomial é, em particular, muito rico para elucidar determinados conceitos matemáticos que, no caso contínuo, seriam de difícil indução. A Teoria dos Processos Estocásticos, conforme introduziremos neste capítulo, possui inúmeros conceitos que podemos induzir a partir de idéias simples de um modelo binomial. Assim como iniciamos a discussão da Teoria da Probabilidade no Capítulo I a partir de um exemplo, procederemos neste da mesma forma, através da discussão de um modelo muito usado na prática: *o modelo binomial da dinâmica de preços de ações* ([6], [22], [31]).

Consideremos que um dado investidor possua ações cujo preço inicial seja S_0 . Suponhamos que tal investidor deseje investir em um mercado onde no período seguinte o preço de suas ações ou cresça por um fator u ou decresça por um fator d . Assim, a evolução dos preços destas ações a cada período segue uma árvore binomial:

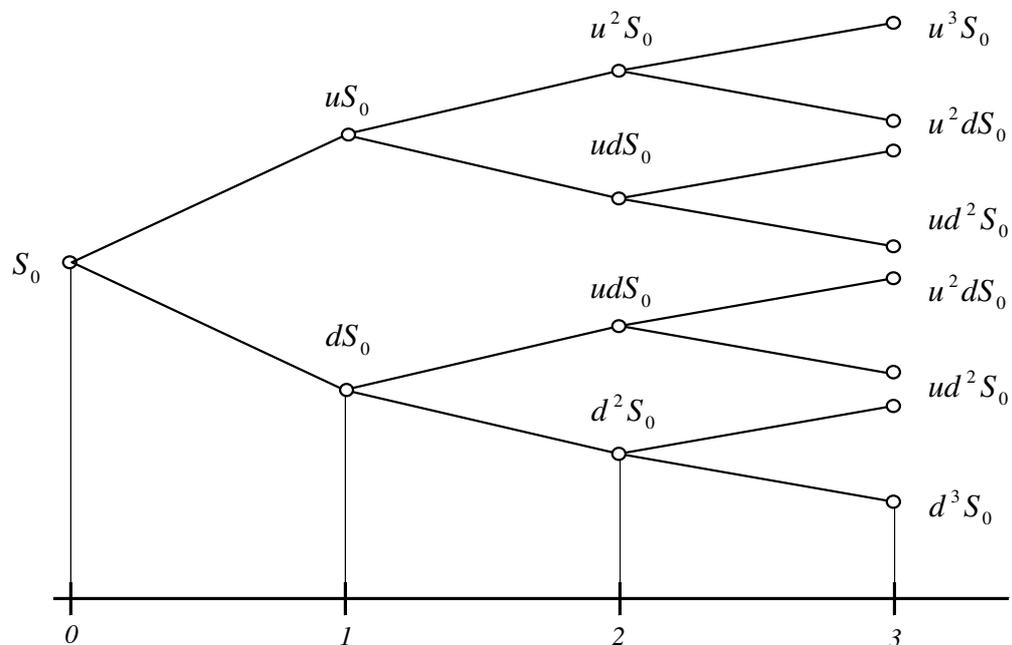


Fig. II.1 – Modelo Binomial para a Dinâmica de Preços

Notemos que neste modelo binomial não temos ainda definido um espaço amostral de fundo, mas tão-somente os valores das ações a cada período de tempo. É fácil definirmos um espaço amostral para este modelo: observemos que, como a cada passo o valor da ação ou

sobe, ou desce, podemos simplesmente associar um movimento de alta como a obtenção de uma ‘cara’ em um jogo de moedas, e um movimento de baixa como a obtenção de uma ‘coroa’ em um jogo de moedas. Ou seja, o fator aleatório que determina os movimentos de alta ou de baixa seria representado por um jogo de moedas. Com isso, o espaço amostral seria dado por

$$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\} \quad \text{II.1}$$

onde consideramos um jogo com apenas três lances. A sigma-álgebra associada a este espaço amostral é simplesmente dada por $\mathcal{A} = \mathcal{P}(\Omega)$.

Definamos agora a variável aleatória $X : \Omega \longrightarrow \mathcal{R}$ da seguinte forma:

$$X\left(\underbrace{HH\dots H}_m, \underbrace{TT\dots T}_n\right) = u^m d^n S_0 \quad \text{II.2}$$

onde fica claro que

$$(\Omega, \mathcal{A}) \xrightarrow{X} (\mathcal{R}, \mathcal{B}[\mathcal{R}]) \quad \text{II.3}$$

Por meio das equações (II.1 – 3) determinamos, assim, todas as informações concernentes à dinâmica de preços da árvore representada pela figura II.1. É a partir destas informações que iremos inferir os primeiros importantes conceitos da Teoria dos Processos Estocásticos.

Para tanto, façamos uma análise detalhada da estrutura de eventos deste modelo binomial. No primeiro passo a informação resultante poderia ser tão-somente ‘cara’ ou ‘coroa’; já na segunda, poderíamos ter ‘duas caras’, ‘cara na primeira jogada e coroa na segunda’, ‘coroa na primeira jogada e cara na segunda’ ou ‘duas coroas’. A partir destes eventos é fácil notar que ‘cara’ e ‘cara na primeira jogada e coroa na segunda’ fariam parte do evento ‘cara na primeira jogada’; e assim sucessivamente. Com isso, para o jogo de três lançamentos introduzamos os seguintes eventos:

$$A_H = \{HHH, HHT, HTH, HTT\} = \{H \text{ na primeira jogada}\}$$

$$A_T = \{TTT, TTH, THT, THH\} = \{T \text{ na primeira jogada}\}$$

$$A_{HH} = \{HHH, HHT\} = \{HH \text{ nas duas primeiras jogadas}\}$$

$$A_{HT} = \{HTH, HTT\} = \{HT \text{ nas duas primeiras jogadas}\}$$

$$A_{TH} = \{THH, THT\} = \{TH \text{ nas duas primeiras jogadas}\}$$

$$A_{TT} = \{TTT, TTH\} = \{TT \text{ nas duas primeiras jogadas}\}$$

II.4

A partir da estrutura de eventos acima, as sigma-álgebras subjacentes são dadas por:

$$\mathcal{A}_0 = \{\emptyset, \Omega\},$$

$$\begin{aligned}\mathcal{A}_1 &= \{\emptyset, \Omega, A_H, A_T\}, \\ \mathcal{A}_2 &= \{\emptyset, \Omega, A_H, A_T, A_{HH}, A_{HT}, A_{TH}, A_{TT}, A_{HH} \cup A_{TH}, A_{HH} \cup A_{TT}, \\ &A_{HT} \cup A_{TH}, A_{HT} \cup A_{TT}, A_{HH}^c, A_{HT}^c, A_{TH}^c, A_{TT}^c\}, \\ \mathcal{A}_3 &= \mathcal{A};\end{aligned}\tag{II.5}$$

A estrutura acima nos permite separar os eventos de forma a contemplar todos os passos deste modelo binomial. Por exemplo, se soubermos que o resultado após o lançamento de três moedas pertence à \mathcal{A}_1 , e que, dentre os eventos possíveis o resultado se encontra em A_H , sabemos *no máximo* que o primeiro lançamento resultou em uma ‘cara’. Assim, a sigma-álgebra \mathcal{A}_1 nos fornece a informação do experimento *até a primeira jogada*. Da mesma maneira, a sigma-álgebra \mathcal{A}_2 nos fornece a informação *até a segunda jogada* e \mathcal{A}_3 nos fornece a *informação total*. A sigma-álgebra \mathcal{A}_0 , ou sigma-álgebra *trivial* não nos fornece informação alguma.

Com isso, vemos que a informação concernente à este jogo de moedas pode ser fragmentada em informações associadas aos passos imediatamente anteriores. Em outras palavras, a sigma-álgebra do último passo em si pode ser decomposta em sub-sigma-álgebras referente aos passos imediatamente anteriores. Assim, para sabermos os resultados possíveis associados a um determinado passo basta simplesmente nos referirmos a sigma-álgebra associada a este passo.

Retornando ao movimento de ações objeto de nossa discussão, sabemos da figura II.1 que a cada passo da árvore corresponde uma variável aleatória, e em função das sub-sigma-álgebras discutidas acima, vemos que é possível introduzir variáveis aleatórias também associadas à cada passo. Para isso, basta introduzirmos mais um índice na expressão da variável aleatória, que representará o passo; portanto, se X_t for a variável aleatória associada ao *t-ésimo* passo, segue-se então que

$$X_t : \Omega \times T \rightarrow \mathcal{R}_X^t \tag{II.6}$$

onde $T = \{0,1,2,3\}$ e \mathcal{R}_X^t corresponde ao espaço dos resultados no passo t . Os valores de X_t são facilmente calculados passo a passo: para o último vale a seguinte expressão

$$X_t(\omega_1 \omega_2 \cdots \omega_t) = u^m d^n S_0, \tag{II.6a}$$

onde $\omega_k \in \{H, T\}$ (com $k \leq t$) e $m = \#(H)$, $n = \#(T)$ tal que $m + n = t$. Para o penúltimo passo vale

$$\begin{aligned}X_{t-1}(\omega_1 \omega_2 \cdots \omega_t) &= X_{t-1}(\omega_1 \omega_2 \cdots \omega_{t-1}) \\ &= u^{m^1} d^{n^1} S_0\end{aligned}\tag{II.6b}$$

onde novamente $\omega_k \in \{H, T\}$, com $k \leq t-1$ e $m^1 = \#(H)$, $n^1 = \#(T)$ tal que $m^1 + n^1 = t-1$. Para o $t-s$ -ésimo passo é válida a equação

$$\begin{aligned}
X_{t-s}(\omega_1 \omega_2 \cdots \omega_{t-s} \cdots \omega_t) &= X_{t-s}(\omega_1 \omega_2 \cdots \omega_{t-s}) \\
&= u^{m^s} d^{n^s} S_0
\end{aligned}
\tag{II.6c}$$

onde mais uma vez $\omega_k \in \{H, T\}$, com $k \leq t-s$ e $m^s = \#(H)$, $n^s = \#(T)$ tal que $m^s + n^s = t-s$.

A partir das equações (II.6) e (II.6 a – c), obtemos todos os valores dos processos estocásticos X_t para um passo arbitrário t , válido para um número de jogadas qualquer. Ainda a partir destas equações é fácil de observar que a sigma-álgebra gerada por X_t , $\sigma[X_t]$, é dada por

$$\sigma[X_t] = \mathcal{A}_t \tag{II.7}$$

onde \mathcal{A}_t é, conforme vimos acima, a sigma-álgebra associada ao t -ésimo passo. Assim, podemos considerar que a variável aleatória X_t *carrega* a informação associada a cada passo. Formalizando-se esta afirmação obtemos a seguinte

II.1. Definição : Seja X uma variável aleatória sobre (Ω, \mathcal{A}) ; dizemos que um evento $A \subset \Omega$ é *determinado pela variável aleatória* X se, em sabendo-se somente o valor de $X(\omega)$, podemos decidir se $\omega \in A$ ou não. Isto é, para todo $x \in \mathcal{R}$ ou $X^{-1}(x) \in A$ ou $\{X^{-1}(x)\} \cap A = \emptyset$. A classe dos eventos de Ω determinados por X é uma sigma-álgebra, denotada *sigma-álgebra gerada por* X , e representada por $\sigma[X]$.

Uma vez abordado este importante tema acerca da informação propiciada por variáveis aleatórias, passemos à discussão de uma questão muito importante quando se pretende abordar sistemas físicos: o *tempo*. Pela estrutura do exemplo que consideramos até aqui, somos tentados a tomar o parâmetro que denominamos de *passo* como sendo o tempo. Isso é plenamente justificável, pois afinal cada passo representa um instante onde o preço de cada ação ou aumenta ou diminui. No entanto, para realmente podermos tomar estas variáveis aleatórias também como função do tempo, faz-se necessária a introdução de algum ordenamento para estas novas variáveis transmitam a noção de evolução temporal. De fato, conforme veremos em seqüência, é possível ordenarmos uma seqüência de variáveis aleatórias e conferirmos à estas a idéia de evolução temporal.

Com isso, ao introduzir um novo parâmetro e interpretá-lo como o tempo, podemos efetivamente analisar os sistemas físicos, introduzindo o conceito de aleatoriedade no processo de evolução temporal dos sistemas dinâmicos e estudá-los assim do ponto de vista da teoria da Probabilidade. Isso representa um grande avanço conceitual, a ponto de a teoria dos Processos Estocásticos se tornar fundamental como instrumento de análise da teoria da Complexidade, do Caos, e assim por diante.

Introduzamos primeiramente, antes de abordarmos mais detalhadamente a estrutura temporal das funções X_t , o primeiro importante conceito que nos ocorre à luz de nossa discussão até o momento:

II.2. Definição : Seja T um conjunto e (Ξ, \mathcal{K}) um espaço de probabilidade mensurável. Um *processo estocástico indexado por T* consiste em uma família de variáveis aleatórias $\{X_t, t \in T\}$, tal que

$$X_t : \Omega \times T \longrightarrow \Xi$$

O espaço (Ξ, \mathcal{K}) é denominado *espaço de estados*.

A definição acima é a mais abrangente possível, pois se relaciona a espaços de estados gerais e conjuntos indexadores T arbitrários. Para interpretarmos o conjunto indexador T como o tempo, faz-se necessária a introdução do ordenamento que discutimos um pouco acima. Façamos, para tanto, uma análise das sigma-álgebras II.5; vimos que cada sigma-álgebra \mathcal{A}_t representa uma informação associada *até* o passo t . Deste modo, a informação do passo seguinte *engloba* a informação do passo anterior, o que resulta na relação geral $\mathcal{A}_t \subset \mathcal{A}_{t+1}$. Portanto, as sigma-álgebras indexadas pelo conjunto $T = \{0,1,2,3\}$ apresentam uma *relação de ordem* entre si, e é exatamente tal relação que corresponde ao desejado ordenamento mencionado acima. Assim, sabemos que a variável aleatória X_t *precede* X_{t+1} pois a informação que X_{t+1} carrega engloba a informação que X_t por sua vez carrega. Portanto, é exatamente em função desta seqüência de sub-sigma-álgebras é que o desejado ordenamento se faz presente no contexto da teoria dos Processos Estocásticos. Isso, porque podemos dizer que a sigma-álgebra \mathcal{A}_t é a classe de todos os eventos que representam *situações passadas ou presentes* no instante t .

Assumiremos em toda a nossa discussão que o conjunto indexador T será igual a um conjunto discreto ou a $[0, \infty)$ representando assim o parâmetro temporal. Tomaremos também em nossa discussão que $\Xi = \mathcal{R}$ ou $\Xi = \mathcal{R}^D$ e que $\mathcal{K} = \mathcal{B}[\mathcal{R}]$.

O ordenamento introduzido está formalizado através da seguinte

II.3. Definição : Uma *filtração* sobre um espaço mensurável (Ω, \mathcal{A}) , consiste em uma seqüência crescente $\{\mathcal{A}_t\}_{t \geq 0}$ de sub-sigma-álgebras de \mathcal{A} . Em outras palavras, para cada t temos uma sub-sigma-álgebra \mathcal{A}_t tal que $\mathcal{A}_t \subset \mathcal{A}_s$ para $s < t$. Um espaço mensurável (Ω, \mathcal{A}) equipado com uma filtração $\{\mathcal{A}_t\}_{t \geq 0}$ é denominado *espaço filtrado*.

Uma filtração que decorre imediatamente da definição acima é a denominada *filtração natural*, dada por

$$\mathcal{A}_t^0 := \sigma[X_s : s \leq t] \quad \text{II.8}$$

Antes de introduzirmos o próximo conceito façamos um pequeno resumo do que foi discutido até aqui. Construímos o conceito de filtrações a partir de um exemplo discreto, o modelo binomial da dinâmica de preços. Pela árvore binomial, pudemos visualizar como cada classe de eventos poderia ser construída a cada passo. Estas sigma-álgebras eram geradas pelos processos X_t , sendo, portanto, mensuráveis em relação a estes. A mensurabilidade de uma família de funções indexadas por um parâmetro t sobre uma filtração é um critério importante para qualificar esta como um processo estocástico, uma vez que assim podemos

interpretar tal parâmetro como o tempo. Desta forma, introduzimos um conceito que será muito importante para as discussões que seguem:

II.4. Definição : Um processo estocástico X sobre (Ω, \mathcal{A}) é chamado de *adaptado à filtração* $\{\mathcal{A}_t\}_{t \geq 0}$ se X_t for mensurável para cada $t \in [0, \infty)$.

É evidente que qualquer processo estocástico X é adaptado à sua filtração natural e também é fácil de notar que $\{\mathcal{A}_t^0\}_{t \geq 0}$ é a filtração mínima para a qual X é adaptado.

Os conceitos de filtração e de adaptação consistem, portanto, nos dois conceitos que permitem, como vimos, a inclusão do fator tempo dentro da teoria dos Processos Estocásticos. Vamos exemplificar agora como os processos $X_t(\omega)$ evoluem em nosso exemplo da dinâmica binomial de preços. Tomemos na Figura II.1, por exemplo, o resultado $u^2 dS_0$; notemos que o preço da ação, de acordo com a figura, pode evoluir a partir do instante 0 de três formas diferentes (vide Figura II.2): ascensão – ascensão – queda (linha grossa), ascensão – queda – ascensão (linha tracejada fina) ou então queda – ascensão – ascensão (linha tracejada grossa). Cada uma destas linhas representa uma *trajetória* possível para denotar a evolução do preço da ação. Notemos que nos três casos o ponto amostral pode ser tomado como *constante*: na primeira trajetória o ponto amostral $\omega_1 = HHT$ corresponde exatamente, de acordo com as equações (II.6 a – c), aos valores intermediários do preço da ação para se atingir o resultado $u^2 dS_0$: uS_0 e $u^2 S_0$. Do mesmo modo os pontos amostrais $\omega_2 = HTH$ e $\omega_3 = THH$ correspondem exatamente aos resultados intermediários para a formação do preço $u^2 dS_0$: uS_0 e udS_0 no caso de $\omega_2 = HTH$, e dS_0 e duS_0 no caso de $\omega_3 = THH$.

Vemos pelo exemplo acima que, no caso específico do modelo binomial, para chegarmos a um dado resultado no instante t podemos *percorrer algumas diferentes trajetórias*, cada uma com uma certa probabilidade associada. Cada trajetória é determinada em fixando-se um dado ponto amostral e em fazendo-se o parâmetro temporal variar. A figura abaixo ilustra claramente este fato:

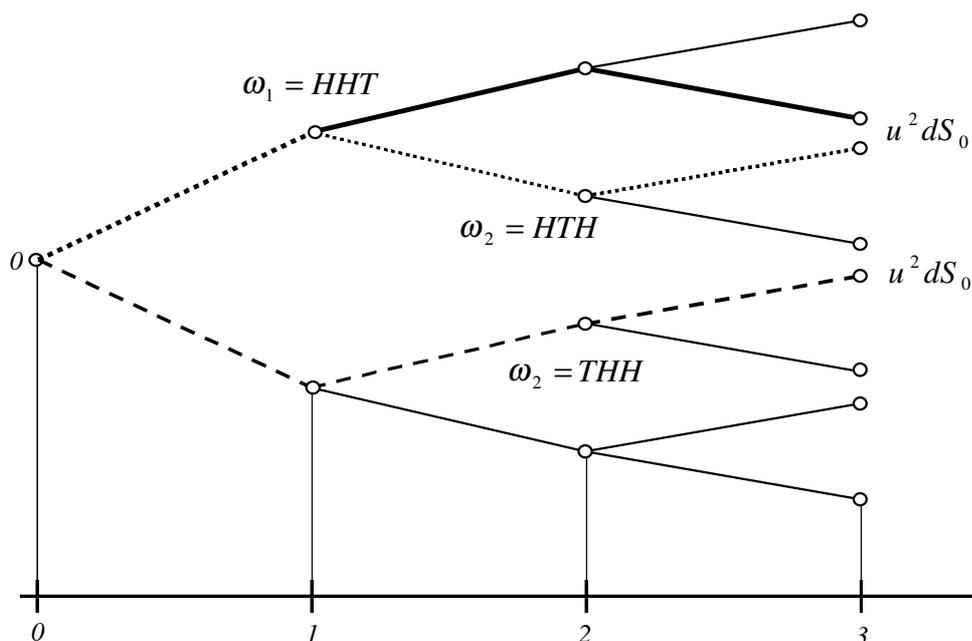


Figura II.2 – Possíveis evoluções temporais para se obter o preço $u^2 dS_0$

Com isso, as três trajetórias possíveis associadas ao resultado $u^2 dS_0$ são representadas pelo conjunto $\mathcal{X}_3(\omega_k) = \{X_t(\omega_k) : X_3(\omega_k) = u^2 dS_0, t \leq 3\}$, onde $k \in \{1,2,3\}$ e $\{\omega_1, \omega_2, \omega_3\} = \{HHT, HTH, THH\}$.

Portanto, de acordo com o exemplo apresentado acima, vemos que, em fixando-se os pontos amostrais $\omega = \tilde{\omega}$ obtemos todas as trajetórias possíveis para se alcançar um dado resultado no tempo a partir de um instante inicial. Deste modo, o processo

$$X_\bullet(\tilde{\omega}) : T \longrightarrow \mathcal{R} \quad \text{II.9}$$

é uma função ordinária cuja variável é o tempo.

Poderíamos também, ao invés de fixar os pontos amostrais ω , fixar o parâmetro temporal t ; com isso, em tomando-se $t = \tau = \text{constante}$ vemos que

$$X_\tau(\bullet) : \Omega \longrightarrow \mathcal{R}_X^\tau \quad \text{II.10}$$

o que demonstra o fato de as funções $X_\tau(\bullet)$ serem simplesmente variáveis aleatórias ordinárias. Retomamos assim o conceito de variável aleatória desvinculada do fator tempo, conforme vimos na teoria da Probabilidade.

À luz do exposto acima, é conveniente introduzirmos a seguinte definição, que oportunamente será de grande valia:

II. 5 Definição : Um processo estocástico X que toma valores em $(\mathcal{R}, \mathcal{B}[\mathcal{R}])$ é denominado *contínuo quase em toda a parte* se o mapeamento $t \mapsto X_t$ for contínuo para quase todo $\omega \in \Omega$.

Para finalizar nosso exemplo de processo estocástico discreto, observemos agora sua distribuição de probabilidades. Por ser um passeio aleatório, segue-se que a probabilidade de ocorrência do valor $X_t / S_0 = u^s d^{t-s}$ é dada pela equação

$$w_t(s) = \binom{t}{s} p^s q^{t-s}. \quad \text{II.11a}$$

Com isso, a distribuição de probabilidades associada ao modelo binomial da dinâmica de preços é dada por

$$P \left[\frac{X_t}{S_0} \leq u^m d^{n-m} \right] = \sum_{k=0}^{n+m} w_n(k). \quad \text{II.11b}$$

Os dois primeiros momentos do processo

$$Z := \frac{X_t}{S_0} = u^m d^{n-m} \quad \text{II.12a}$$

são dados pelas equações

$$\begin{aligned}
E[\ln Z] &= E[\ln(u^m d^{n-m})] \\
&= np \ln \frac{u}{d} + n \ln d
\end{aligned} \tag{II.12b}$$

$$\text{var}[\ln Z] = npq \left[\left(\ln \frac{u}{d} \right) \right]^2. \tag{II.12c}$$

Veremos, na seção seguinte, que o Teorema de De Moivre-Laplace também se aplica neste caso, bastando para tanto assumir algumas simples condições.

O passeio aleatório típico, conforme vimos no capítulo anterior, também é um exemplo de processo estocástico: tomando-se o processo *iid* $\{X_t, t \in \mathcal{N}\}$, tal que $X_t(\omega) \in \{-1, 1\}$, $\forall \omega \in \Omega$, segue-se que as somas parciais

$$S_t = \sum_{s=1}^t X_s \tag{II.13}$$

consistem em um processo estocástico seguindo uma distribuição de Bernoulli:

$$P[S_t = \xi] = \binom{t}{\frac{t+\xi}{2}} p^{\frac{1}{2}(t+\xi)} (1-p)^{\frac{1}{2}(t-\xi)} \tag{II.14}$$

onde as probabilidades p e $1-p$ são dadas pelas equações

$$P[X_t = -1] = 1-p \quad \text{e} \quad P[X_t = 1] = p \tag{II.15}$$

Com isso, caracterizamos completamente o processo do passeio aleatório. Este processo consiste no mais importante exemplo de processo estocástico discreto, pois é a partir deste que, através dos teoremas de limite, obtemos dois dos principais processos presentes nas aplicações da Física: *os processos de Wiener e os de Poisson*, temas de nossa próxima seção.

II.2 Processos de Poisson e Processos de Wiener

Na seção anterior analisamos o passeio aleatório como um processo estocástico. Um fator importante de analisarmos no presente momento é o limite contínuo do processo do passeio aleatório discutido na seção anterior. Consideremos que a duração de cada passo da árvore binomial seja h , o que indica que o tempo total transcorrido após n passos seja $t = nh$. Com isso, as equações (II.15 a – c) se escrevem como

$$E[\ln Z] = \frac{t}{h} \mu \tag{II.16a}$$

$$\text{var}[\ln Z] = \frac{t}{h} \sigma^2 \tag{II.16b}$$

onde definimos

$$\mu = p \ln \frac{u}{d} + \ln d, \quad \sigma^2 = pq \left[\ln \left(\frac{u}{d} \right) \right]^2 \quad \text{II.17}$$

O limite contínuo será obtido quando tomarmos $n \rightarrow \infty$ ou equivalentemente $h \rightarrow 0$. Nosso problema consiste em encontrar valores para p , u e d de tal forma que os dois primeiros momentos de $\log Z$ sejam finitos em tomando-se o limite mencionado. Tais escolhas nos permitem obter dois limites distintos: o de Poisson e o de Gauss.

✧ O Limite de Poisson ([6], [31])

O limite de Poisson consiste em assumir que a probabilidade de cada passo está relacionada com a sua duração h de acordo com a equação $p = \lambda h$. Cabe-nos agora escolher os valores de u e de d de acordo com nossa discussão acima. Escolhamos portanto

$$u = u, \quad d = e^{-\xi h}$$

Com isso

$$\frac{\mu}{h} = \lambda \ln u - \xi + O(h) \quad \text{II.18a}$$

$$\frac{\sigma^2}{h} = \lambda (\ln u)^2 + O(h) \quad \text{II.18b}$$

Para uma distribuição de Poisson vale $\mu = \sigma^2$, e, portanto, pelas equações acima obtemos

$$(\ln u)^2 - \ln u + \frac{\xi}{\lambda} = 0$$

de onde calculamos

$$u = \exp \left[\frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{\xi}{\lambda}} \right] \quad \text{II.19}$$

Para obtermos o limite de Poisson, basta procedermos como fizemos no capítulo I: reescrevendo a equação (II.11a) como

$$w_n(m) = \frac{1}{m!} (np)^m \left(1 - \frac{np}{n} \right)^n \left[\frac{1}{(1-p)^m} \prod_{k=1}^{m-1} \left(1 - \frac{k}{n} \right) \right] \quad \text{II.20}$$

e substituindo-se os valores $n = t/h$ e $p = \lambda h$ obtemos, após tomarmos o limite $h \rightarrow 0$, a seguinte expressão:

$$w_i = \frac{(\lambda t)^m}{m!} e^{-\lambda t} \quad \text{II.21}$$

No limite $h \rightarrow 0$ obtemos, portanto, uma distribuição de Poisson. Assim,

$$X = m \ln u + (n - m) \ln d$$

é um processo estocástico de Poisson com distribuição

$$P[X \leq \zeta] = \sum_{k=0}^{\zeta} \frac{(\lambda t)^k}{k!} e^{-\lambda t} \quad \text{II.22}$$

✧ O Limite Gaussiano ([6], [31])

Nos casos anteriores, consideramos arbitrários os valores de u e de d , sujeitos a uma única restrição, a de que $d < u < 1$. Consideremos agora um modelo binomial simétrico, ou seja, um modelo onde os movimentos ascendentes e descendentes difiram por um mesmo fator:

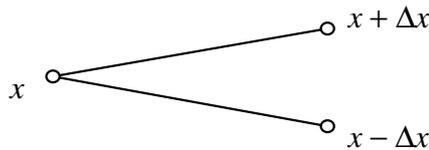


Figura II.3 : Ilustração de um Passeio Aleatório Simétrico

Somente modelos binomiais aditivos (*i. e.*, aqueles cujo termo no instante $t+1$ difere do termo no instante t apenas por um fator adicionado ou subtraído) satisfazem a propriedade indicada na figura. No caso do modelo binomial discutido neste capítulo vemos que em se tomando seu logaritmo natural este passa a ser aditivo, e portanto através da escolha conveniente dos fatores u e d podemos então obter um modelo simétrico. Como

$$\ln Z_{n+1} = \ln Z_n + \ln u \quad \text{ou} \quad \ln Z_{n+1} = \ln Z_n + \ln d$$

segue-se que, se $\ln u = -\ln d$,

$$\ln Z_{n+1} = \ln Z_n + \ln u \quad \text{ou} \quad \ln Z_{n+1} = \ln Z_n - \ln u$$

Pelas considerações acima vemos que, em tomando-se

$$u = \frac{1}{d} \quad \text{II.23}$$

obtemos então um modelo binomial simétrico. Para obtermos um processo cuja distribuição seja gaussiana, usaremos este modelo como base. Fazemos então as seguintes escolhas para u , d e p :

$$u = e^{\sigma\sqrt{h}}, \quad d = e^{-\sigma\sqrt{h}} \quad \text{II.24}$$

$$p = \frac{1}{2} \left(1 + \frac{\mu}{\sigma} \sqrt{h} \right)$$

Com isso, substituindo-se as equações (II.24) em (II.17) obtemos

$$\begin{aligned} E \ln Z &= \mu t \\ \text{var}[\ln Z] &= (\sigma^2 - \mu^2 h) t \end{aligned} \quad \text{II.25}$$

Das equações acima, podemos observar que os dois primeiros momentos deste processo são finitos no limite $h \rightarrow 0$:

$$\begin{aligned} \lim_{h \rightarrow 0} E \ln Z &= \mu t \\ \lim_{h \rightarrow 0} \text{var}[\ln Z] &= \sigma^2 t. \end{aligned} \quad \text{II.26}$$

Para aplicarmos o Teorema de De Moivre-Laplace, faz-se necessário primeiramente que calculemos os seguintes parâmetros:

$$\begin{aligned} x_k &= (k - np)h \\ &= \frac{1}{\sqrt{\text{var}[\ln Z]}} (\ln Z - E \ln Z) \end{aligned} \quad \text{II.27a}$$

$$\begin{aligned} h &= \sqrt{npq} \\ &= \frac{1}{\sqrt{\text{var}[\ln Z]}} \ln Z, \end{aligned} \quad \text{II.27b}$$

a partir dos quais obtemos

$$\begin{aligned} \lim_{h \rightarrow 0} h x_k^3 &= \lim_{h \rightarrow 0} \frac{\ln \frac{u}{d} (\ln Z - E \ln Z)^3}{\text{var}[\ln Z]^2} \\ &= 0. \end{aligned} \quad \text{II.28}$$

Com isso, verificam-se as condições do teorema de De Moivre-Laplace, o que nos leva à seguinte conclusão:

$$\lim_{t \rightarrow \infty} P \left[\frac{\ln Z - \mu t}{\sigma \sqrt{t}} \leq \zeta \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\zeta} \exp \left[-\frac{x^2}{2} \right]; \quad \text{II.29}$$

ou seja, o processo

$$X := \frac{\ln Z - \mu t}{\sigma \sqrt{t}} \quad \text{II.30}$$

apresenta como função de distribuição uma gaussiana com valor esperado 0 e variância 1.

Por meio da equação (II.30) observamos que é possível representar um processo com distribuição gaussiana com valor esperado μt e variância igual a $\sigma^2 t$ através da equação

$$Y_t = \mu t + \chi \sigma \sqrt{t}, \quad \text{II.31}$$

onde $\chi \sim N(0,1)$. A equação acima nos permite ainda uma certa liberdade em termos de mudarmos sua expressão sem no entanto alterar sua distribuição e momentos. Será através desta propriedade que construiremos o processo que posteriormente identificaremos com o movimento Browniano.

II.3 Martingales ([16], [18], [19], [27], [31])

Introduzimos na seção anterior um conceito extremamente importante, que consistirá em um dos pontos fundamentais de toda nossa discussão seguinte acerca dos Processos Estocásticos: o conceito de *informação e sua relação com o tempo*. Vimos que, pela definição de filtração, o conteúdo informacional associado a cada instante está inserido na sigma-álgebra associada ao instante imediatamente posterior; com isso deu-se o ordenamento no tempo das informações carregadas pelos processos estocásticos. Com isso, vemos que os processos estocásticos são, ao longo do tempo, *dependentes entre si*: isto é, a informação associada a um instante específico depende da informação associada ao instante imediatamente anterior.

Vamos analisar um pouco mais detalhadamente a questão da dependência entre processos. Suponhamos que, em um dado instante t , seja obtido um elemento ω em um experimento aleatório com estrutura de árvore binomial. Não sabemos exatamente o valor exato de ω , mas temos alguma informação a seu respeito, como, por exemplo, que o resultado do primeiro lançamento foi uma coroa. Como não temos a informação total, não é possível calcular o valor exato de $X_t(\omega)$, mas tão-somente uma estimativa sua. Esta estimativa depende da informação parcial que temos a respeito de ω , e portanto depende da sigma-álgebra dos eventos *anteriores* a t . Em função destas características, deduzimos então que tal estimativa é dada pelo valor esperado condicional de X_t em relação à uma dada sigma-álgebra associada ao tempo $\tau \leq t-1$.

Para ilustrarmos o cálculo desta estimativa utilizemos o exemplo dado no início desta seção, o da árvore binomial de preços. Iniciemos por $E[X_1 | \mathcal{A}_0]$; para tanto, usemos a propriedade de média parcial sobre os eventos de \mathcal{A}_0 :

$$\int_{\Omega} E[X_1 | \mathcal{A}_0] dP = \int_{\Omega} X_1 dP \quad \text{II.32a}$$

O lado direito da equação (II.32a) é igual a

$$\begin{aligned} \int_{\Omega} E[X_1 | \mathcal{A}_0] dP &= P[\Omega] E[X_1 | \mathcal{A}_0](\omega) \\ &= E[X_1 | \mathcal{A}_0](\omega), \quad \forall \omega \in \Omega \end{aligned}$$

ao passo que o lado esquerdo resulta em

$$\int_{\Omega} X_1 dP = puS_0 + qdS_0 \quad \text{II.32b}$$

Portanto,

$$E [X_1 | \mathcal{A}_0] (\omega) = (pu + qd)S_0 . \quad \text{II.33a}$$

Um cálculo similar nos leva aos seguintes resultados:

$$E [X_2 | \mathcal{A}_1] (\omega) = (pu + qd)X_1 (\omega) \quad \text{II.33b}$$

$$E [X_3 | \mathcal{A}_2] (\omega) = (pu + qd)X_2 (\omega) \quad \text{II.33c}$$

A partir das equações (II.33 a – c), concluímos que é válida a seguinte expressão:

$$E [X_{t+1} | \mathcal{A}_t] (\omega) = (pu + qd)X_t (\omega) \quad \text{II.34}$$

Pela equação acima, vemos que a estimativa do valor de um processo estocástico em um instante t , em considerando-se seu passado imediato $t-1$, é proporcional a seu valor associado a este passado imediato. Com isso, podemos verificar as seguintes conseqüências:

- i. Se $(pu + qd)X_t (\omega) < 1$, então a estimativa $E [X_{t+1} | \mathcal{A}_t] (\omega)$ *tende* a decrescer com o tempo;
- ii. Se $(pu + qd)X_t (\omega) = 1$, então a estimativa $E [X_{t+1} | \mathcal{A}_t] (\omega)$ não *tende* nem a crescer, nem a decrescer com o tempo.

Pelas relações acima, vemos que muito da informação do valor esperado condicional possui as mesmas características do processo subjacentes; assim, como sabemos já *a priori* os valores de p e de q , e de u e de d , torna-se simples a tarefa de visualizarmos as tendências dos valores estimados futuros. Introduzamos então a seguinte definição, que generaliza as idéias discutidas acima:

II.6. Definição : Seja (Ω, \mathcal{A}, P) um espaço de probabilidades e $\{ \mathcal{A}_t \}_{t \geq 0}$ uma filtração associada a este espaço. O processo estocástico $\{ X_t \}_{t \geq 0}$ sobre (Ω, \mathcal{A}, P) adaptado à $\{ \mathcal{A}_t \}_{t \geq 0}$ é denominado de *martingale* se satisfizer as seguintes condições:

- i. $EX_t < \infty$
- ii. $E [X_t | \mathcal{A}_s] = X_s$ para $s < t$;

O processo estocástico $\{ X_t \}_{t \geq 0}$ é denominado *supermartingale* (*submartingale*) se satisfizer a primeira condição acima juntamente com $E [X_t | \mathcal{A}_s] \geq X_s$ para $s < t$ (respectivamente $E [X_t | \mathcal{A}_s] \leq X_s$).

Assim, se um processo estocástico é um martingale, seu valor esperado condicional não tenderá nem a ascender, nem a decrescer. Importantes processos estocásticos são martingales, como por exemplo [32],

- i. *Passeio Aleatório* : Seja $\{X_t, t \in \mathcal{N}_G\}$ um processo estocástico *iid* tal que $EX_t = 0$ para todo $t \in \mathcal{N}_G$. O passeio aleatório

$$Y_t = \sum_{n=1}^t X_n, t \in \mathcal{N}_G$$

é um martingale.

- ii. *Processos Gaussianos* : Seja $\{X_t, t \in \mathcal{N}_G\}$ um processo gaussiano tal que $EX_t = 0$ e $E[X_t^2] = 1$. Então X é um martingale.

II.4 Processos de Markov ([16], [18], [19], [27], [31])

Em nossa discussão anterior acerca dos martingales, vimos uma maneira de como os conceitos de passado, presente e futuro podem ser correlacionados. Usamos como base o fato de a estimativa do processo no instante t depender da informação disponível até o momento imediatamente anterior; assim, o conteúdo informacional levado em consideração para o cálculo dos valores esperados condicionais refletia *todo o passado anterior* ao momento t . No entanto, há sistemas na natureza nos quais não é necessário conhecermos todo o seu passado para podermos fazer previsões sobre seu futuro; o próprio exemplo da árvore binomial ilustra este fato, conforme nos indica a figura II.3. Nesta figura, observamos que para obtermos o valor esperado condicional de X_3 é necessário somente que conheçamos a estrutura de eventos do passo *imediatamente* anterior; com isso, basta sabermos os valores do processo no instante imediatamente anterior para sabermos qual é seu conteúdo informacional. Conforme veremos, há inúmeros exemplos de relevância física que apresentam esta característica; veremos também que a *memória* de um sistema físico está diretamente relacionada com este conceito.

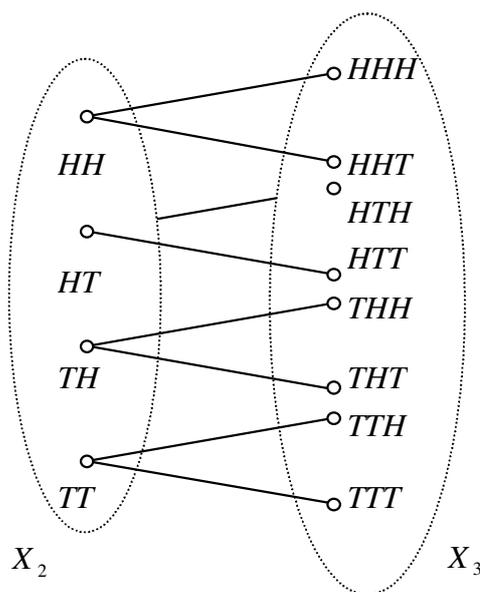


Fig. II.3 – Diagrama que ilustra a dependência entre dois passos na árvore binomial

Diante do exposto, introduzamos este novo conceito através da seguinte

II.7. Definição : Seja (Ω, \mathcal{A}, P) um espaço de probabilidades e $\{\mathcal{A}_t\}_{t \geq 0}$ uma filtração associada à este espaço. O processo estocástico $\{X_t\}_{t \geq 0}$ sobre (Ω, \mathcal{A}, P) adaptado à $\{\mathcal{A}_t\}_{t \geq 0}$ é denominado de *Processo de Markov* se satisfizer à *Propriedade de Markov*:

$$P[X_{t+1} \in B | \mathcal{A}_t] = P[X_{t+1} \in B | \sigma[X_t]] \quad \text{II.35}$$

onde $B \in \mathcal{B}[\mathcal{R}]$. Podemos também escrever a equação (II.35) da seguinte forma:

$$P[X_{t+1} \in B_{t+1} | X_t \in B_t, \dots, X_1 \in B_1] = P[X_{t+1} \in B_{t+1} | X_t \in B_t]$$

onde os eventos B são borelianos.

A propriedade de Markov possui outras formas equivalentes, conforme demonstra a seguinte

II.8. Proposição : As relações seguintes são equivalentes à propriedade de Markov:

- i.* Para toda função $\phi: \mathcal{R} \longrightarrow \mathcal{R}$, mensurável em um conjunto de Borel, tal que $E|\phi(X_{t+1})| < \infty$, segue-se que

$$E[\phi(X_{t+1}) | \mathcal{A}_t] = E[\phi(X_{t+1}) | \sigma[X_t]]$$

(concordância de valores esperados de funções reais arbitrárias)

- ii.* Para todo $\xi \in \mathcal{R}$ tal que $E \exp[\xi X_{t+1}] < \infty$, segue-se que

$$E[\exp[\xi X_{t+1}] | \mathcal{A}_t] = E[\exp[\xi X_{t+1}] | \sigma[X_t]]$$

(concordância de transformadas de Laplace)

- iii.* Para todo $\xi \in \mathcal{R}$ segue-se que

$$E[\exp[i\xi X_{t+1}] | \mathcal{A}_t] = E[\exp[i\xi X_{t+1}] | \sigma[X_t]]$$

(concordância de funções características)

Estas propriedades equivalentes à propriedade de Markov nos permitem provar em diversas situações que determinados processos são markovianos. Vamos provar formalmente que o modelo da dinâmica de preços é um processo de Markov. Para tanto, usaremos a propriedade *i.* da Proposição II.7. Definamos a variável aleatória

$$Z = \frac{X_{\tau+1}}{X_\tau} \quad \text{II.36}$$

onde consideramos que τ é um dado tempo fixo. Tomando-se um ponto amostral qualquer como sendo da forma $\omega = \omega_1 \omega_2 \cdots \omega_t$, onde $\omega_k = H$ ou $\omega_k = T$, segue-se da relação (II.36) que $Z(\omega) = u$ se $\omega_{\tau+1} = H$ ou que $Z(\omega) = d$ se $\omega_{\tau+1} = T$. Notemos que o Z independe do valor de ω_τ ; portanto, como a informação contida no instante τ não interfere no resultado de Z , concluímos que Z é independente de \mathcal{A}_t . Introduzamos agora uma função arbitrária ϕ ; portanto, calculando-se o valor esperado condicional de $\phi(X_{\tau+1})$ em relação a \mathcal{A}_t , obtemos

$$\begin{aligned} E[\phi(X_{\tau+1}) | \mathcal{A}_t] &= E\left[\phi\left(\frac{X_{\tau+1}}{X_\tau}\right) X_\tau \mid \mathcal{A}_t\right] \\ &= E[\phi(ZX_\tau) | \mathcal{A}_t] \\ &= p\phi(uX_\tau) + q\phi(dX_\tau); \end{aligned}$$

Lembrando-se que $\sigma[X_t] \subset \mathcal{A}_t$, obtemos, condicionando-se ambos os membros da equação acima à $\sigma[X_t]$ a seguinte expressão:

$$\begin{aligned} E[E[\phi(X_{\tau+1}) | \mathcal{A}_t] | \sigma[X_t]] &= pE[\phi(uX_\tau) | \sigma[X_t]] + qE[\phi(dX_\tau) | \sigma[X_t]] \\ &= p\phi(uX_\tau) + q\phi(dX_\tau). \end{aligned}$$

O processo X_τ é mensurável em relação a $\sigma[X_t]$, e, portanto, pela propriedade de torre (ver [16], [18], [19], [27], [31]), obtemos no lado direito da equação acima

$$E[E[\phi(X_{\tau+1}) | \mathcal{A}_t] | \sigma[X_t]] = E[\phi(X_{\tau+1}) | \sigma[X_t]].$$

Portanto,

$$E[\phi(X_{\tau+1}) | \mathcal{A}_t] = E[\phi(X_{\tau+1}) | \sigma[X_t]],$$

o que significa dizer que o modelo binomial da dinâmica de preços é um processo de Markov.

CAPÍTULO III

A TEORIA DE LÉVY

III.1 Introdução

Dentre todos os tipos de variáveis aleatórias discutidas até o momento, destacam-se indubitavelmente as variáveis aleatórias gaussianas, cujas propriedades especiais tornam-nas as variáveis mais adequadas para se tratar inúmeros problemas, não só os de ordem matemática, mas também os provenientes da Física, Economia ([3], [24]), e de muitos outros ramos do conhecimento. No entanto, para muitos efeitos práticos, as distribuições gaussianas se demonstraram insuficientes para uma descrição mais exata de muitos fenômenos físicos, uma vez que estas, conforme veremos posteriormente, não contemplam os eventos raros. Isto é, há muitos fenômenos cuja probabilidade de ocorrência, se calculada levando-se em consideração uma distribuição gaussiana, assumirá valor zero, ao passo que os mesmos fenômenos empiricamente poderão apresentar uma probabilidade de ocorrência diferente de zero. A figura abaixo nos permite ilustrar este fato:

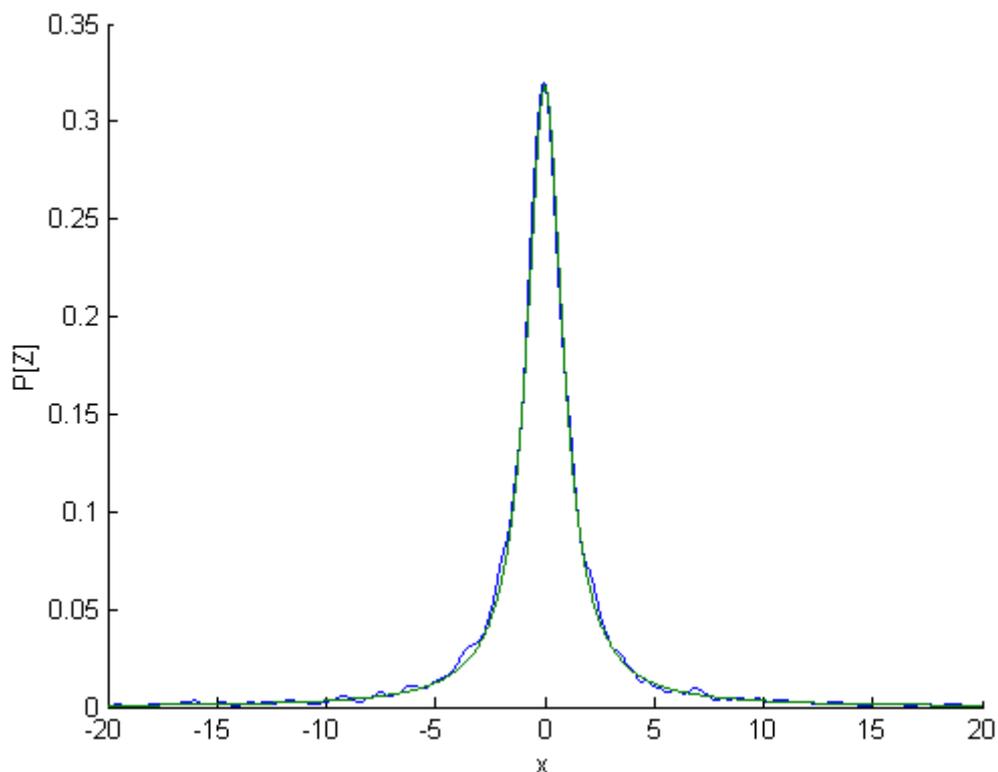


Fig. III.1 – Densidade de probabilidade com ‘caudas grossas’.

Nesta figura, podemos observar que a densidade acima apresenta um decaimento mais lento na região dos eventos raros $x \gg 1$; ademais, estes eventos contribuem para que

variância da distribuição apresente um alto valor; com isso, de modo a podermos descrever tais situações, é necessário desenvolvermos uma teoria na qual possamos analisar distribuições com grandes variâncias ou até mesmo com *variâncias infinitas*. É exatamente este o propósito da chamada *teoria de Lévy* ([13], [21], [23], [29]), tema central deste capítulo.

III.2 Distribuições Infinitamente Divisíveis – Conceitos Iniciais

Antes de lançarmos os alicerces para a construção da teoria de Lévy, analisemos mais detidamente algumas interessantes propriedades das variáveis aleatórias gaussianas. Consideremos para tanto um conjunto de variáveis aleatórias $\{X_k\}_{k=1}^N$ distribuídas de acordo com a lei $X_k \sim N(\mu_k; \sigma_k^2)$. Em termos de funções características, expressamos o fato de tais variáveis aleatórias serem gaussianas por meio da seguinte equação

$$\phi_k(\xi) = \exp\left[i\mu_k \xi - \frac{1}{2}\sigma_k^2 \xi^2\right]. \quad \text{III.1}$$

A soma destas variáveis aleatórias resulta na função característica

$$\phi(\xi) = \exp\left[i\mu\xi - \frac{1}{2}\sigma^2 \xi^2\right], \quad \text{III.2}$$

onde

$$\mu = \sum_{k=1}^N \mu_k \quad \text{e} \quad \sigma^2 = \sum_{k=1}^N \sigma_k^2. \quad \text{III.3}$$

Em termos de funções de distribuição, a mesma soma acima se escreve como

$$F(x) = F(x_1) * F(x_2) * \dots * F(x_N), \quad \text{III.4}$$

onde $F(x)$ e $F(x_k)$ são as funções de distribuição de X e de X_k , respectivamente. Portanto, a soma de N variáveis aleatórias gaussianas independentes resulta em uma variável aleatória também gaussiana. Dizemos, por este fato, que a distribuição gaussiana é *fechada sob convoluções*.

A propriedade de fechamento sob convoluções nos permite uma liberdade muito grande para escolhermos os valores de μ e de σ que comparecem em (III.2). Suponhamos, por exemplo, que $\mu_1 = \mu_2 = \dots = \bar{\mu}$ e que $\sigma_1^2 = \sigma_2^2 = \dots = \bar{\sigma}^2$; a função característica (III.2) torna-se

$$\phi(\xi) = \exp\left[iN\bar{\mu}\xi - \frac{N}{2}\bar{\sigma}^2 \xi^2\right], \quad \text{III.5}$$

ou

$$\phi(\xi) = \bar{\phi}(\xi)^N. \quad \text{III.6}$$

De modo inverso,

$$\check{\phi}(\xi) = \phi(\xi)^{1/N}. \quad \text{III.7}$$

Assim, tanto a N -ésima potência quanto a N -ésima raiz de uma função característica de uma distribuição gaussiana resulta igualmente em uma função característica de uma distribuição gaussiana. Isto é, podemos criar cópias da distribuição original simplesmente em tomando-se potências ou raízes de sua função característica. Este procedimento nos permite, por exemplo, criar seqüências de variáveis aleatórias cujas propriedades de suas componentes se assemelhem às de sua soma.

Passemos então a criar uma seqüência de variáveis aleatórias através da radiciação de uma função característica gaussiana; para tanto, introduzamos o conjunto de variáveis aleatórias independentes $\{X_k\}_{k=1}^N$ e suas funções características $\phi_k(\xi)$ associadas

$$\phi_k(\xi) = Ee^{i\xi X_k}, \quad \text{III.8}$$

Definamos agora a variável aleatória X (com momentos μ e σ^2) de tal modo que sua função característica $\phi(\xi)$ se relacione às funções $\phi_k(\xi)$ por

$$\phi_k(\xi) = \phi(\xi)^{1/k}. \quad \text{III.9}$$

É fácil de ver que os dois primeiros momentos dos termos da seqüência $\{X_k\}_{k=1}^N$ são dados por

$$\mu_k = \mu / k \quad \text{e} \quad \sigma_k^2 = \sigma^2 / k. \quad \text{III.10}$$

Assim, podemos escrever a variável aleatória X como sendo a soma

$$X = X_1 + X_2 + \cdots + X_N, \quad \text{III.11}$$

onde cada componente da seqüência possui momentos dados por (III.10). Do mesmo modo, se uma seqüência de variáveis aleatórias $\{Y_l\}_{l=1}^M$ apresentar como funções características $\chi_l(\xi)$ relacionadas a $\phi(\xi)$ por meio de uma equação análoga à (III.9), segue-se também que

$$X = Y_1 + Y_2 + \cdots + Y_M. \quad \text{III.12}$$

Das equações (III.11) e (III.12), podemos concluir que, independentemente da seqüência de variáveis aleatórias escolhida que satisfaça a prescrição (III.9), sempre a soma de suas componentes resultará na mesma variável aleatória X . Ou seja: *podemos decompor uma variável aleatória gaussiana em infinitas seqüências de variáveis aleatórias também gaussianas, e seu valor numérico a cada ponto amostral será dado pela soma dos termos destas seqüências calculadas neste mesmo ponto*. Esta propriedade das variáveis aleatórias gaussianas é denominada por *divisibilidade infinita*.

A propriedade de divisibilidade infinita das variáveis aleatórias gaussianas nos demonstra que, de um modo geral, as *partes* (isto é, as componentes de uma dada seqüência), se assemelham ao *todo* (isto é, sua soma). Este é o princípio que consiste a base da *geometria fractal*: o todo e as partes se assemelham a menos de um fator de escala. Conforme veremos, é

este princípio que generaliza o conceito do movimento Browniano, e, portanto, a base probabilística de toda a Física.

Diante destes fatos, somos levados a nos indagar: há mais famílias de variáveis aleatórias que apresentam a propriedade de divisibilidade infinita? Em caso afirmativo, qual seria sua representação geral? Para respondermos estas perguntas inicialmente introduzamos a seguinte ([13], [21], [23], [29])

III.1. Definição : Uma variável aleatória X é denominada por *infinitamente divisível* se for possível sua decomposição em uma soma de variáveis aleatórias *iid.* $\{X_{m_n}\}_{1 \leq m \leq n}$ para todo valor $n \in \mathcal{N}$:

$$X = X_{n_1} + X_{n_2} + \dots + X_{n_l} + \dots + X_{m_n} .$$

Em termos de funções características a propriedade acima se escreve como

$$\phi_n(\xi) = \phi(\xi)^{1/n} .$$

Como exemplos de funções infinitamente divisíveis podemos enumerar as seguintes funções características:

✧ **Distribuição Unitária :**

$$\phi(\xi) = \exp[i\xi x]$$

$$\phi_n(\xi) = \exp\left[i \frac{\xi}{n} x\right]$$

✧ **Distribuição de Poisson :**

$$\phi(\xi) = \exp[\lambda(e^{i\xi} - 1)]$$

$$\phi_n(\xi) = \exp\left[\frac{\lambda}{n}(e^{i\xi} - 1)\right]$$

✧ **Distribuição de Cauchy :**

$$\phi(\xi) = \exp[i\alpha\xi - \beta|x|]$$

$$\phi_n(\xi) = \exp\left[i \frac{\alpha}{n} \xi - \frac{\beta}{n}|x|\right]$$

✧ **Distribuição Gama :**

$$\phi(\xi) = \left[1 - i \frac{\xi}{\sigma}\right]^{-\lambda}$$

$$\phi_n(\xi) = \left[1 - i \frac{\xi}{\sigma} \right]^{-\lambda/n}$$

Por meio dos exemplos acima, podemos observar que algumas das distribuições mais comuns na Teoria da Probabilidade são infinitamente divisíveis. Este fato reforça a necessidade de uma representação mais geral para uma função característica infinitamente divisível que contenha todas as equações acima como casos particulares. Tal representação poderia contemplar, a princípio, os pontos nos quais as distribuições gaussianas falham, como, por exemplo, no caso dos eventos raros. Veremos, mais adiante, que tais funções características consistirão em um dos dois principais sustentáculos da teoria que nos permitirá solucionar este e alguns outros problemas decorrentes das distribuições gaussianas.

Passemos agora à discussão de algumas propriedades elementares das variáveis aleatórias infinitamente divisíveis ([13], [21], [23], [29]). Enunciaremos alguns teoremas e, nos casos mais fundamentais, discutiremos alguns argumentos usados em suas demonstrações.

III. 2. Teorema : A função característica de uma variável aleatória infinitamente divisível *nunca* assume valor zero.

A recíproca do teorema acima não é verdadeira: há funções características que jamais se anulam, no entanto não são infinitamente divisíveis.

III. 3. Teorema : A soma de um número finito de variáveis aleatórias infinitamente divisíveis também é infinitamente divisível.

Esta propriedade decorre imediatamente da definição III.1. Em decorrência deste teorema podemos provar o seguinte

III. 4. Corolário : Se $\phi(\xi)$ é uma função característica infinitamente divisível, então $|\phi(\xi)|$ também o é.

Após estes teoremas elementares, passemos à discussão das propriedades de convergência de seqüências de variáveis aleatórias infinitamente divisíveis. A primeira e mais elementar questão que podemos nos colocar é a seguinte: o limite de uma seqüência de variáveis aleatórias infinitamente divisíveis resulta em uma outra variável aleatória infinitamente divisível? Para respondermos esta questão, consideremos uma seqüência $\{X_m\}_{1 \leq m \leq n}$, $n \in \mathcal{N}$, de variáveis aleatórias infinitamente divisíveis com funções características $\phi_m^{(k)}(\xi)$. Introduzamos agora a expressão

$$X^{(k)} = \sum_{l=1}^n X_l^{(k)}, \quad \text{III.13}$$

e suponhamos que a seqüência $X^{(k)}$ convirja a um dado limite X para todo $\omega \in \Omega$. De acordo com o teorema (I.32), item III, o limite X é uma variável aleatória. A relação

$$X(\omega) = \lim_k X^{(k)}(\omega) \quad \text{III.14}$$

traduz-se, em termos de funções características, como

$$\phi(\xi) = \lim_k \phi^{(k)}(\xi), \quad \text{III.15}$$

uma vez que funções características são exponenciais, e portanto vale o conhecido resultado $x = \lim_n x_n \Rightarrow e^x = \lim_n e^{x_n}$. Em (III.15), $\phi^{(k)}(\xi)$ e $\phi(\xi)$ são as funções características associadas a $X^{(k)}$ e X respectivamente. Da definição III.1 segue-se que

$$\lim_k \phi^{(k)}(\xi) = \lim_k [\phi^{(k)}_m(\xi)]^m, \quad \text{III.16}$$

e, portanto, de (III.15) e (III.16),

$$\phi(\xi) = \lim_k [\phi^{(k)}_m(\xi)]^m. \quad \text{III.17}$$

Definindo-se

$$\phi_m(\xi) := \lim_k \phi^{(k)}_m(\xi), \quad \text{III.18}$$

segue-se de (III.17) e (III.18) que

$$\phi(\xi) = [\phi_m(\xi)]^m, \quad \text{III.19}$$

ou seja, a variável aleatória X é infinitamente divisível. Com isso, podemos enunciar o seguinte ([13], [21], [23], [29])

III. 5. Teorema : A variável aleatória cujo valor é igual ao limite de uma seqüência de variáveis aleatórias infinitamente divisíveis é infinitamente divisível.

Com o auxílio do teorema III.5 podemos explorar melhor a equação fundamental das funções características infinitamente divisíveis. Para tanto, usemos a conhecida representação para logaritmos naturais

$$\ln x = \lim_n n[x^{1/n} - 1], \quad \text{III.20}$$

que pode ser obtida por uma simples expansão em série de Taylor do termo $x^{1/n} = \exp[(1/n)\ln x]$. Inserindo-se uma função característica na equação (III.20) obteremos como resultado a seguinte expressão:

$$\ln \phi(\xi) = \lim_n n[\phi(\xi)^{1/n} - 1]. \quad \text{III.21}$$

Se $\phi(\xi)$ for infinitamente divisível, segue-se que

$$\ln \phi(\xi) = \lim_n n[\phi_n(\xi) - 1]. \quad \text{III.22}$$

Definamos agora

$$\psi_n(\xi) = n[\phi_n(\xi) - 1]; \quad \text{III.23}$$

vemos a partir desta que

$$\phi_n(\xi) := e^{\psi_n(\xi)} \quad \text{III.24}$$

é uma função característica de Poisson de parâmetro $\lambda = n$ e, portanto, infinitamente divisível. Com isso, o teorema III.5 mostra que o limite (III.22) é válido, pois ambos os membros desta equação são funções características infinitamente divisíveis.

A representação (III.22) é válida pois $\phi(\xi)$ é infinitamente divisível; no entanto, obteríamos maior generalidade se pudéssemos escolher funções arbitrárias $\chi_n(\xi)$ em sua expressão. Tal generalidade é permitida, conforme podemos demonstrar no seguinte teorema [23] :

III. 6. Teorema : Seja $\{\chi_n\}$ uma seqüência de funções características. Para que uma dada função característica $\phi(\xi)$ seja infinitamente divisível é necessário e suficiente que exista o limite contínuo

$$\psi(\xi) = \lim_n n[\chi_n(\xi) - 1] \quad \text{III.25}$$

onde, neste caso,

$$\phi(\xi) := e^{\psi(\xi)}. \quad \text{III.26}$$

A equação (III.25) nos permite ainda obter uma representação semelhante, conforme podemos observar através do seguinte [23]

III. 7. Lema : Seja $\chi(\xi)$ uma dada função característica e λ uma constante positiva. Então,

$$\phi(\xi) = \exp\{\lambda[\chi(\xi) - 1]\}$$

é uma função característica infinitamente divisível.

A expressão acima nos fornece a primeira representação geral para uma função característica infinitamente divisível, e esta se dá em termos de uma função de Poisson. Uma combinação dos resultados obtidos no teorema III. 6 e no lema III. 7. origina o importante [23]

III. 8. Teorema de De Finetti : Sejam $\{\chi_n(\xi)\}$ e $\{\lambda_n\}$ seqüências de funções características e de números positivos, respectivamente. Para que uma dada função característica $\phi(\xi) := e^{\psi(\xi)}$ seja infinitamente divisível, é necessário e suficiente que

$$\psi(\xi) = \lim_n \lambda_n [\chi_n(\xi) - 1]. \quad \text{III.27}$$

O teorema de De Finetti traz uma importantíssima implicação para a discussão que se segue. Para deduzirmo-la, consideremos inicialmente a representação das funções $\chi_n(\xi)$ em termos de suas funções de distribuição associadas $G_n(x)$:

$$\chi_n(\xi) = \int e^{i\xi x} dG_n(x). \quad \text{III.28}$$

Deste modo,

$$n[\chi_n(\xi) - 1] = n \int [e^{i\xi x} - 1] dG_n(x). \quad \text{III.29}$$

Feito isso, dividamos o intervalo $-\infty < x < \infty$ na partição

$$-\alpha = x_0 < x_1 < \dots < x_{n-1} < x_n = \alpha,$$

onde tomaremos posteriormente o limite $\alpha \rightarrow \infty$; com isso, discretizando-se a integral (III.29) obtemos as somas de Darboux

$$\begin{aligned} n \int [e^{i\xi x} - 1] dG_n(x) &= \lim_{m \rightarrow \infty} n \sum_{l=1}^m \exp(i\xi x_l - 1) [G_n(x_l) - G_n(x_{l-1})] \\ &= \lim_{m \rightarrow \infty} \sum_{l=1}^m \beta_{ln} \exp(i\xi x_l - 1), \end{aligned} \quad \text{III.30}$$

onde definimos

$$\beta_{ln} := n[G_n(x_l) - G_n(x_{l-1})].$$

De (III.25), (III.26), (III.29) e (III.30) segue-se que

$$\ln \phi(\xi) = \lim_{m \rightarrow \infty} \sum_{l=1}^m \beta_{ln} \exp(i\xi x_l - 1). \quad \text{III.31}$$

Com isso provamos um teorema que consiste na base de toda a nossa discussão acerca da representação geral de uma função característica infinitamente divisível ([13], [21], [23], [29]) :

III. 9. Teorema : O limite de uma seqüência de produtos finitos de funções características de Poisson é infinitamente divisível. De modo recíproco, toda função infinitamente divisível é o limite de uma seqüência de produtos finitos de funções características de Poisson.

Nas duas próximas seções aplicaremos os conceitos desenvolvidos até aqui para perseguirmos as representação mais geral para uma função característica infinitamente divisível.

III.3 A Representação Canônica de Kolmogorov

Uma vez obtidos estes teoremas fundamentais, passemos agora à obtenção da representação geral de uma função característica infinitamente divisível. Consideremos, inicialmente, uma seqüência de funções características da forma $\phi_n(\xi) := e^{\psi_n(\xi)}$, onde

$$\psi_n(\xi) = \lambda_n [\phi_n(\xi) - 1]. \quad \text{III.32}$$

As funções acima se relacionam com uma função característica infinitamente divisível de acordo com o teorema de De Finetti:

$$\psi(\xi) = \lim_n \psi_n(\xi). \quad \text{III.33}$$

Suponhamos agora que as funções de distribuição associadas às funções (III.32) sejam dadas pela seqüência $F_n(x)$, e que seus dois primeiros momentos sejam finitos, cujos valores são dados por

$$\mu_n = -i\psi'(0), \quad \text{III.34a}$$

$$\sigma_n^2 = -\psi''(0). \quad \text{III.34b}$$

Em considerando-se o limite (III.33) como sendo contínuo, é razoável supormos que os dois primeiros momentos (III.34a e b) também converjam para os limites μ e σ^2 . Para analisarmos esta convergência, reescrevamos (III.32) da seguinte forma:

$$\psi_n(\xi) = i\mu_n\xi + \chi_n(\xi) \quad \text{III.35}$$

onde vemos que as relações (III.34a e b) são satisfeitas se impusermos os vínculos

$$\chi(0) = 0, \quad \text{III.36a}$$

$$\chi'(0) = 0. \quad \text{III.36b}$$

Com essas relações em mãos, passemos agora ao cálculo dos termos individuais que comparecem na equação (III.35). Usando a representação

$$\phi_n(\xi) = \int e^{i\xi x} dF_n(x) \quad \text{III.37}$$

e as equações (III. 34a) e (III.35), obtemos

$$\begin{aligned} \mu_n &= -i\psi'(0) \\ &= \lambda_n \int x dF_n(x). \end{aligned} \quad \text{III.38}$$

Já o segundo momento é calculado diretamente a partir das equações (III.34b) e (III.35):

$$\begin{aligned} \sigma_n^2 &= -\psi''(0) \\ &= -\chi''(0) \\ &= -\lambda_n \int x^2 dF_n(x). \end{aligned} \quad \text{III.39}$$

A condição de σ_n^2 ser finita implica que as medidas $\lambda_n x^2 dF_n(x)$ sejam, por sua vez, também finitas; ademais, a existência do limite contínuo

$$\chi''(\xi) = \lim_n \chi_n''(\xi) \quad \text{III.40}$$

implica que estas mesmas medidas também converjam para um dado limite. Para melhor analisarmos estas questões, definamos as seguintes medidas de Riemann-Lebesgue

$$G_n(x) := \kappa_n \lambda_n \int_{-\infty}^x z^2 dF_n(z), \quad \text{III.41}$$

onde as constantes κ_n normalizam estas medidas de modo que tenhamos $G_n(+\infty) = 1$:

$$\kappa_n = \left\{ \lambda_n \int_{-\infty}^{\infty} z^2 dF_n(z) \right\}^{-1}. \quad \text{III.42}$$

As funções $\chi_n''(\xi)$ se expressam, em termos das medidas $G_n(x)$, da seguinte forma:

$$\chi_n''(\xi) = -\frac{1}{\kappa_n} \int_{-\infty}^{\infty} \exp[i\xi x] dG_n(x); \quad \text{III.43}$$

dividindo-se a equação acima por $\chi_n''(0)$, vemos que

$$\gamma_n(\xi) := \frac{\chi_n''(\xi)}{\chi_n''(0)} \quad \text{III.44}$$

são as funções características associadas às medidas $G_n(x)$,

$$\gamma_n(\xi) = \int_{-\infty}^{\infty} \exp[i\xi x] dG_n(x), \quad \text{III.45}$$

que, à luz da equação (III.40), convergem para o limite

$$\gamma(\xi) = \int_{-\infty}^{\infty} \exp[i\xi x] dG(x). \quad \text{III.46}$$

Com isso, concluímos que as medidas $G_n(x)$ satisfazem a equação

$$G(x) = \lim_n G_n(x). \quad \text{III.47}$$

A partir do limite (III.40) e da equação (III.46), a função $\chi''(\xi)$ assume a forma

$$\chi''(\xi) = -\int_{-\infty}^{\infty} \exp[i\xi x] dK(x), \quad \text{III.48}$$

onde definimos a medida

$$K(x) = \psi''(0) G(x). \quad \text{III.49}$$

Integrando-se (III.48) duas vezes, encontramos

$$\chi(\xi) = \int_{-\infty}^{\infty} \frac{1}{x^2} \exp[i\xi x] dK(x) + \alpha\xi + \beta, \quad \text{III.50}$$

onde as constantes α e β são obtidas a partir do uso dos vínculos (III.36a e b):

$$\alpha = -i \int_{-\infty}^{\infty} \frac{1}{x} dK(x), \quad \text{III.51a}$$

$$\beta = - \int_{-\infty}^{\infty} \frac{1}{x^2} dK(x). \quad \text{III.51b}$$

Deste modo, encontramos finalmente ([13], [21])

$$\chi(\xi) = \int_{-\infty}^{\infty} \frac{1}{x^2} \{ \exp[i\xi x] - i\xi x - 1 \} dK(x). \quad \text{III.52}$$

A convergência do termo $\chi_n(\xi)$ e da expressão (III.35) nos garantem a convergência do termo $i\mu_n \xi$ para um limite $i\mu \xi$ com μ real.

Analisemos agora o comportamento do integrando de (III.52) para medidas concentradas na origem. Devemos verificar o seguinte: como distribuições normais são casos particulares de distribuições infinitamente divisíveis, espera-se que a representação (III.52) se reduza à expressão de uma função característica normal para medidas concentradas na origem. Portanto, neste caso, a equação (III.52) deve fornecer um termo $-\frac{1}{2}\sigma^2 \xi^2$:

$$\begin{aligned} \frac{1}{x^2} \{ \exp[i\xi x] - i\xi x - 1 \} &= \frac{1}{x^2} \left\{ 1 + i\xi x - \frac{1}{2!} \xi^2 x^2 - \frac{i}{3!} \xi^3 x^3 + \dots \right. \\ &\quad \left. \dots - i\xi x - 1 \right\} \\ &= \frac{1}{x^2} \left\{ -\frac{1}{2!} \xi^2 x^2 - \frac{i}{3!} \xi^3 x^3 + \dots \right\} \\ &= -\frac{1}{2!} \xi^2, \quad x \rightarrow 0. \end{aligned} \quad \text{III.53}$$

De (III.49) e (III.52),

$$\begin{aligned} \chi(\xi) &= -\frac{1}{2} \xi^2 \int_{-\infty}^{+\infty} dK(x) \\ &= -\frac{1}{2} \xi^2 K(+\infty) \end{aligned}$$

$$= -\frac{1}{2}\xi^2\sigma^2. \quad \text{III.54}$$

Vemos, portanto, que a representação de $\chi(\xi)$ para uma medida concentrada na origem e a adição do termo $i\mu\xi$ resulta exatamente na expressão de uma função característica gaussiana. Confirmamos, assim, a eficiência da representação (III.52).

Com essas considerações feitas até aqui, podemos enunciar o seguinte ([13], [21], [23], [29])

III. 10. Teorema – A Representação Canônica de Kolmogorov : Seja K uma medida não decrescente, limitada, com $K(-\infty)=0$ e α uma constante positiva. Então, para que $\phi(\xi)$ seja uma função característica infinitamente divisível de variância finita é necessário e suficiente que esta seja representável sob a forma

$$\ln \phi(\xi) = i\alpha\xi + \int_{-\infty}^{\infty} \frac{1}{x^2} \{ \exp[i\xi x] - i\xi x - 1 \} dK(x), \quad \text{III.55}$$

onde o integrando de (III.55) na origem é calculado por continuidade, resultando em

$$\frac{1}{x^2} \{ \exp[i\xi x] - i\xi x - 1 \} \Big|_{x=0} = -\frac{1}{2!}\xi^2. \quad \text{III.56}$$

Ademais, esta representação é única.

III.4 As Representações Canônicas de Lévy-Khinchine e de Lévy

A representação canônica de Kolmogorov, teorema III.10, nos propiciou a forma mais geral para uma função característica infinitamente divisível dotada de variância infinita. A pergunta mais natural que poderíamos fazer no momento seria: é concebível uma generalização da representação canônica de Kolmogorov para o caso de distribuições com variância infinita? E, em caso de resposta afirmativa, seriam tais distribuições com variância infinita relevantes dentro do contexto da teoria da Probabilidade?

A resposta à primeira questão é afirmativa: de fato há uma representação para distribuições dotadas de variância infinita, conforme nos prova o seguinte teorema ([13], [21], [23], [29]) :

III. 11. Teorema – A Representação Canônica de Lévy-Khinchine : Seja G uma medida não decrescente, limitada, com $G(-\infty)=0$ e β uma constante positiva. Então, para que $\phi(\xi)$ seja uma função característica infinitamente divisível é necessário e suficiente que esta seja representável sob a forma

$$\ln \phi(\xi) = i\beta\xi + \int_{-\infty}^{\infty} \left\{ \exp[i\xi x] - 1 - \frac{i\xi x}{1+x^2} \right\} \frac{1+x^2}{x^2} dG(x), \quad \text{III.57}$$

onde o integrando de (III.57) na origem é calculado por continuidade, resultando em

$$\left\{ \left[\exp[i\xi x] - 1 - \frac{i\xi x}{1+x^2} \right] \frac{1+x^2}{x^2} \right\} \Big|_{x=0} = -\frac{1}{2!} \xi^2. \quad \text{III.58}$$

Ademais, esta representação é única.

A representação de Lévy-Khinchine ainda pode ser estendida para o caso de medidas que sejam descontínuas na origem. Consideremos, pois, as medidas

$$M(z) := \int_{-\infty}^z \frac{1+x^2}{x^2} dG(x), \quad x < 0 \quad \text{III.59a}$$

$$N(z) := -\int_z^{\infty} \frac{1+x^2}{x^2} dG(x), \quad x > 0 \quad \text{III.59b}$$

$$\sigma^2 = G(+0) - G(-0), \quad \text{III.59c}$$

que são claramente não-decrescentes nos intervalos $(-\infty, 0)$ e $(0, \infty)$, respectivamente. Ademais, tais medidas satisfazem claramente $M(-\infty) = N(\infty) = 0$. Para todo $\varepsilon > 0$, segue-se que as integrais

$$\int_{-\varepsilon}^0 x^2 dM(x) \quad \text{III.60}$$

e

$$\int_0^{\varepsilon} x^2 dN(x) \quad \text{III.61}$$

são finitas. De posse destes elementos, podemos enunciar o seguinte ([13], [21], [23], [29])

III. 12. Teorema – A Representação Canônica de Lévy : Sejam M e N medidas tais que:

- ✱ $M(x)$ e $N(x)$ são não-decrescentes nos intervalos $(-\infty, 0)$ e $(0, \infty)$;
- ✱ $M(-\infty) = N(\infty) = 0$;
- ✱ As integrais

$$\int_{-\varepsilon}^0 x^2 dM(x) \quad \text{e} \quad \int_0^{\varepsilon} x^2 dN(x)$$

são finitas para todo $\varepsilon > 0$.

Então, para que a função característica $\phi(\xi)$ seja infinitamente divisível, é necessário e suficiente que seja representável na forma

$$\begin{aligned} \ln \phi(\xi) = & i\gamma\xi - \frac{\sigma^2}{2}\xi^2 + \int_{-\infty}^{-0} \left\{ \exp[i\xi x] - 1 - \frac{i\xi x}{1+x^2} \right\} dM(x) \\ & + \int_{+0}^{+\infty} \left\{ \exp[i\xi x] - 1 - \frac{i\xi x}{1+x^2} \right\} dN(x), \end{aligned} \quad \text{III.62}$$

onde a constante σ^2 é dada por

$$\sigma^2 = G(+0) - G(-0)$$

e γ é uma constante real e positiva. Ademais, a representação acima é única.

A representação de Lévy é a mais geral dentre todas as representações possíveis para as funções características infinitamente divisíveis. É imediato verificar que a representação (III.62) se reduz à (III.57) se a medida G for contínua na origem. Também é fácil identificar a representação de Lévy-Khinchine com a de Kolmogorov no caso de variâncias finitas. Consideremos que (III.57) seja uma função característica com variância finita. Então, as integrais

$$\int_{-\infty}^{\infty} (1+x^2) dG(x) \quad \text{e} \quad \int_{-\infty}^{\infty} x dG(x)$$

são finitas. Definindo-se então

$$K(z) := \int_{-\infty}^z (1+x^2) dG(x)$$

e

$$\beta = \gamma + \int_{-\infty}^{\infty} x dG(x)$$

obtemos a representação de Kolmogorov (III.55). A recíproca é verdadeira, em admitindo-se a representação de Kolmogorov e tomando-se

$$G(z) := \int_{-\infty}^z \frac{1}{1+x^2} dK(x).$$

As demais condições do teorema de Lévy-Khinchine seguem do fato de a função característica em questão ser representável pelo teorema de Kolmogorov (III.55).

Para ilustrar a aplicabilidade dos conceitos até então desenvolvidos, observemos a tabela abaixo, que nos fornece alguns exemplos de como podemos escrever as funções características mais conhecidas em termos das três representações analisadas:

Função Característica	Representação							
	Kolmogorov		Lévy-Khinchine		Lévy			
Expressão	α	$K(x)$	β	$G(x)$	γ	σ	$M(x)$	$N(x)$
$\exp(i\xi x)$	ξ	0	ξ	0	ξ	0	0	0
$\exp\left[i\mu\xi - \frac{1}{2}\sigma^2\xi^2\right]$	μ	$\sigma^2\varepsilon(x)$	μ	$\sigma^2\varepsilon(x)$	μ	σ	0	0
$\exp[\lambda e^{i\xi} - 1]$	λ	$\lambda\varepsilon(x-1)$	$\frac{\lambda}{2}$	$\frac{\lambda}{2}\varepsilon(x-1)$	$\frac{\lambda}{2}$	0	0	$\lambda\varepsilon(x-1)$
$\exp(-\theta \xi)$	-	-	0	$\left(\frac{\theta}{\pi}\right)\tan^{-1}x + \frac{\theta}{2}$	0	0	$-\frac{\theta}{\pi x}$	$-\frac{\theta}{\pi x}$

Tabela III.1 : Representações Canônicas das principais funções características (adaptado de [Lukács]).

Pela tabela acima vemos que a função característica de Cauchy não é representável por Kolmogorov, exatamente pelo fato de possuir o segundo momento infinito.

A representação de Lévy, por sua maior generalidade, constitui o sustentáculo da teoria das distribuições infinitamente divisíveis. O próximo tópico a ser discutido, o que concerne a noção de *distribuições estáveis*, consistirá, juntamente com a teoria acima exposta, no maior trunfo da moderna teoria da Probabilidade e Processos Estocásticos para uma melhor descrição dos fenômenos naturais.

III.5 Distribuições Estáveis

Em conexão com a propriedade de divisibilidade infinita das variáveis aleatórias gaussianas, há uma outra propriedade das mesmas que nos chama a atenção. Como vimos, a soma de uma ou duas variáveis aleatórias gaussianas resulta em uma variável aleatória também gaussiana. Podemos generalizar ainda mais este conceito; para tanto notemos que

$$X = X_1 + X_2 + \cdots + X_N \quad \text{III.63}$$

é um caso particular da equação mais geral

$$Y = \alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_N X_N, \quad \text{III.64}$$

onde α_k , $k=1, \dots, N$, são números reais.

Tal variável aleatória é também gaussiana, conforme podemos facilmente demonstrar. Em termos de funções de distribuição, a relação acima se escreve como

$$F(y) = F(\alpha_1 x_1) * F(\alpha_2 x_2) * \dots * F(\alpha_N x_N). \quad \text{III.65}$$

Com isso, a relação mais geral possível para uma soma de variáveis aleatórias gaussianas é

$$F(y) = F(\alpha_1 x_1 + \beta_1) * F(\alpha_2 x_2 + \beta_2) * \dots * F(\alpha_N x_N + \beta_N), \quad \text{III.66}$$

onde α_k e β_k , $k = 1, \dots, N$, são números reais.

Da mesma maneira que a propriedade de divisibilidade infinita, a propriedade acima, que denominaremos de *estabilidade*, nos mostra que as mesmas propriedades valem tanto para os termos da seqüência de variáveis aleatórias gaussianas quanto para sua soma, indicando que, neste caso, as partes guardam uma relação intrínseca com o todo. Então, somos levados novamente a perguntar: há outras classes de variáveis aleatórias que apresentem tal propriedade? A resposta é afirmativa, e mostraremos que as leis estáveis nos levam à classe mais geral de processos que podem surgir na teoria da probabilidade: os *processos de Lévy*.

Vamos agora formalizar o conceito de estabilidade através da seguinte ([13], [21], [23], [29])

III. 13. Definição : A função de distribuição $F(x)$ é denominada *estável* se para toda constante $\alpha_k > 0$, β_k , com $k = 1, 2$ existam constantes $\alpha > 0$ e β tais que a igualdade

$$F(\alpha_1 x + \beta_1) * F(\alpha_2 x + \beta_2) = F(\alpha x + \beta)$$

seja válida.

Em termos de funções características, a relação acima pode ser escrita como

$$\phi\left(\frac{\xi}{\alpha}\right) = \phi\left(\frac{\xi}{\alpha_1}\right) \phi\left(\frac{\xi}{\alpha_2}\right) \exp(i\eta\xi) \quad \text{III.67}$$

onde $\eta = \beta - \beta_1 - \beta_2$. Para um número arbitrário de constantes κ_k , $k = 1, 2, \dots, n$, podemos deduzir da equação acima a seguinte expressão:

$$\phi\left(\frac{\xi}{\alpha}\right) = \phi\left(\frac{\xi}{\kappa_1}\right) \phi\left(\frac{\xi}{\kappa_2}\right) \dots \phi\left(\frac{\xi}{\kappa_n}\right) \exp(i\eta\xi). \quad \text{III.68}$$

Escolhendo-se todas as constantes κ_k como sendo iguais a 1, segue-se de (III.68) que

$$\phi(\xi) = \phi(\delta\xi) \exp(i\gamma\xi). \quad \text{III.69}$$

Tal expressão nos sugere que uma distribuição estável é infinitamente divisível; de fato, em definindo-se

$$\phi(\xi) = \left\{ \phi\left(\frac{\xi}{\delta}\right) \exp\left[-\frac{i\gamma\xi}{n\delta}\right] \right\}^n, \quad \text{III.70}$$

demonstramos de maneira simplificada o seguinte ([13], [21], [23], [29])

III. 14. Teorema : Toda função característica estável é infinitamente divisível.

Do mesmo modo que buscamos na seção anterior a representação mais geral para uma distribuição infinitamente divisível, passemos à investigação da representação mais geral para uma lei estável. De acordo com o teorema III.14., toda lei estável é infinitamente divisível; com isso, podemos representar a função característica de uma lei estável por meio da expressão (III.62). Cabe-nos agora a tarefa de descobrir as medidas $M(x)$ e $N(x)$ de tal modo que sua função característica satisfaça a relação (III.67). De (III.62) e (III. 67) obtemos

$$\begin{aligned} \ln \phi\left(\frac{\xi}{\kappa}\right) &= \ln \phi\left(\frac{\xi}{\kappa_1}\right) + \ln \phi\left(\frac{\xi}{\kappa_2}\right) + i\gamma\xi \\ &= i\gamma_1\xi - \frac{\sigma^2}{2\kappa_1^2} \xi^2 + \int_{-\infty}^0 \left\{ \exp[i\xi x] - 1 - \frac{i\xi x}{1+x^2} \right\} dM(\kappa_1 x) \\ &\quad + \int_{+0}^{\infty} \left\{ \exp[i\xi x] - 1 - \frac{i\xi x}{1+x^2} \right\} dN(\kappa_1 x) \\ &= i\gamma_2\xi - \frac{\sigma^2}{2\kappa_2^2} \xi^2 + \int_{-\infty}^0 \left\{ \exp[i\xi x] - 1 - \frac{i\xi x}{1+x^2} \right\} dM(\kappa_2 x) \\ &\quad + \int_{+0}^{\infty} \left\{ \exp[i\xi x] - 1 - \frac{i\xi x}{1+x^2} \right\} dN(\kappa_2 x) + i\eta\xi \\ &\equiv i\gamma\xi - \frac{\sigma^2}{2\kappa^2} \xi^2 + \int_{-\infty}^0 \left\{ \exp[i\xi x] - 1 - \frac{i\xi x}{1+x^2} \right\} dM(\kappa x) \\ &\quad + \int_{+0}^{\infty} \left\{ \exp[i\xi x] - 1 - \frac{i\xi x}{1+x^2} \right\} dN(\kappa x). \end{aligned} \quad \text{III.71}$$

Devido à unicidade da representação (III.62), segue-se que

$$\sigma^2 \left(\frac{1}{\kappa^2} - \frac{1}{\kappa_1^2} - \frac{1}{\kappa_2^2} \right) = 0 \quad \text{III.72a}$$

$$M(\kappa x) = M(\kappa_1 x) + M(\kappa_2 x), \quad x < 0; \quad \text{III.72b}$$

$$N(\kappa x) = N(\kappa_1 x) + N(\kappa_2 x), \quad x > 0, \quad \text{III.72c}$$

Determinemos agora as expressões de $M(x)$ e $N(x)$ por meio das expressões (III.72). É suficiente determinar apenas uma destas, pois ambos os casos são similares. Tomemos, por exemplo, $N(x)$. Por recorrência, obtemos de (III.72) a seguinte expressão:

$$N(\kappa x) = N(\kappa_1 x) + N(\kappa_2 x) + \cdots + N(\kappa_n x), \quad \text{III.73}$$

onde κ_k são constantes positivas. Escolhendo-se o caso particular $\kappa_k = 1$, segue-se de (III.73)

$$N(\kappa x) = nN(x), \quad \text{III.74}$$

onde $\kappa = \kappa(n)$. A medida $N(x)$ satisfaz portanto a equação

$$N(\kappa(\zeta)x) = \zeta N(x), \quad \text{III.75}$$

onde ζ é um número real e positivo. Podemos calcular, a partir da equação acima, a derivada de $N(x)$ em relação à x , fornecendo

$$\zeta \frac{d}{dx} N(x) = \kappa \frac{d}{d(\kappa x)} N(\kappa(\zeta)x). \quad \text{III.76}$$

Dividindo-se (III.76) por $N(x)$ e realizando alguns cálculos simples chega-se a equação diferencial

$$\frac{N'(x)}{N(x)} = \kappa \frac{1}{N(\kappa x)} \frac{d}{d(\kappa x)} N(\kappa x). \quad \text{III.77}$$

Tomando-se $x = 1$, e, em definindo-se

$$\frac{N'(1)}{N(1)} = -\alpha, \quad \text{III.78}$$

obtemos

$$\frac{dN(\kappa)}{N(\kappa)} = -\alpha \frac{d\kappa}{\kappa}, \quad \text{III.79}$$

cuja solução é

$$N(\kappa) = -\lambda \kappa^{-\alpha}, \quad \text{III.80}$$

onde λ é uma constante positiva.

O resultado $N(\infty) = 0$ nos mostra que a constante α também deve ser positiva. Procedendo de forma absolutamente análoga, encontramos para M a seguinte expressão:

$$M(\kappa) = -\frac{v}{|\kappa|^\beta}. \quad \text{III.81}$$

Por meio do teorema III.12, vemos que as restrições

$$\int_{-\varepsilon}^0 x^2 dM(x) < \infty, \quad \int_0^{\varepsilon} x^2 dN(x) < \infty$$

implicam que as constantes α e β devem assumir valores no intervalo $(0,2]$. Ademais, podemos relacioná-las empregando as equações (III.72b, c), (III.80) e (III.81), o que resulta na relação

$$\frac{1}{\kappa^\alpha} = \frac{1}{\kappa^\beta} = 2. \quad \text{III.82}$$

Pela equação acima vemos facilmente que $\alpha = \beta$.

Vamos agora obter a expressão da função característica de uma lei estável usando para tanto as medidas obtidas acima, (III.80) e (III.81). Dividamos estes cálculos em duas partes (uma vez que a solução para $0 < \alpha < 1$ é idêntica à solução para $1 < \alpha < 2$):

✱ $\alpha = 1$:

Calculemos a primeira integral em (III.62) em substituindo-se a medida (III.81) em sua expressão:

$$\begin{aligned} \int_{+0}^{+\infty} \left\{ \exp[i\xi x] - 1 - \frac{i\xi x}{1+x^2} \right\} \frac{dx}{x^2} &= \int_0^{+\infty} \left\{ \frac{\cos \xi x - 1}{x^2} \right\} dx \\ &+ i \int_{+0}^{+\infty} \left\{ \frac{\text{sen } \xi x - \frac{\xi x}{1+x^2}}{x^2} \right\} dx \\ &= -\frac{\pi}{2} \xi - i\xi \ln \xi + i\xi \int_{+0}^{+\infty} \left\{ \frac{\text{sen } x}{x^2} - \frac{1}{x(1+x^2)} \right\} dx. \end{aligned}$$

III.83

Para a segunda integral em (III.62) obtemos, após substituição da equação (III.80), a seguinte expressão:

$$\begin{aligned} \int_{+0}^{+\infty} \left\{ \exp[-i\xi x] - 1 + \frac{i\xi x}{1+x^2} \right\} \frac{dx}{x^2} &= -\frac{\pi}{2} \xi + i\xi \ln \xi \\ &- i\xi \int_{+0}^{+\infty} \left\{ \frac{\text{sen } x}{x^2} - \frac{1}{x(1+x^2)} \right\} dx. \end{aligned}$$

III.84

Introduzindo-se as constantes

$$\delta := (\lambda + \nu) \frac{\pi}{2}, \quad \text{III.85}$$

$$\beta := \frac{\nu - \lambda}{\lambda + \nu}, \quad \text{III.86}$$

segue-se que, para todo $\xi \in \mathcal{R}$,

$$\ln \phi(\xi) = i\gamma\xi - \delta|\xi| \left\{ 1 + i\beta \frac{2\xi}{\pi|\xi|} \ln|\xi| \right\}. \quad \text{III.87}$$

✱ $0 < \alpha < 1$:

De (III.62), (III.80) e (III.81) obtemos

$$\begin{aligned} \ln \phi(\xi) = i\gamma\xi + \nu \int_{-\infty}^{-0} \{\exp[i\xi x] - 1\} \frac{dx}{|x|^{1+\alpha}} \\ + \lambda \int_{+0}^{+\infty} \{\exp[i\xi x] - 1\} \frac{dx}{x^{1+\alpha}}. \end{aligned} \quad \text{III.88}$$

Fazendo-se a mudança de variáveis $\zeta = \xi x$, segue-se que

$$\begin{aligned} \ln \phi(\xi) = i\gamma\xi + \xi^\alpha \left\{ \nu \int_{-\infty}^{-0} \{\exp[-i\zeta] - 1\} \frac{d\zeta}{\zeta^{1+\alpha}} \right. \\ \left. + \lambda \int_{+0}^{+\infty} \{\exp[i\zeta] - 1\} \frac{d\zeta}{\zeta^{1+\alpha}} \right\} \\ = i\gamma\xi + \xi^\alpha (\nu e^{i\pi\alpha/2} + \lambda e^{-i\pi\alpha/2}) I(\alpha), \end{aligned} \quad \text{III.89}$$

onde definimos

$$I(\alpha) = \int_0^{\infty} \{\exp[-i\zeta] - 1\} \frac{d\zeta}{\zeta^{1+\alpha}}. \quad \text{III.90}$$

Introduzindo-se os termos

$$\beta := -I(\alpha)(\lambda + \nu) \cos \frac{\pi\alpha}{2}, \quad \text{III.91}$$

$$\delta := \frac{\nu - \lambda}{\lambda + \nu}, \quad \text{III.92}$$

e solucionando-se a integral (III.90) pelo teorema de Cauchy, obtemos, finalmente,

$$\ln \phi(\xi) = i\gamma\xi - \beta|\xi|^\alpha \left\{ 1 + i\delta \frac{2\xi}{\pi|\xi|} \tan \frac{\pi\alpha}{2} \right\}. \quad \text{III.93}$$

Com os elementos apresentados acima, podemos enunciar o seguinte

III. 15. Teorema - A Representação de uma Lei Estável : Para que uma dada distribuição $F(x)$ seja estável, é necessário e suficiente que o logaritmo de sua função característica seja representável por

$$\ln \phi(\xi) = i\gamma\xi - \beta|\xi|^\alpha \left\{ 1 + i\delta \frac{\xi}{|\xi|} \Phi(\xi, \alpha) \right\}, \quad \text{III.94}$$

onde a função $\Phi(\xi, \alpha)$ é definida por

$$\Phi(\xi, \alpha) = \begin{cases} \tan \frac{\pi\alpha}{2}, & \text{se } \alpha \neq 1 \\ \frac{2}{\pi} \ln|\xi|, & \text{se } \alpha = 1 \end{cases} \quad \text{III.95}$$

O parâmetro α , denominado *expoente de estabilidade*, satisfaz o vínculo

$$0 < \alpha \leq 2, \quad \text{III.96}$$

(ver discussão correspondente à dedução da relação (III.82)). Já o parâmetro δ , denominado *parâmetro de assimetria*, por sua vez, satisfaz o vínculo

$$-1 \leq \delta \leq 1.$$

(ver ([13], [21], [23], [29])).

Denominamos um processo estocástico dotado de uma função característica igual a (III.94) por *processo estocástico de Lévy*. Sua distribuição associada recebe o nome de *distribuição de Lévy*.

A representação III.94 é a mais geral possível para uma distribuição estável. Podemos demonstrar facilmente que o caso onde $\alpha = 2$ e $\beta = 1/2$ corresponde exatamente à expressão de uma função característica gaussiana. Vemos, com isso, que uma distribuição de Lévy, ao englobar duas importantes distribuições, a de Cauchy e a de Gauss, generaliza o próprio conceito de distribuição normal.

Podemos assim concluir face à discussão acima que os processos de Lévy generalizam o conceito de Movimento Browniano! Analisaremos as implicações desta importante descoberta no capítulo VII, ao investigarmos a estrutura probabilística de uma série de dados financeiros.

CAPÍTULO IV

O MOVIMENTO BROWNIANO

IV.1 Introdução

O primeiro cientista a demonstrar a existência deste fenômeno não somente para corpúsculos orgânicos, mas também para partículas inorgânicas foi o botanista inglês Robert Brown [25]. Brown analisou ao microscópio o comportamento de grãos de pólen de diferentes plantas dispersas em água, e observou que estas se comportavam como inúmeras partículas dotadas de um movimento irregular e ininterrupto.

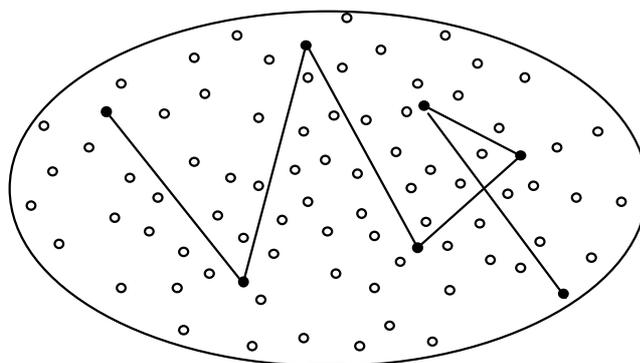


Fig. IV.1 : Movimento intermitente de um grão de pólen (ponto cheio) observado ao microscópio

Brown repetiu o experimento com outros corpúsculos orgânicos, e observou que o mesmo fenômeno ocorria. Concluiu inicialmente que tal movimento, conforme outros biólogos já haviam suposto anteriormente, era devido a uma ‘força vital’, presente em todos os compostos orgânicos. Com isso, Brown imaginou ter encontrado nestes corpúsculos a ‘molécula primitiva’ dos organismos vivos.

Brown repetiu o experimento com vários tipos de partículas inorgânicas, e, em observando-se que o mesmo fenômeno ocorria, concluiu que toda a matéria era composta por tais moléculas primitivas. Este trabalho foi publicado em 1828, e deste então tal movimento aleatório recebeu o nome de *Movimento Browniano*.

O grande mérito de Brown foi ter demonstrado que o fenômeno em questão não era devido à natureza orgânica da substância dispersada, refutando assim os argumentos ‘vitalistas’. Com isso, o movimento Browniano deixou de ser um fenômeno puramente ‘biológico’ para ser um fenômeno físico, necessitando, assim, de um tratamento matemático mais acurado.

Após Brown, outros pesquisadores estudaram o movimento Browniano visando uma explicação mais plausível para sua causa. Em 1858, Regnault propôs que o movimento era devido ao aquecimento irregular causado pela luz incidente. Posteriormente, Exner, em 1867, descobriu que o movimento se torna mais rápido com partículas menores, e também através de exposição a raios caloríficos e luminosos. Um pouco mais tarde, em 1870, Jevons sugeriu

que o movimento se deve a forças elétricas existentes entre as partículas em suspensão. Tal idéia não encontrou sustentação, sendo refutada no mesmo ano por Dancer, que demonstrou que tais forças não influenciam o movimento em si.

O primeiro passo rumo a uma teoria mais precisa do movimento Browniano foi dado em 1877, quando Delsaux formulou a hipótese de que a origem deste movimento intermitente era devida às colisões entre as partículas dispersadas e as moléculas do líquido dispersante.

Posteriormente, as investigações de Gouy levaram a importantes conclusões que suportaram os resultados obtidos anos depois por Einstein. Gouy descobriu que o movimento é tão mais intermitente quanto maior for a viscosidade do líquido dispersante; também descobriu que movimentos moleculares de origem térmica influenciavam o movimento das partículas dispersadas. Mostrou também que nem as mudanças da iluminação, nem campos magnéticos fortes influenciavam o movimento Browniano.

A teoria completa do movimento Browniano foi obtida por Einstein em uma série de importantes artigos escritos a partir de 1905 [8]. Em seu primeiro artigo, visando estudar o movimento de pequenas partículas de dimensão visível dispersas em líquidos, Einstein obteve pela primeira vez o coeficiente de difusão em termos de outras variáveis macroscópicas, como a viscosidade e a temperatura absoluta do líquido dispersante. Obteve também a famosa expressão que caracteriza um movimento Browniano, demonstrando que a distribuição do número de partículas suspensas em um ponto x do espaço em um instante temporal t é gaussiana. Neste mesmo trabalho Einstein deduziu uma fórmula para o cálculo das dimensões moleculares.

Com isso, Einstein lançou as bases de uma das teorias mais bem sucedidas para a modelagem de sistemas naturais. O movimento Browniano pode ser aplicado a qualquer sistema dotado de movimentos aleatórios cuja distribuição de probabilidades seja gaussiana. A aplicabilidade desta teoria é enorme: podemos analisar desde o estudo de partículas dispersas em um líquido até a precificação de uma opção no mercado financeiro.

IV.2 A Teoria de Einstein do Movimento Browniano

IV.2.1 A Teoria da Difusão

O primeiro grande fenômeno que discutiremos como base para a teoria do movimento Browniano denomina-se *difusão*. Para analisarmos este fenômeno consideremos um gás monoatômico encerrado em um recipiente com ν moléculas por unidade de volume, em estado de equilíbrio (fig. IV.2)

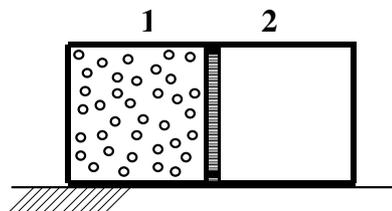


Fig. IV.2 : Gás encerrado em uma câmara

Suponhamos que, em um dado instante, um orifício de área A é aberto, permitindo assim a passagem deste gás para uma outra câmara:

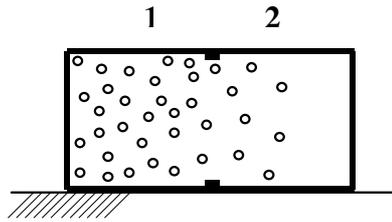


Fig. IV.3 : O processo de difusão

O transporte das moléculas do gás de uma câmara para outra cria um fluxo molecular que atua no sentido de restaurar a situação de equilíbrio do sistema. Isso pois, como as moléculas do gás se encontram em movimento intermitente, colisões entre estas e com as paredes do recipiente impelem-nas a executar um movimento em todas as direções. Este processo de transporte criado pela diferença de distribuições do gás nas duas câmaras é denominado de processo de *difusão*. Notemos que, nesta situação de não-equilíbrio, a distribuição de moléculas do gás v é função da coordenada espacial, e, portanto $v = v(x)$.

Procedamos agora à obtenção da expressão da densidade de fluxo Φ das partículas transportadas do gás. Para tanto, observemos a seguinte figura:

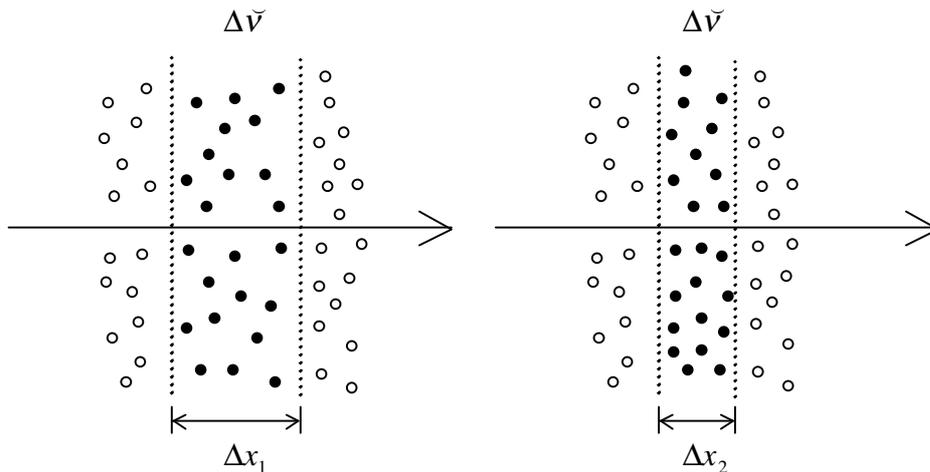


Fig. IV.4 : O fluxo molecular em um processo de difusão

Suponhamos que fosse possível rotularmos um certo número de moléculas do gás; suponhamos também que sua distribuição seja de \check{v} moléculas por unidade de volume. O fluxo molecular pode ser entendido como sendo a quantidade de moléculas que atravessa uma dada faixa de área A , espessura Δx , em um dado instante de tempo Δt ; no entanto, sem perda de generalidade, podemos considerar $A=1$ e $\Delta t=1$, a fins de simplificarmos a expressão. O fluxo molecular será tão maior quanto maior for a quantidade de moléculas e menor for a espessura Δx , conforme vemos facilmente pela fig.IV.4. Com isso, obtemos para o fluxo molecular a seguinte expressão:

$$\Phi = -D \frac{\partial}{\partial x} \check{v}(x), \quad \text{IV.1}$$

onde D é uma constante de proporcionalidade que se denomina *coeficiente de difusão*. Se $dv/dx > 0$, então o movimento de moléculas se dá no sentido de maior concentração; no

entanto, como o fluxo real se move no sentido de menor concentração, colocamos o sinal negativo de modo a tornarmos a constante D positiva, e, conseqüentemente, o fluxo Φ . A expressão (IV.1) será denominada por *Lei de Fick* se a distribuição de moléculas do gás for também função do tempo.

A partir agora do conceito de fluxo molecular é possível determinar a equação diferencial que rege o processo de difusão. Consideremos então que $v(x, t)$ seja a distribuição média de moléculas por unidade de volume, situadas na posição x ao instante t . Suponhamos que, no exato instante t , fotografássemos a superfície A pela qual o gás escoo (fig.IV.5), e escolhêssemos uma faixa de espessura dx para observarmos a difusão de suas moléculas:

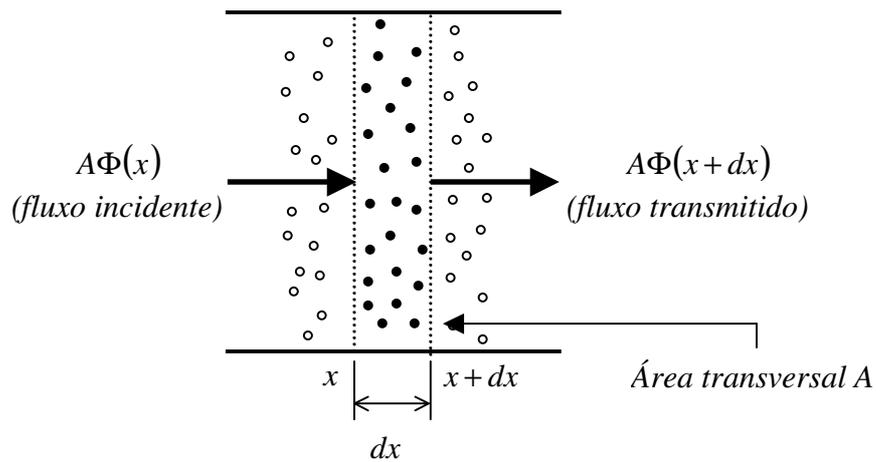


Fig. IV.5 : Processo de transporte no gás em termos dos fluxos moleculares

Nesta faixa encontram-se $v(x, t)A dx$ moléculas. Notemos que qualquer aumento nesta quantidade de moléculas será devido ao fluxo incidente de moléculas pela superfície delimitada por x ; conseqüentemente, serão emitidas pela superfície delimitada por $x+dx$ uma quantidade tal de moléculas que será igual ao fluxo transmitido. Com isso, a variação temporal do número total de moléculas compreendido entre as superfícies delimitadas por x e $x+dx$ será igual a

$$\frac{\partial}{\partial t} \{v(x, t)A dx\} = A\Phi(x) - A\Phi(x+dx). \quad \text{IV.2}$$

Expandindo-se o lado esquerdo de (IV.2), obtemos

$$\frac{\partial}{\partial t} v(x, t) = -\frac{\partial}{\partial x} \Phi(x). \quad \text{IV.3}$$

Pela lei de Fick, equação (IV.1), obtemos finalmente a *equação de difusão*

$$\frac{\partial}{\partial t} v(x, t) = D \frac{\partial^2}{\partial x^2} v(x, t). \quad \text{IV.4}$$

O exemplo usado acima pode ser ainda estendido a situações mais complexas. Consideremos, por exemplo, que na câmara 2 estivesse contido um gás diferente do da câmara 1. Suponhamos que ambos estivessem inicialmente isolados, e que, de algum modo,

podéssemos rotular as moléculas de ambos os gases. Após termos removido a separação, veríamos que as moléculas de ambos os gases se difundiriam em sentidos opostos, visando ocupar todo o volume disponível. Após certo tempo estariam os gases totalmente misturados. O processo de difusão seria, neste caso, idêntico à difusão de um gás em uma câmara vazia; com isso, vemos que a teoria, e conseqüentemente as equações desenvolvidas, se aplicam aos casos mais gerais.

IV.2.2 O Conceito de Pressão Osmótica

As experiências conduzidas ao longo dos últimos cinco séculos que culminaram nos trabalhos de Robert Brown tiveram sempre como foco principal o movimento de corpúsculos em suspensão aquosa. Independentemente da investigação realizada (por exemplo, os esporos de bactérias, como nas pesquisas de Spallanzani, ou grãos de pólen, como nas pesquisas de Brown), sempre as causas daquele movimento intermitente foram questionadas. Inicialmente, conforme já mencionamos na introdução a este capítulo, acreditava-se que o movimento era devido a uma força vital presente nos corpúsculos orgânicos. No entanto, conforme demonstrado empiricamente por Brown, partículas inorgânicas também realizavam aquele movimento intermitente quando em suspensão. Com isso, a causa do movimento destes corpúsculos deixou de pertencer exclusivamente ao campo da Biologia e ingressou definitivamente no campo da investigação da Física.

Nesta seção discutiremos os princípios físicos fundamentais que regem o movimento de partículas suspensas em um dado líquido. Esta discussão, associada ao conceito de difusão apresentado na seção anterior, consistirá na base fundamental para a construção da teoria do movimento Browniano.

Consideremos pois um líquido de volume total V no qual dispersaremos uma certa quantidade de material sólido composto por grãos, por exemplo. Se o dispersado for solúvel no líquido dispersante, observaremos a formação de uma solução homogênea (consideramos aqui que a quantidade de soluto é suficiente apenas para saturar a solução). Caso contrário, veremos que os grãos dispersos ficarão suspensos neste líquido, em um movimento aleatório intermitente. No primeiro caso, sabemos que as moléculas do soluto se dissolvem totalmente no solvente, ao passo que, no segundo caso, as moléculas continuam em um estado de agregação tal que a estrutura formada (o corpúsculo) se mantém intacta.

Para melhor investigarmos estas duas situações, suponhamos que em ambos os casos o material dispersado esteja envolto no interior de uma membrana semipermeável de volume v , que permite a livre passagem do dispersante mas não das moléculas ou dos corpúsculos. No caso de uma solução não-iônica, a existência desta membrana internamente à solução promove o fenômeno da *osmose*: haverá um movimento do líquido de fora para dentro da membrana, visando diluir ainda mais o soluto encerrado. Neste caso, uma *pressão osmótica* será exercida sobre a membrana, sendo dada pela *equação de Van't Hoff* [8]

$$p = \frac{nRT}{v}, \quad \text{IV.5}$$

onde n é o número de moles do soluto e T é a temperatura absoluta da solução.

No caso de corpúsculos suspensos, podemos imaginar que também haja uma pressão a ser exercida sobre a membrana semipermeável, pois o líquido, do ponto de vista empírico, também flui para o interior da membrana. No caso anterior, a lei de Van't Hoff pode ser obtida experimentalmente; no caso presente, teremos de utilizar argumentos da Mecânica Estatística para justificarmos a existência desta pressão 'osmótica'.

Consideremos, portanto, que hajam n corpúsculos encerrados no volume v de líquido contido na membrana semipermeável; também consideremos que o volume total dos corpúsculos seja desprezível se comparado com o volume v de líquido. Assumamos também que o espaço de fase deste sistema seja completamente caracterizado pelas variáveis $(q_1, \dots, q_m; p_1, \dots, p_m)$. A função de partição associada a este sistema será então dada pela expressão [8] (ver também [28] e [34])

$$Z = \int \exp\left\{-\frac{\varepsilon}{k_B T}\right\} dq dp, \quad \text{IV.6}$$

onde utilizamos as notações abreviadas $dq = dq_1 \cdots dq_m$ e $dp = dp_1 \cdots dp_m$. A densidade de pontos do espaço de fase no ensemble canônico considerado é então dada pela expressão

$$\rho(q, p, t) = \frac{n}{Z} \exp\left\{-\frac{\varepsilon}{k_B T}\right\}. \quad \text{IV.7}$$

Com esta expressão para a densidade, calculamos facilmente a probabilidade de que um dado ponto do espaço de fase se encontre nos intervalos $\lambda_k + d\lambda_k$, onde λ_k representa qualquer uma das coordenadas de fase do sistema:

$$dW = \frac{dZ}{Z}. \quad \text{IV.8}$$

Uma vez estabelecidos estes conceitos associados ao sistema em questão, passemos à sua análise. Separemos na expressão (IV.6) os termos associados às posições dos n corpúsculos das variáveis restantes; assim poderemos verificar como a função de partição (e, conseqüentemente, a entropia e a função de energia livre do sistema) varia em relação a v . Para tanto, sejam (x_k, y_k, z_k) , $k=1, \dots, n$ as coordenadas dos centros de massa dos corpúsculos suspensos. Consideremos que os corpúsculos estejam contidos nos elementos de volume infinitesimais (dx_k, dy_k, dz_k) . Com isso, a função de partição (IV.6) escreve-se como

$$\begin{aligned} dZ &= \exp\left\{-\frac{\varepsilon}{k_B T}\right\} dq dp \\ &= dJ dv_1 \cdots dv_n. \end{aligned} \quad \text{IV.9}$$

onde definimos os volumes infinitesimais $dv_k = dx_k dy_k dz_k$. Se as funções J forem independentes tanto das coordenadas do sistema em questão, quanto do volume v , vemos que a função de partição acima guarda com o volume v uma simples relação de proporcionalidade.

Provemos agora que a função J é independente das coordenadas de posição dos centros de massa das partículas suspensas e do volume de líquido encerrado v ; para tanto, consideremos que o sistema em si tenha, de um certo modo, sido deslocado de sua situação inicial (por exemplo, por uma agitação da membrana). Com isso, o novo sistema terá (x'_k, y'_k, z'_k) como as coordenadas do centro de massa das partículas suspensas. Deste modo, a função de partição deste novo sistema será dada por

$$dZ' = dJ' dv'_1 \cdots dv'_n. \quad \text{IV.10}$$

Como os volumes infinitesimais permanecem os mesmos, segue-se que $dv_k = dv'_k$, e, portanto, de (IV.9) e (IV.10),

$$\frac{dZ}{dZ'} = \frac{dJ}{dJ'}. \quad \text{IV.11}$$

Em termos das medidas de probabilidade (IV.8),

$$\frac{dW}{dW'} = \frac{dJ}{dJ'}. \quad \text{IV.12}$$

Da equação acima podemos facilmente obter o resultado desejado. Se o líquido for homogêneo e não exercer forças sobre as partículas, então qualquer novo sistema criado a partir do sistema inicial por uma alteração *apenas* das coordenadas espaciais dos corpúsculos não implica em uma alteração das distribuições dos mesmos. Com isso, a probabilidade de encontrarmos as partículas em uma dada configuração espacial será idêntica em ambos os sistemas. Então, de (IV.12) segue-se que

$$J = J' \quad \text{IV.13}$$

Portanto, a expressão final da função de partição passa a ser dada por

$$Z = Jv^n, \quad \text{IV.14}$$

ao passo que a entropia e a energia livre de Helmholtz do sistema tem como expressão

$$S = \frac{E}{T} + k_B \{\ln J + n \ln v\} \quad \text{IV.15}$$

$$F = -k_B T \{\ln J + n \ln v\}. \quad \text{IV.16}$$

A partir da equação

$$p = -\frac{\partial F}{\partial v}, \quad \text{IV.17}$$

obtemos finalmente a relação

$$\begin{aligned} p &= k_B T \frac{n}{v} \\ &= \frac{RT}{N} v. \end{aligned} \quad \text{IV.18}$$

Portanto, argumentos da Mecânica Estatística nos asseguram a existência de uma pressão osmótica para o caso de partículas suspensas em um dado líquido. Assim, vemos que o comportamento tanto de moléculas dissolvidas quanto de partículas em suspensão é idêntico

quando em um estado de grande diluição (ou seja, o volume total das partículas é ínfimo se comparado ao volume do líquido). Ademais, podemos também concluir que a diferença entre as partículas do soluto dissolvidas e corpúsculos em suspensão reside *somente* em suas dimensões.

IV.2.3 O Movimento de Partículas suspensas em um Líquido

Uma vez discutidos os dois importantes conceitos de difusão e pressão osmótica, passemos agora à análise detalhada do movimento de partículas suspensas em um líquido.

Consideremos, como ponto de partida, um cilindro Z de secção transversal unitária preenchido por certo um líquido; no instante $t = 0$ colocamos uma certa massa de uma dada substância granulada em uma região imaginária X :

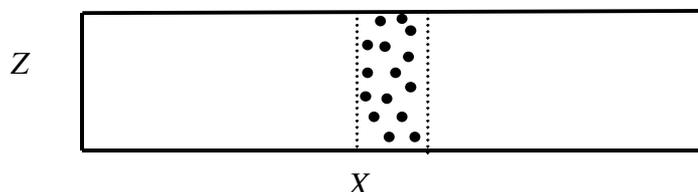


Fig. IV.6 : Processo de Difusão em um Líquido

Imediatamente após termos colocado a substância no líquido verificaremos o desencadear de um processo de difusão, pois a pequena massa das partículas e a viscosidade do líquido farão com que estas fiquem suspensas e, portanto, propensas a colisões. As pressões exercidas pelo líquido em ambos os lados da região X geram uma pressão osmótica que contribui no sentido de intensificar as colisões entre as partículas suspensas, e estas assim tenderão a percorrer espaços cada vez maiores, impulsionadas pelas colisões com as demais partículas que encontrarem pelo caminho. Este movimento se expande em todas as direções, visando uma ocupação maior do volume do líquido disponível, até que o mesmo esteja homogeneamente preenchido pelas partículas em suspensão. Após esta fase final, observamos que as partículas passarão a executar um movimento intermitente e irregular.

No processo de difusão descrito acima observamos que são duas as forças exercidas sobre as partículas suspensas: a força gerada pela pressão osmótica e a força de resistência imposta pelo líquido ao movimento das partículas. Considerando-se, por fins de simplicidade, que o processo de difusão ocorra apenas no sentido do eixo das abscissas, observaremos que tais forças se dispõem em sentidos contrários, cancelando-se, portanto, mutuamente. Com isso, o sistema 'líquido + partículas suspensas' se encontra em equilíbrio dinâmico, o que nos permite a utilização, neste caso, das grandezas termodinâmicas desenvolvidas na seção anterior.

Analisemos agora as equações que regem este processo físico. Consideremos que a substância dispersada, em função do processo de difusão, sofra um deslocamento arbitrário δx :

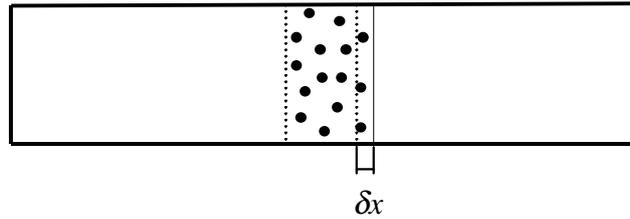


Fig. IV.7 : Deslocamento da Substância dispersada para fins de cálculo

Como o sistema se encontra em equilíbrio dinâmico, a energia livre permanece constante em se considerando tal deslocamento arbitrário δx ; com isso,

$$\delta F = \delta E - T\delta S = 0. \quad \text{IV.19}$$

Seja K a força atuante sobre cada partícula e v a densidade de partículas suspensas por unidade de volume no estado de equilíbrio dinâmico. A variação de energia por partícula em termos do deslocamento δx é dada por [8]

$$-K\delta x, \quad \text{IV.20a}$$

ao passo que a variação infinitesimal da energia da massa da substância dispersa é dada por

$$-K\delta x v dx. \quad \text{IV.20b}$$

Deste modo, obtemos [8]

$$\delta E = -\int K v \delta x dx. \quad \text{IV.20}$$

Consideramos o sinal negativo na equação (IV.20 a) pelo fato de a força exercida sobre as partículas ser resistente ao movimento.

A partir da equação (IV.15) obtemos para a variação da entropia

$$\delta S = k_B \frac{n}{v} \delta v. \quad \text{IV.21a}$$

O termo

$$\frac{n}{v} \delta v \quad \text{IV.21b}$$

representa o número de partículas suspensas encerradas no volume δv ; este termo é idêntico à expressão

$$\int v d\delta x. \quad \text{IV.21c}$$

Com isso,

$$\begin{aligned}
\delta S &= k_B \int v d\delta x \\
&= k_B \int v \frac{\partial \delta x}{\partial x} dx \\
&= -k_B \int \frac{\partial v}{\partial x} \delta x dx.
\end{aligned} \tag{IV.21}$$

Substituindo-se (IV.20) e (IV.21) em (IV.19), obtemos finalmente [8]

$$Kv - k_B T \frac{\partial v}{\partial x} = 0. \tag{IV.22}$$

A equação (IV.22) rege portanto o processo de difusão das partículas suspensas no líquido no estado de equilíbrio dinâmico. A força K representa a resistência imposta pelo líquido ao movimento de cada partícula, que, no caso de partículas esféricas de raio a , é dada pela *lei de Stokes*

$$K = -6\pi a \eta u, \tag{IV.23}$$

onde η é o coeficiente de viscosidade do líquido e u é a velocidade de cada partícula.

Como o processo de difusão desempenha um papel fundamental na análise do movimento das partículas suspensas, podemos, a partir da equação (IV.22), obter uma expressão para o cálculo do coeficiente de difusão D . Para tanto, substituamos a expressão (IV.23) em (IV.22):

$$6\pi a \eta u v = -k_B T \frac{\partial v}{\partial x},$$

o que fornece

$$vu = -\frac{k_B T}{6\pi a \eta} \frac{\partial v}{\partial x}. \tag{IV.24}$$

O lado esquerdo da equação acima fornece o número de partículas que atravessam a superfície unitária por unidade de tempo; com isso, a grandeza vu representa um fluxo de partículas. Como tal fluxo é devido à difusão, segue-se que este deve ser igual ao fluxo difusional da lei de Fick:

$$\Phi_D = -D \frac{\partial v}{\partial x}. \tag{IV.25}$$

Com isto, obtemos a partir de (IV.24) e (IV.25) a expressão final para o coeficiente de difusão de uma substância suspensa em um líquido ([8], [28], [34])

$$D = \frac{k_B T}{6\pi \eta a}. \tag{IV.26}$$

A expressão obtida para o coeficiente de difusão, equação (IV.26), contém as informações necessárias para podermos explicar as causas do movimento intermitente das partículas suspensas. Conforme vimos, são as colisões existentes entre as partículas suspensas que permitem o processo de difusão; tais colisões serão tão mais frequentes quanto maior for a temperatura da substância dispersada, pois as partículas em si disporão de uma maior energia cinética. Com isso, o processo de difusão será facilitado, o que implica em um maior valor para o coeficiente de difusão. Portanto, empiricamente, o coeficiente de difusão de um sistema deve ser diretamente proporcional à temperatura absoluta da substância dispersada, fato este comprovado teoricamente pela equação (IV.26).

Um outro fator que interfere diretamente no processo de difusão se refere à resistência oferecida pelo líquido ao movimento das partículas. Quanto menor for o coeficiente de viscosidade do líquido dispersante, menor será esta resistência, e, conseqüentemente, mais frequentes serão as colisões entre as partículas. Deste modo, será facilitado o processo de difusão, que implica em um maior valor para o coeficiente de difusão. Podemos assim dizer que, do ponto de vista empírico, o coeficiente de difusão é inversamente proporcional à viscosidade do líquido, fato este novamente comprovado teoricamente pela equação (IV.26).

Por último, sabemos que partículas menores se difundirão mais rapidamente, pois necessitam de uma menor quantidade de energia para serem colocadas em movimento. Com isso, também se confirma a relação de proporcionalidade inversa entre o coeficiente de difusão e o raio das partículas suspensas conforme consta na relação (IV.26).

Podemos ver, face ao resultado obtido pela equação (IV.26), que o modelo acima exposto fornece justificativas plausíveis a muitos dos fenômenos que ocorrem no processo de dispersão de partículas em um líquido. Em face da discussão acima, podemos então inferir que a causa principal do processo de difusão de tais partículas consiste nas colisões existentes entre estas. Devemos notar, entretanto, que tanto o deslocamento efetuado por uma partícula entre duas colisões consecutivas quanto à direção que esta segue depende do ‘arranjo’ instantâneo das partículas que estiverem à sua frente. Observemos a figura IV.8:

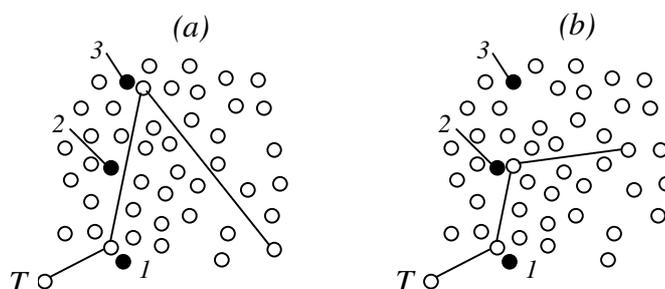


Fig. IV.8 : ‘Arranjos’ de partículas diferentes

Suponhamos que, em um dado instante t , a partícula T colidisse com a partícula 1; se a partícula 2 sofresse uma colisão anterior de tal modo que não fosse projetada em direção à partícula T , observaríamos a configuração (a), ou seja, a partícula T , após colidir com a partícula 1, colidiria em seguida com a partícula 3. Caso contrário, a partícula T se encontraria com a partícula 2 conforme a configuração (b), ou seja, seguiria uma outra direção, e, conseqüentemente, não colidiria com a partícula 3 logo em seguida. Como as situações (a) e (b) acima representam arranjos diferentes, vemos que não é possível, *a priori*, predizermos qual seria a posição da partícula T em um dado instante posterior a t , pois desconhecemos o arranjo futuro das partículas ao redor de T . Tanto poderia ocorrer o arranjo (a), quanto o arranjo (b), ou mesmo muitíssimos outros, imprevisíveis, pois nosso sistema em consideração tem infinitos graus de liberdade. Podemos concluir então, pelas figuras acima, que o

deslocamento futuro de uma dada partícula é *independente* de seu deslocamento *imediatamente anterior*.

Com isso, podemos dizer que as colisões entre as partículas ocorrem *aleatoriamente*, e, em função disso, somente podemos inferir a probabilidade de uma dada partícula se situar em uma posição x em um dado instante t .

Para calcularmos esta probabilidade, consideremos novamente um cilindro Z contendo um líquido onde, no instante $t=0$, depositamos uma certa quantidade da substância dispersada em uma certa região X (figura IV.9a):

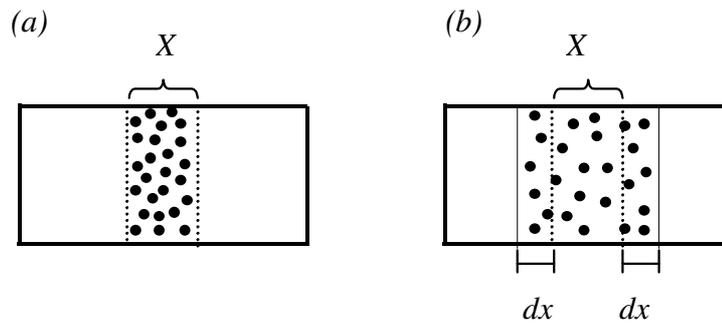


Fig. IV.9 : Em (a), a substância dispersada se encontra encerrada em uma região X no instante $t=0$; em (b) vemos seu estado após transcorrido um intervalo de tempo τ

Após transcorrido um certo intervalo de tempo τ , cada partícula suspensa sofrerá, devido ao processo de difusão, um acréscimo x diferente em suas coordenadas espaciais; tais acréscimos poderão tanto ser positivos quanto negativos (Fig. IV.9b). Por esta mesma figura podemos facilmente observar que somente as partículas que sofrem acréscimo cuja magnitude se encontra no intervalo $x+dx$ ocuparão as duas faixas de espessura dx . Sendo $f(x,t)dx$ a probabilidade de a posição de uma dada partícula ter magnitude entre x e $x+dx$ no instante t , segue-se que o número de partículas que se encontram neste intervalo é dado pela expressão

$$dn = nf(x,t)dx, \quad \text{IV.27}$$

onde $f(-x,t) = f(x,t)$. Em termos da densidade de partículas por unidade de volume, a expressão acima se escreve como

$$dn = v(x,t)dx, \quad \text{IV.28}$$

o que nos fornece

$$v(x,t) = nf(x,t). \quad \text{IV.29}$$

A expressão da função $f(x,t)$ pode ser obtida em solucionando-se a equação de difusão deste processo, dada pela expressão (IV.4), e em usando-se a relação (IV.29):

$$\frac{\partial}{\partial t} f(x,t) = D \frac{\partial^2}{\partial x^2} f(x,t). \quad \text{IV.30}$$

Para solucionarmos esta equação, tomemos a transformada de Fourier da função $f(x, t)$,

$$f(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ikx} \phi(k, t) dk, \quad \text{IV.31}$$

e a substituamos na equação (IV.30), resultando na expressão

$$\frac{\partial}{\partial t} \phi(x, t) = -Dk^2 \phi(x, t), \quad \text{IV.32}$$

cuja solução é dada por

$$\phi(x, t) = C e^{-Dk^2 t}. \quad \text{IV.33}$$

Com isso,

$$\begin{aligned} f(x, t) &= \frac{C}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ikx - Dk^2 t} dk \\ &= \frac{C}{\sqrt{2Dt}} \exp\left\{-\frac{x^2}{4Dt}\right\}. \end{aligned} \quad \text{IV.34}$$

Usando-se a equação (IV.27) obtemos, finalmente, a expressão

$$f(x, t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left\{-\frac{x^2}{4Dt}\right\}. \quad \text{IV.35}$$

Assim, o número de partículas por unidade de volume que, entre os instantes $t = 0$ e $t = t$ sofreram um acréscimo de magnitude entre x e $x + dx$ em suas coordenadas de posição é dado pela equação

$$v(x, t) = \frac{n}{\sqrt{4\pi Dt}} \exp\left\{-\frac{x^2}{4Dt}\right\}. \quad \text{IV.36}$$

Com isso, a lei de probabilidade que rege o movimento de partículas dispersadas em um líquido é *gaussiana*. Isso somente é possível graças ao fato de, conforme a discussão que se seguiu à figura IV.8, os eventos *deslocamento entre os instantes 0 e t*, e *deslocamento entre os instantes t e Δt*, serem *mutuamente exclusivos*.

Uma vez determinada a lei de probabilidade que rege o fenômeno estudado, podemos calcular sua média e desvio médio quadrático. O valor médio de x é dado pela expressão

$$\langle x \rangle = \int_{-\infty}^{+\infty} x f(x, t) dx, \quad \text{IV.37}$$

o qual, pela equação (IV.35), fornece

$$\langle x \rangle = 0. \quad \text{IV.38}$$

O desvio quadrático médio, dado pela equação

$$\langle x^2 \rangle = \int x^2 f(x, t) dx, \quad \text{IV.39}$$

apresenta, a partir da equação (IV.35), o valor

$$\begin{aligned} \sigma &:= \sqrt{\langle x^2 \rangle} \\ &= \sqrt{2Dt}. \end{aligned} \quad \text{IV.40}$$

Com base nas expressões acima, a covariância de $x(t)$ e de $x(s)$, $0 \leq s < t$, é dada por

$$\text{cov}(x(t), x(s)) = \langle x(t)x(s) \rangle, \quad \text{IV.41}$$

o que fornece

$$\begin{aligned} \langle x(t)x(s) \rangle &= \langle x(s)[x(t) - x(s) + x(s)] \rangle \\ &= \langle x(s)[x(t) - x(s)] \rangle + \langle x(s)^2 \rangle; \end{aligned} \quad \text{IV.42}$$

como $x(s)$ e $x(t) - x(s)$ são independentes, segue-se de (IV.38), que

$$\langle x(s)[x(t) - x(s)] \rangle = \langle x(s) \rangle \langle x(t) - x(s) \rangle = 0. \quad \text{IV.43}$$

Portanto, usando-se a expressão (IV.40), obtemos

$$\text{cov}(x(t), x(s)) = 2Ds. \quad \text{IV.44}$$

De modo geral, ([28], [34])

$$\text{cov}(x(t), x(s)) = 2D \min(s, t) \quad \text{IV.45}$$

A partir da expressão do desvio quadrático médio podemos obter uma expressão mais detalhada para o cálculo do valor teórico do número de Avogadro. Para tanto, substituamos a expressão (IV.26) em (IV.40), obtendo assim

$$\sigma = \sqrt{\frac{k_B T}{3\pi\eta a}} t, \quad \text{IV.46}$$

ou

$$N = \frac{1}{\sigma^2} \frac{RT}{3\pi\eta a} t. \quad \text{IV.47}$$

A relação acima fecha de forma triunfal a teoria desenvolvida até aqui. Veremos, na seção seguinte, as importantíssimas implicações que esta teoria, desenvolvida por Einstein, trouxe ao campo da Física.

IV.2.4 O Movimento Browniano

De posse de toda a teoria desenvolvida na seção anterior, podemos lançar então a seguinte

IV.1 Definição : Denominamos por *Movimento Browniano* o movimento intermitente de partículas suspensas em um líquido estacionário, causado pelas sucessivas colisões experimentadas por tais partículas umas com as outras. A probabilidade de encontramos uma partícula entre x e $x+dx$, entre os instantes $t=0$ e $t=t$, é dada por

$$f(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left\{-\frac{x^2}{4Dt}\right\}, \quad \text{IV.48}$$

onde D é o coeficiente de difusão do processo, dado pela equação

$$D = \frac{k_B T}{6\pi\eta a}, \quad \text{IV.49}$$

onde η é o coeficiente de viscosidade do líquido, a o raio das partículas, e T a temperatura absoluta do sistema. O valor médio da coordenada espacial de uma partícula em um instante qualquer é dado por

$$\langle x \rangle = 0, \quad \text{IV.50}$$

ao passo que o desvio médio quadrático experimentado por esta partícula é igual a

$$\sigma = \sqrt{2Dt}. \quad \text{IV.51}$$

Nestas condições, o número de Avogadro tem como equação

$$N = \frac{1}{\sigma^2} \frac{RT}{3\pi\eta a} t. \quad \text{IV.52}$$

Os deslocamentos efetuados por uma partícula são independentes entre si, constituindo portanto em eventos mutuamente exclusivos.

A teoria desenvolvida na seção anterior, cujo resultado em suma apresentamos acima, foi lançada em um artigo de Einstein em 1905, intitulado “*Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung in ruhenden Flüssigkeiten suspendierten Teilchen*”, e publicado na famosa revista ‘*Annalen der Physik*’. Inicialmente Einstein não considerou que sua teoria explicasse o movimento Browniano (após o artigo de 1827 de Brown), conforme sua própria menção no artigo supracitado: “É possível que os movimentos a serem discutidos aqui sejam idênticos ao denominado ‘Movimento Browniano’; no entanto, a informação a mim disponível é tão imprecisa que não posso efetuar

qualquer julgamento a respeito”. Após sua publicação, Einstein fora informado através dos trabalhos de Sidentopf e de Gouy que sua teoria descrevia realmente o movimento Browniano conforme primeiramente descrito por Robert Brown.

Em seu artigo, além de fornecer uma explicação plausível para a existência daquele movimento intermitente de partículas, Einstein lançou duas equações fundamentais que seriam cruciais para a verificação da validade experimental de sua teoria: a equação (IV.49), que fornece uma fórmula para o cálculo do coeficiente de difusão do sistema, e a equação (IV.51), que fornece o deslocamento quadrático médio das partículas. Juntas, estas equações permitiam a obtenção de uma terceira, a (IV.52), que permite o cálculo teórico da constante de Avogadro. O físico Seddig promoveu a primeira prova experimental de que a teoria de Einstein do movimento Browniano estava correta; logo em seguida, Jean Perrin não somente demonstrou a plena concordância da teoria de Einstein com os experimentos, mas também obteve experimentalmente o valor do número de Avogadro usando a equação (IV.52). Graças a estas medidas e conseqüente obtenção do valor do número de Avogadro, Jean Perrin obteve, em 1926, o prêmio Nobel de Física.

Os trabalhos de Einstein e Perrin forneceram pela primeira vez evidências acerca da natureza molecular da matéria. Também, graças à equação (IV.52), Einstein pode, pela primeira vez, mensurar as dimensões moleculares. Assim, a teoria do Movimento Browniano, conforme desenvolvida por Einstein, constituiu a última grande contribuição da Física Clássica para a elucidação dos fenômenos microscópicos da matéria antes do advento da Mecânica Quântica.

IV.3 A Teoria de Langevin

Posteriormente aos trabalhos de Einstein, o físico francês Paul Langevin lançou uma teoria na qual obteve o mesmo resultado que Einstein ([28], [34]). O fator que diferencia a teoria de Langevin da de Einstein é a introdução de uma força aleatória estacionária $F(t)$ que seria responsável pelas colisões entre as partículas suspensas. Sendo m a massa de cada partícula de dotada de movimento Browniano, e $\dot{x}(t)$ sua velocidade, a equação do movimento será dada por

$$m \frac{d^2}{dt^2} x(t) = -6\pi\eta a \frac{d}{dt} x(t) + F(t). \quad \text{IV.52}$$

Pelo fato de $F(t)$ ser estacionária, podemos supor que

$$\langle F(t) \rangle = 0. \quad \text{IV.53}$$

Multiplicando-se ambos os lados da equação (IV.45) por $x(t)$, e usando-se a expressão

$$x(t) \frac{d^2}{dt^2} x(t) = \frac{d}{dt} [x(t)\dot{x}(t)] - [\dot{x}(t)]^2, \quad \text{IV.54}$$

segue-se que

$$m \frac{d}{dt} [x(t)\dot{x}(t)] = m\dot{x}(t)^2 - 6\pi\eta a x(t)\dot{x}(t) + x(t)F(t). \quad \text{IV.55}$$

Tomando-se os valores médios de ambos os lados de (IV.55), obtemos

$$m \left\langle \frac{d}{dt} [x(t)\dot{x}(t)] \right\rangle = m \langle \dot{x}(t)^2 \rangle - 6\pi\eta a \langle x(t)\dot{x}(t) \rangle + \underbrace{\langle x(t)F(t) \rangle}_{=0}, \quad \text{IV.56}$$

onde usamos a relação (IV.53). Pela Teoria Cinética dos Gases sabemos que

$$\frac{1}{2} m \langle \dot{x}(t)^2 \rangle = \frac{1}{2} k_B T, \quad \text{IV.57}$$

o que resulta, portanto, na equação

$$\frac{d}{dt} \langle x(t)\dot{x}(t) \rangle = \frac{k_B T}{m} - \frac{6\pi\eta a}{m} \langle x(t)\dot{x}(t) \rangle, \quad \text{IV.58}$$

cuja solução é

$$\langle x(t)\dot{x}(t) \rangle = C e^{-\gamma t} + \frac{k_B T}{6\pi\eta a}, \quad \text{IV.59}$$

onde definimos $\gamma = 6\pi\eta a / m$. Usando-se a condição inicial

$$\langle x(0)\dot{x}(0) \rangle = 0, \quad \text{IV.60}$$

obtemos

$$\langle x(t)\dot{x}(t) \rangle = \frac{k_B T}{6\pi\eta a} (1 - e^{-\gamma t}). \quad \text{IV.61}$$

Empregando-se a relação

$$\langle x(t)\dot{x}(t) \rangle = \frac{1}{2} \frac{d}{dt} \langle x(t)^2 \rangle, \quad \text{IV.62}$$

e a condição inicial

$$\langle x(0)^2 \rangle = 0, \quad \text{IV.63}$$

obtemos a equação diferencial

$$\frac{d}{dt} \langle x(t)^2 \rangle = \frac{k_B T}{3\pi\eta a} (1 - e^{-\gamma t}), \quad \text{IV.64}$$

cuja solução é igual a

$$\langle x(t)^2 \rangle = \frac{k_B T}{3\pi\eta a} \left[t - \frac{1}{\gamma} (1 - e^{-\gamma t}) \right]. \quad \text{IV.65}$$

A solução acima da equação de Langevin possui dois limites específicos:

✧ **Limite $t \ll \gamma^{-1}$:**

Neste caso, a solução é dada por

$$\langle x(t)^2 \rangle = \frac{k_B T}{m} t^2; \quad \text{IV.66}$$

✧ **Limite $t \gg \gamma^{-1}$:**

A solução (IV.65) assumirá a forma

$$\langle x(t)^2 \rangle = \frac{k_B T}{3\pi\eta a} t; \quad \text{IV.67}$$

com isso, devido ao fato de estarmos analisando um processo difusivo e, portanto, valer a relação

$$\begin{aligned} \sigma &= \sqrt{\langle x(t)^2 \rangle} \\ &= \sqrt{2Dt}, \end{aligned}$$

segue, da equação (IV.67), a expressão

$$D = \frac{k_B T}{3\pi\eta a}, \quad \text{IV.68}$$

ou seja, resultado idêntico ao obtido pela teoria de Einstein.

Deste modo, a teoria de Langevin da difusão apresenta resultados idênticos ao da teoria de Einstein no limite $t \gg \gamma^{-1}$.

CAPÍTULO V

A TEORIA MATEMÁTICA DO MOVIMENTO BROWNIANO

V.1 O Movimento Browniano como um Processo Estocástico

V.1.1 O Passeio Aleatório e o Movimento Browniano

Após termos investigado o fenômeno físico do movimento Browniano através da Termodinâmica, analisaremos, neste capítulo, suas propriedades matemáticas fundamentais ([2], [4], [5], [6], [7], [11], [16], [18], [19], [25], [27] e [31]). Isso pois, conforme vimos, há uma estrutura probabilística por trás do conceito, devida à natureza aleatória das colisões existentes entre as partículas suspensas no líquido. Conforme veremos, muitas das variáveis associadas aos processos físicos poderão ser traduzidas em termos puramente matemáticos, o que muito enriquece a Teoria do Movimento Browniano.

Para iniciarmos a descrição matemática do movimento Browniano deveremos introduzir primeiramente o espaço de probabilidade sobre o qual construiremos a teoria. Iniciemos pelo caso mais simples, o de um Movimento Browniano a tempos discretos ([16], [19], [27] e, principalmente, [31]).

No capítulo anterior vimos que a função densidade de probabilidade associada a um Movimento Browniano é escrita em termos da variável que representa o deslocamento sofrido pela partícula, além do tempo gasto para efetuá-lo; diante disso podemos, então, considerar o deslocamento da partícula como a variável aleatória associada ao Movimento Browniano. De acordo também com o que discutimos no capítulo anterior, os deslocamentos sofridos pelas partículas ocorriam ao longo do eixo das abscissas; assim, em termos desta variável aleatória, a evolução temporal do Movimento Browniano poderia tomar a seguinte forma, em um instante arbitrário:

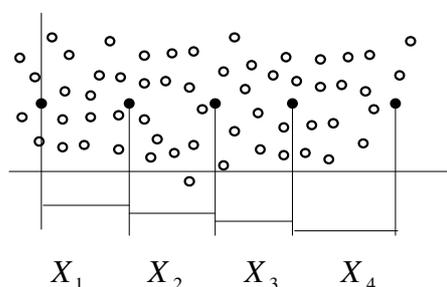


Fig. V.1 : Deslocamento de uma partícula arbitrária em solução aquosa

No esquema acima consideramos, por simplificação gráfica, o movimento no sentido crescente; no entanto, em nossas considerações matemáticas o sentido do movimento é arbitrário, podendo ser tanto para a direita quanto para a esquerda.

Pela situação anterior verificamos que o movimento Browniano pode ser representado, no caso de tempos discretos, em termos de um passeio aleatório. Para simplificar, podemos considerar cada passo na figura como sendo de comprimento fixo; com isso, representemo-los por meio do passeio aleatório

$$X_k(\omega) = \begin{cases} 1, & \omega = H \\ -1, & \omega = T \end{cases} \quad \text{V.1}$$

Considerando-se que a probabilidade de obtermos uma cara é igual à probabilidade de obtermos uma coroa, é fácil verificarmos as seguintes condições:

$$EX_k = 0; \quad \text{V.2a}$$

$$EX_k^2 = 1; \quad \text{V.2b}$$

$$EX_k X_l = EX_k EX_l, \quad k \neq l. \quad \text{V.2c}$$

A última destas nos demonstra que estas variáveis aleatórias são independentes.

O deslocamento total da partícula em relação à origem é então dado, a cada instante, pela soma das variáveis (V.1):

$$B_0 = 0,$$

$$B_n = \sum_{k=1}^n X_k, \quad n \geq 1. \quad \text{V.3}$$

Esta variável aleatória apresenta as seguintes propriedades:

$$EB_k = 0; \quad \text{V.4a}$$

$$EB_k^2 = k; \quad \text{V.4b}$$

$$EB_k B_l = \min\{k, l\}, \quad k \neq l; \quad \text{V.4c}$$

assim, comparando os resultados acima com as expressões (IV.38), (IV.40) e (IV.45), vemos que a variável aleatória B_k apresenta, a menos de um fator constante, as mesmas propriedades de um movimento Browniano. Ademais, os *deslocamentos*

$$X_k = B_k - B_{k-1}, \quad \text{V.5}$$

são *independentes* entre si, conforme é fácil de ver. Para um deslocamento arbitrário, a variância do processo é dada por

$$\text{var}[B_k - B_l] = k - l,$$

onde $k > l$. Esta expressão também esta de conformidade com a teoria exposta no capítulo anterior, mais precisamente com a expressão (IV.40).

Resta-nos obter a função densidade de probabilidade desta variável aleatória para verificarmos finalmente a correlação de um passeio aleatório e o movimento Browniano. Para fins de facilitar o cálculo, re-escalonemos a variável aleatória B_k de modo que esta apresente variância unitária:

$$B_k \longrightarrow \frac{1}{\sqrt{k}} B_k.$$

Com isso, a função característica associada desta variável aleatória será dada por

$$\phi_k(\xi) = E \exp\left\{\frac{\xi}{\sqrt{k}} B_k\right\}. \quad \text{V.6}$$

Após algumas manipulações triviais obtemos, usando-se as relações (V.4a – c), a seguinte expressão:

$$\phi_k(\xi) = \left\{ \frac{1}{2} \exp[\xi x] + \frac{1}{2} \exp[-\xi x] \right\}^{1/x^2}, \quad \text{V.7}$$

onde definimos $x = k^{-1/2}$ para fins de simplificação.

Calculemos agora o limite do logaritmo da expressão acima para $x \rightarrow 0$:

$$\begin{aligned} \lim_{k \rightarrow \infty} \ln \phi_k(\xi) &= \lim_{x \rightarrow 0} \frac{1}{x^2} \ln \left\{ \frac{1}{2} \exp[\xi x] + \frac{1}{2} \exp[-\xi x] \right\} \\ &= \frac{1}{2} \xi^2; \end{aligned} \quad \text{V.8}$$

então,

$$\lim_{k \rightarrow \infty} \phi_k(\xi) = e^{\frac{1}{2} \xi^2}, \quad \text{V.9}$$

ou seja, as funções características deste passeio aleatório convergem para uma função característica gaussiana. Pelo teorema I.59, a relação (V.9) implica em $(1/\sqrt{n})B_n \xrightarrow{D} B$ onde $B \sim N(0,1)$.

Portanto, o passeio aleatório definido pelas variáveis (V.3), tendo como propriedades as expressões (V.4a – c), assume, no limite $k \rightarrow \infty$, a forma de um movimento Browniano. Denominaremos este caso por *Movimento Browniano discreto a Tempos discretos* ou simplesmente *caso discreto*.

De acordo com o exposto no capítulo II, todo passeio aleatório apresenta tanto as propriedades de martingale quanto de Processo de Markov; demonstramos estas propriedades para o caso de um processo multiplicativo, ou seja, cuja forma é $S_n = S_0 u^m d^{n-m}$. No entanto, para passarmos ao caso aditivo, como na discussão presente, basta tomarmos o logaritmo da

variável aleatória multiplicativa. Com isso, concluímos que o movimento Browniano discreto, a tempos discretos, é um martingale e um Processo de Markov.

V.1.2 O Movimento Browniano Contínuo a Tempos Discretos

Uma vez discutido o caso do passeio aleatório, passemos a um modelo semicontínuo, com tempos discretos e espaço amostral arbitrário ([16], [19], [27], [31]). De modo análogo ao passeio aleatório, definiremos variáveis com propriedades semelhantes; isto é, definiremos processos análogos ao deslocamento da partícula, X_k , e ao movimento Browniano B_k . Para tanto, consideremos uma seqüência $\{X_k(\omega)\}_{k=0}^{\infty}$ de variáveis aleatórias gaussianas contínuas definidas sobre um espaço de probabilidade (Ω, \mathcal{A}, P) que satisfaçam as seguintes propriedades

$$EX_k = 0; \quad \text{V.9a}$$

$$EX_k^2 = 1; \quad \text{V.9b}$$

$$EX_k X_l = EX_k EX_l, \quad k \neq l. \quad \text{V.9c}$$

A partir destas variáveis, definimos os processos estocásticos

$$B_0 = 0,$$

$$B_n = \sum_{k=1}^n X_k, \quad n \geq 1. \quad \text{V.10}$$

Pelas expressões acima, é fácil verificar que estes processos estocásticos satisfazem as seguintes relações:

$$EB_k = 0; \quad \text{V.11a}$$

$$EB_k^2 = k; \quad \text{V.11b}$$

$$EB_k B_l = \min\{k, l\}, \quad k \neq l. \quad \text{V.11c}$$

ou seja, idênticas ao caso do passeio aleatório. Também segue-se que, em função das variáveis X_k serem gaussianas, o processo B_k também será normalmente distribuído. Portanto, o processo B_k será um *movimento Browniano contínuo a tempos discretos*, ou simplesmente *caso semicontínuo*.

No caso anterior, vimos que o movimento Browniano discreto a tempos discretos é tanto um martingale quanto um Processo de Markov. Para demonstrarmos que para o caso semicontínuo é válido o mesmo princípio, precisaremos inicialmente construir uma filtração sobre a estrutura de sigma-álgebra deste movimento Browniano; para tanto, usemos a filtração natural associada a esse processo:

$$\mathcal{A}_0^0 := \{\emptyset, \Omega\},$$

$$\mathcal{A}_n^0 := \sigma[X_m : m \leq n].$$

Calculemos agora o valor esperado $E[B_{n+1} | \mathcal{A}_n]$ associado ao movimento Browniano semicontínuo:

$$\begin{aligned} E[B_{n+1} | \mathcal{A}_n] &= E[X_{n+1} + B_n | \mathcal{A}_n] \\ &= EX_{n+1} + B_n \\ &= B_n. \end{aligned}$$

Com isso, vemos que a definição acima do movimento Browniano é também um martingale; resta-nos, a fins de concluir a discussão, provar que o mesmo é um Processo de Markov. Para tanto, definamos primeiramente

$$f(\xi) := E\phi(X_{n+1} + \xi);$$

empreguemos agora a proposição II.8, item *i*:

$$\begin{aligned} E[\phi(B_{n+1}) | \mathcal{A}_n] &= E[\phi(X_{n+1} + B_n) | \mathcal{A}_n] \\ &= f(B_n), \end{aligned}$$

o que demonstra o fato de o modelo Browniano semicontínuo também ser um Processo de Markov.

V.1.3 A Definição Geral do Movimento Browniano – Processos de Wiener

Após termos introduzido os dois casos mais simples, passemos agora à definição mais geral do Movimento Browniano. O caso discreto, tanto a tempos discretos quanto a tempos contínuos, apresentam estruturas idênticas no que se refere às propriedades dos processos estocásticos-base, conforme podemos facilmente verificar por meio das equações (V.2.a – c) e (V.9.a – c), e das equações (V.4.a – c) e (V.11.a – c). O movimento Browniano contínuo a tempos contínuos também deve, portanto, possuir a mesma estrutura ([16], [19], [27], [31]):

V. 1. Definição : Seja (Ω, \mathcal{A}, P) um espaço de probabilidades e $T = [0, \infty)$; então, o processo estocástico

$$B_t(\omega): \Omega \times T \longrightarrow \mathcal{R}$$

é denominado *Movimento Browniano* se

i. $B_0 = 0$;

ii. B_t é uma função contínua de t ;

iii. Para $0 = t_0 < t_1 < \dots < t_n$ os deslocamentos

$$W_{t_k} := B_{t_k} - B_{t_{k-1}}$$

são *independentes*:

$$P\left[\bigcap_{k=1}^n \{W_{t_k} \in A_k\}\right] = \prod_{k=1}^n \{W_{t_k} \in A_k\}.$$

iv. Para $0 \leq s < t$, o deslocamento $B_t - B_s$ é *distribuído normalmente*

$$P[\omega \in \Omega : B_t(\omega) - B_s(\omega) \in A] = \frac{1}{\sqrt{2\pi(t-s)}} \int_A \exp\left\{-\frac{x^2}{t-s}\right\} dx.$$

v. A média e a variância são dadas por:

$$EB_t = 0,$$

$$\text{var}[B_t - B_s] = t - s,$$

onde $0 \leq s < t$.

A demonstração formal que o processo acima existe é dada através do Teorema da extensão de Kolmogorov (ver [2], [15] e [16] para maiores detalhes), que não comentaremos aqui. Na definição acima consideramos, sem perda de generalidade, as constantes que comparecem no capítulo anterior como sendo unitárias.

Com a definição acima associamos definitivamente o fenômeno físico do movimento Browniano à teoria matemática dos Processos Estocásticos. Este suporte matemático à teoria nos permitirá estendê-la a casos bem mais abrangentes, e será exatamente esta identificação que nos possibilitará, no capítulo VII, generalizar o conceito de Movimento Browniano.

Conforme vimos nos casos anteriores, o movimento Browniano também é um martingale e um Processo de Markov. No caso presente, também podemos afirmar este fato: ([16], [19], [27], [31]):

V.2. Teorema : O Movimento Browniano, conforme a Definição V.1, é também um martingale.

V.3. Teorema : O Movimento Browniano, conforme a Definição V.1, é também um Processo de Markov.

A definição V.1 estabelece as restrições que um dado processo deve satisfazer de modo que este seja um movimento Browniano. Tal definição nos permite ‘criar’ novos processos que tenham as características de um movimento Browniano. Consideremos, por exemplo, o item v. da definição V.1; como o lado direito da equação da variância apresenta um termo dependente do tempo, e de natureza não-aleatória, podemos representar o deslocamento $B_t - B_s$ em função de um termo aleatório, ζ , e a diferença $t - s$:

$$B_t - B_s = \zeta \sqrt{t - s} . \quad \text{V.12}$$

Deste modo, tomando-se a variância da expressão acima veremos que a variável aleatória ζ deverá satisfazer $E\zeta^2 = 1$, de modo que o item v. da definição V.1 seja obedecido. Se tomarmos a variável ζ como sendo gaussiana, e tendo valor esperado igual a zero, então a distribuição de probabilidades de $B_t - B_s$ também será gaussiana. Com isso, construímos, *a priori*, um movimento Browniano, tendo simplesmente adotado algumas regras impostas pela definição V.1.

Podemos estender ainda mais a equação (V.12) de modo a contemplar processos dotados de valores esperados não-nulos. Basta, para tanto, adicionar um termo $\mu\Delta t$ no lado direito da expressão (V.12). Este termo, conforme discutimos na seção II.2, surge da própria definição da variável aleatória associada a uma árvore binomial, e é necessária sua presença de modo que este processo discreto convirja a um movimento Browniano no limite infinito.

Com base no discutido acima, introduzamos então a seguinte [31]

V.4. Definição : Seja $\{\zeta\}$ uma seqüência de variáveis aleatórias gaussianas que satisfaçam

$$E\zeta = 0 ,$$

$$E\zeta^2 = 1 ;$$

Então, o processo estocástico

$$X_{t+\Delta t} = X_t + \mu\Delta t + \zeta \sqrt{\Delta t} , \quad \text{V.13}$$

denominado *Processo de Wiener*, é um movimento Browniano com valor esperado $\mu\Delta t$ e variância Δt .

Os processos de Wiener, conforme definidos acima, constituem a base sobre a qual construiremos toda a teoria dos processos estocásticos nas discussões que se seguem. Também veremos que, a título de simulações computacionais, os processos de Wiener são facilmente implementáveis, trazendo resultados muito satisfatórios.

Vamos explorar um pouco mais as propriedades matemáticas dos processos de Wiener. Tomando-se Δt tendendo a zero, verificamos que tais processos são funções contínuas do tempo,

$$\lim_{\Delta t \rightarrow 0} X_{t+\Delta t} = X_t, \quad \text{V.14}$$

conforme a definição de movimento Browniano exige. Entretanto, tais processos são não-diferenciáveis em todos os pontos:

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} \frac{\Delta X}{\Delta t} &= \mu + \xi \Delta t^{-1/2} \\ &= \infty. \end{aligned} \quad \text{V.15}$$

A expressão (V.15) nos remonta a uma equação diferencial ordinária. Uma pergunta pertinente que podemos colocar a esta altura em nossa discussão é: existem equações diferenciais para processos estocásticos? A resposta é afirmativa, a qual introduziremos na próxima seção.

V.2 A Integral de Itô

Já discutimos, até o presente momento, derivadas de processos estocásticos. Sempre as tomamos em relação a parâmetros não-probabilísticos, como, por exemplo, em relação ao tempo. Entretanto, expressões como

$$df = \frac{d}{dX_t} f[X_t] dX_t, \quad \text{V.16}$$

onde X_t é um processo estocástico qualquer, ainda não foram discutidas; nestas, há uma derivação explícita em relação ao processo X_t . Para podermos efetuar a operação (V.16), é necessário primeiramente introduzirmos o conceito de integral cujo volume de integração seja um processo estocástico. Para facilitar nossa discussão, escrevamos a parte estocástica da equação (V.13) como

$$W_{t+\Delta t} = W_t + \zeta \sqrt{\Delta t}. \quad \text{V.17}$$

Agora, antes de introduzirmos a idéia de integral estocástica, rememoremos o cálculo integral usual. Consideremos uma função contínua a valores reais $F(x)$, definida sobre um intervalo (x_0, x) . Dividindo-se este último em n intervalos de comprimento $\Delta x = (x - x_0)/n$, e introduzindo-se as diferenças

$$\Delta F(x_k) = F(x_k + \Delta x) - F(x_k), \quad \text{V.18}$$

segue-se que a integral de Stieltjes do $dF(x)$ é dada pelo somatório [6]

$$\int_{x_0}^x dF(z) = \lim_{n \rightarrow \infty} \sum_{k=1}^n \Delta F(x_k). \quad \text{V.19}$$

No caso de um processo estocástico, dispomos de todos os ingredientes acima: o processo contínuo W_t , um intervalo temporal (t_0, t) , particionado em n subintervalos de comprimento $\Delta t = (t - t_0)/n$, e um conjunto de deslocamentos

$$\Delta W_{t_k} = W_{t_k + \Delta t} - W_{t_k} . \quad \text{V.20}$$

Usaremos na definição da integral estocástica o modo de convergência no p -ésimo momento, que se revelou ser o mais conveniente. De acordo com a discussão promovida na subseção I.2.4. do capítulo I, dizemos que uma seqüência $\{X_n(\omega)\}_{n=0}^{\infty}$ converge a um limite X no p -ésimo momento [6] se

$$E|X_n - X|^p \rightarrow 0 \quad \text{no limite } n \rightarrow \infty . \quad \text{V.21}$$

Para expressarmos este modo em forma de limite introduzimos o conceito de p -limite:

$$E|X_n - X|^p \rightarrow 0 \quad \Rightarrow \quad p \lim_{n \rightarrow \infty} X_n = X . \quad \text{V.22}$$

Em particular, para nossas discussões, é suficiente tomarmos $p=2$, uma vez que são os dois primeiros momentos que definem uma distribuição gaussiana. Com isso, surge o conceito de *limite médio quadrático*,

$$E|X_n - X|^2 \rightarrow 0 \quad \Rightarrow \quad ms \lim_{n \rightarrow \infty} X_n = X . \quad \text{V.23}$$

Por meio destes conceitos podemos finalmente introduzir a expressão da integral estocástica, analogamente ao que fizemos em (V.19) [6] :

V. 5. Definição : Seja W_t um processo contínuo definido por

$$W_{t+\Delta t} = W_t + \zeta \sqrt{\Delta t} ;$$

a *integral estocástica de Itô* é definida pela expressão

$$\int_{t_0}^t dW_s = ms \lim_{n \rightarrow \infty} \sum_{k=0}^n \Delta W_{t_k} . \quad \text{V.24}$$

Uma vez definida a expressão geral para uma integral estocástica, analisemos algumas de suas propriedades. Isto é, passemos ao cálculo de expressões do tipo

$$\int_{t_0}^t \Phi(s) dW_t , \quad \text{V.25}$$

onde $\Phi(s)$ é uma função contínua de t . O caso mais elementar ocorre quando $\Phi(s) = \kappa$, onde κ é uma constante real; com isso, a partir da definição V.5, obtemos

$$\int_{t_0}^t \kappa dW_s = \kappa(W_t - W_0), \quad \text{V.26}$$

como era de se esperar. Deste modo, se tomarmos o processo estocástico (V.13) como a diferencial

$$dX_t = \mu dt + dW_t, \quad \text{V.27}$$

sua integral será dada pela expressão

$$X_t = X_0 + \mu(t - t_0) + W_t - W_{t_0}, \quad \text{V.28}$$

ou seja, compatível com uma integral de Stieltjes.

As próximas expressões a serem analisadas são as que concernem potências do processo W_t ; comecemos com o caso mais simples: $\Phi(s) = W_s$. Da definição V.5 obtemos

$$\int_{t_0}^t W_s dW_s = ms \lim_{n \rightarrow \infty} \sum_{k=1}^n W_k \Delta W_k; \quad \text{V.29}$$

usando-se a identidade

$$W_k \Delta W_k = \frac{1}{2} [W_{k+1}^2 - W_k^2 - \Delta W_k^2], \quad \text{V.30}$$

segue-se que

$$\sum_{k=1}^n W_k \Delta W_k = \frac{1}{2} [W_n^2 - W_0^2] - \frac{1}{2} \sum_{k=1}^n \Delta W_k^2. \quad \text{V.31}$$

Resta-nos calcular o último termo do lado direito de (V.31). Para tanto, observemos a seguinte expressão:

$$E \left[\sum_{k=1}^n \Delta W_k^2 \right] = t - t_0; \quad \text{V.32}$$

este resultado não se altera se tomarmos o limite $n \rightarrow \infty$; com isso, o limite do valor esperado de ΔW_k^2 é igual à diferença do instante final e o inicial. Podemos então suspeitar que o limite

$$\lim_{n \rightarrow \infty} E \left[\sum_{k=1}^n (\Delta W_k^2 - \Delta t_k) \right]^2$$

se anule; de fato, expandindo-se o valor esperado acima, obtemos

$$E \left[\sum_{k=1}^n (\Delta W_k^2 - \Delta t_k) \right]^2 = \sum_{k=1}^n E \Delta W_k^4 + \sum_{k>l} E [\Delta W_k^2 \Delta W_l^2] - 2 \sum_{k,l} \Delta t_l E W_k^2 + \sum_{k,l} \Delta t_k \Delta t_l; \quad \text{V.33}$$

considerando-se as relações [6]

$$\begin{aligned} i. E\Delta W_k^4 &= 3E[\Delta W_k^2]^2, \\ ii. E[\Delta W_k^2 \Delta W_l^2] &= E\Delta W_k^2 E\Delta W_l^2, \end{aligned} \tag{V.34}$$

segue-se que

$$E\left[\sum_{k=1}^n (\Delta W_k^2 - \Delta t_k)\right]^2 = 2\sum_{k=1}^n \Delta t_k^2. \tag{V.35}$$

Tomando-se o limite da equação acima, obtemos

$$\begin{aligned} \lim_{n \rightarrow \infty} E\left[\sum_{k=1}^n (\Delta W_k^2 - \Delta t_k)\right]^2 &= \lim_{\Delta t \rightarrow 0} \sum_k \Delta t_k \\ &= 0. \end{aligned} \tag{V.36}$$

Deste modo,

$$ms \lim_{n \rightarrow \infty} \sum_{k=1}^n \Delta W_k^2 = t - t_0, \tag{V.37}$$

o que, a partir das expressões (V.29) e (V.31), fornece,

$$\int_{t_0}^t W_s dW_s = \frac{1}{2} [W_t^2 - W_0^2] - \frac{1}{2} (t - t_0). \tag{V.38}$$

O resultado (V.38) difere do esperado para uma integral de Stieltjes devido ao termo $1/2(t - t_0)$; tal termo é justificável devido ao fato de $|dW_t|$ ser de ordem $O(\sqrt{t})$ quase sempre, e por isso termos de segunda ordem não se anulam em ΔW_t quando tomamos o limite médio quadrático.

Antes de procedermos ao cálculo da integral da função $\Phi(s) = W_s^n$, façamos uma pausa oportuna para discutirmos o seguinte ponto: pela definição (V.17), considerando-se Δt infinitesimal, vemos que dW^2 é igual, a menos de um fator aleatório, a dt ; com isso, podemos supor que uma integral estocástica cujo volume de integração seja dW^2 possa ser reduzida a uma integral ordinária integrada em relação a t . Para melhor avaliarmos esta situação, consideremos a integral

$$\int_{t_0}^t \Phi(s) dW_s^2, \tag{V.39}$$

onde $\Phi(s)$ é uma função contínua de t , dotada de propriedades que abordaremos em seguida. Da definição V.5 segue-se que

$$\int_{t_0}^t \Phi(s) dW_t^2 = ms \lim_{n \rightarrow \infty} \sum_{k=1}^n \Phi_k \Delta W_k^2; \quad \text{V.40}$$

o fato de que o termo quadrático ΔW_k^2 ser proporcional a Δt_k nos sugere que o limite

$$\lim_{n \rightarrow \infty} E \left[\sum_{k=1}^n \Phi_k (\Delta W_k^2 - \Delta t_k) \right]^2 \quad \text{V.41}$$

seja nulo. De fato, analogamente ao limite discutido anteriormente, obtemos

$$E \left[\sum_{k=1}^n \Phi_k (\Delta W_k^2 - \Delta t_k) \right]^2 = E \left[\sum_{k=1}^n \Phi_k^2 (\Delta W_k^2 - \Delta t_k)^2 + 2 \sum_{k>l} \Phi_k \Phi_l (\Delta W_k^2 - \Delta t_k) (\Delta W_l^2 - \Delta t_l) \right] \quad \text{V.42}$$

Se a função $\Phi(s)$ satisfizer as propriedades

i. $\Phi(s)$ é estatisticamente independente de $W_t - W_s$ para $s < t$, e

$$\text{ii. } \int_{t_0}^t \Phi(s)^2 ds < \infty, \quad \text{V.43}$$

então, da equação (V.42) obteremos as expressões

$$E \left[\sum_{k=1}^n \Phi_k^2 (\Delta W_k^2 - \Delta t_k)^2 \right] = 2 \sum_{k=1}^n E \Phi_k^2 \Delta t_k^2, \quad \text{V.44}$$

e

$$E \left[\sum_{k>l} \Phi_k \Phi_l (\Delta W_k^2 - \Delta t_k) (\Delta W_l^2 - \Delta t_l) \right] = 0. \quad \text{V.45}$$

Esta função $\Phi(s)$ recebe o nome de *função de não-antecipação*.

Com isso,

$$\begin{aligned} \lim_{n \rightarrow \infty} E \left[\sum_{k=1}^n \Phi_k (\Delta W_k^2 - \Delta t_k) \right]^2 &= \lim_{\Delta t_k \rightarrow \infty} \sum_k E \Phi_k^2 \Delta t_k^2 \\ &= 0, \end{aligned} \quad \text{V.46}$$

o que resulta em

$$ms \lim_{n \rightarrow \infty} \sum_{k=1}^n \Phi_k \Delta W_k^2 = \lim_{n \rightarrow \infty} \sum_{k=1}^n \Phi_k \Delta t_k. \quad \text{V.47}$$

Portanto,

$$\int_{t_0}^t \Phi(s) dW_s^2 = \int_{t_0}^t \Phi(s) ds, \quad \text{V.48}$$

o que implica em

$$dW_t^2 = dt. \quad \text{V.49}$$

Vemos, pela discussão acima, que a integral sobre dW^2 de uma função de não-antecipação $\Phi(s)$ se reduz a uma integral de Riemann. Podemos nos perguntar o que ocorre quando tomamos potências de ordem superior de dW ; tomemos, então, como exemplo, o cálculo da seguinte integral:

$$\int_{t_0}^t \Phi(s) dW_s^3, \quad \text{V.50}$$

onde $\Phi(s)$ é uma função de não-antecipação. Pelas definições e propriedades analisadas anteriormente, a integral acima se escreve como

$$\int_{t_0}^t \Phi(s) dW_s^3 = ms \lim_{n \rightarrow \infty} \sum_{k=1}^n \Phi_k \Delta t_k \Delta W_k.$$

Um exame minucioso da equação acima nos demonstra que seu lado direito se anula ao tomarmos o limite médio quadrático; de fato,

$$\begin{aligned} \lim_{n \rightarrow \infty} E \left[\sum_{k=1}^n \Phi_k \Delta t_k \Delta W_k \right]^2 &= \lim_{n \rightarrow \infty} E \left\{ \sum_{k=1}^n \Phi_k^2 \Delta t_k^2 \Delta W_k^2 + 2 \sum_{k>l} \Phi_k \Phi_l \Delta W_k \Delta W_l \Delta t_k \Delta t_l \right\} \\ &= \lim_{n \rightarrow \infty} \sum_{k=1}^n E \Phi_k^2 \Delta t_k^4 \\ &= 0. \end{aligned}$$

Com isso, vemos que potências superiores de dW se anulam:

$$dW_t^{2+n} = 0. \quad \text{V.51}$$

A identidade acima nos permite calcular facilmente a integral de $\Phi(s) = W_s^n$; da expressão

$$\int_{t_0}^t W_s^n dW_s = \int_{t_0}^t d[W_s^{n+1}] - \int_{t_0}^t W_s d[W_s^n], \quad \text{V.52}$$

calculamos

$$\begin{aligned}
d[W_s^n] &= [W_s + dW_s]^2 - W_s^n \\
&= nW_s^{n-1}dW_s + \frac{n(n-1)}{2}W_s^{n-2}dW_s^2.
\end{aligned} \tag{V.53}$$

Portanto, o resultado final da integral (V.52) é dado por

$$\int_{t_0}^t W_s^n dW_s = \frac{1}{n+1} [W_t^{n+1} - W_{t_0}^{n+1}] - \frac{n(n-1)}{2(n+1)} \int_{t_0}^t W_s^{n-2} ds. \tag{V.54}$$

Com esses conceitos, concluímos nossa discussão sobre as integrais estocásticas. Estas consistirão em uma poderosa ferramenta na discussão das equações diferenciais estocásticas.

V.3 Equações Diferenciais Estocásticas

Retomemos a discussão iniciada na seção V.1. Nosso ensejo ao introduzir o conceito de integral estocástica era o de resolver uma equação do tipo (V.16) [6]; com o ferramental desenvolvido na seção anterior, estamos aptos a solucionar não somente equações deste tipo, mas também equações de ordens superiores. Consideremos, pois, um processo estocástico diferenciável da forma

$$X_t = X[t, W_t]; \tag{V.55}$$

efetuando-se uma expansão em série de Taylor, obtemos

$$dX_t = \frac{\partial X_t}{\partial t} dt + \frac{1}{2} \frac{\partial^2 X_t}{\partial t^2} dt^2 + \dots + \frac{\partial X_t}{\partial W_t} dW_t + \frac{1}{2} \frac{\partial^2 X_t}{\partial W_t^2} dW_t^2 + \dots. \tag{V.56}$$

A prescrição (V.51) nos fornece

$$dX_t = \left[\frac{\partial X_t}{\partial t} + \frac{1}{2} \frac{\partial^2 X_t}{\partial W_t^2} \right] dt + \frac{\partial X_t}{\partial W_t} dW_t. \tag{V.57}$$

Com isso, obtivemos a equação estocástica mais geral possível, uma vez que termos de ordem superior se anulam devido a identidade (V.51). A equação (V.57) pode ser reescrita, para um caso geral, como

$$dX_t = \xi(t, X_t) dt + \zeta(t, X_t) dW_t. \tag{V.58}$$

Assumiremos que $\xi(t, X_t)$ e $\zeta(t, X_t)$ sejam contínuas em (t, x) e sejam contínuas em x no sentido de Lipschitz, isto é

$$|\xi(t, x) - \xi(t, y)| \leq L|x - y| \text{ e } |\zeta(t, x) - \zeta(t, y)| \leq L|x - y|, \tag{V.59}$$

para quaisquer valores de t , x e y . Com isso, considerando-se as condições iniciais (t_0, x_0) , a solução da equação diferencial estocástica (V.58) será dada pelo processo estocástico $\{X_t\}_{t \geq t_0}$:

$$X_t = x_0 + \int_{t_0}^t \xi(s, X_s) ds + \int_{t_0}^t \zeta(s, X_s) dW_s. \quad \text{V.60}$$

Pela forma do processo acima observamos que, se conhecermos a trajetória do movimento Browniano até o instante t , poderemos determinar a evolução temporal de X_t . Em outras palavras, o processo $\{X_t\}_{t \geq t_0}$ é adaptado à filtração gerada pelo movimento Browniano.

Antes de darmos prosseguimento às possíveis soluções de interesse para a equação (V.60), é importante que a estendamos para o caso de funções estocásticas da forma $f = f[X_t]$. Usando-se a forma diferencial (V.58), obtemos

$$\begin{aligned} df[X_t] &= f[X_t + dX_t] - f[X_t] \\ &= f'[X_t] dX_t + \frac{1}{2} f''[X_t] dX_t^2 + \dots \\ &= f'[X_t] [\xi(t, X_t) dt + \zeta(t, X_t) dW_t] dX_t + \frac{1}{2} f''[X_t] dW_t^2, \end{aligned}$$

pois os termos cruzados $dt dW$ e os de ordem superior em dW se anulam. O resultado final destas manipulações é a fórmula

$$df[X_t] = \left\{ f'[X_t] \xi(t, X_t) + \frac{1}{2} f''[X_t] \zeta(t, X_t)^2 \right\} dt + f'[X_t] \zeta(t, X_t) dW_t.$$

Se a função estocástica f apresentar dependência explícita do tempo, $f = f[t, X_t]$, a equação acima assume a forma

$$df[X_t] = \left\{ \frac{\partial f}{\partial t} + f'[X_t] \xi(t, X_t) + \frac{1}{2} f''[X_t] \zeta(t, X_t)^2 \right\} dt + f'[X_t] \zeta(t, X_t) dW_t, \quad \text{V.61}$$

a qual denominamos *fórmula* ou *Lema de Itô* [6]. A fórmula de Itô é assim uma generalização imediata da equação (V.57) para uma função estocástica contínua.

Desenvolvemos, assim, todo um ferramental para solucionarmos equações diferenciais estocásticas. Analisaremos, a título de desenvolvimentos posteriores, os dois casos mais imediatos decorrentes da equação (V.60): o *processo de Wiener* e o *movimento Browniano geométrico*.

V.3.1 O Processo de Wiener

Retomemos a equação (V.60). As expressões mais elementares que as funções $\xi(t, X_t)$ e $\zeta(t, X_t)$ podem assumir ocorrem quando estas são funções constantes. Tomando-se

$$\begin{aligned}\xi(t, X_t) &= \mu, \\ \zeta(t, X_t) &= 1,\end{aligned}\tag{V.62}$$

e assumindo-se a condição inicial $X_{t_0} = x_0$, obtemos como solução da equação (V.60) a expressão

$$X_t = x_0 + \mu(t - t_0) + [W_t - W_{t_0}],\tag{V.63}$$

ou seja, um processo de Wiener. Se $\mu = 0$, o processo acima coincide com um movimento Browniano sem o termo de arrasto.

Podemos solucionar a equação (V.60) sujeita às condições (V.62) de uma forma ligeiramente diferente. Para tanto, basta identificarmos a equação (V.57) com (V.58) e usar as funções (V.62). Obtemos assim as equações diferenciais

$$\frac{\partial X_t}{\partial t} + \frac{1}{2} \frac{\partial^2 X_t}{\partial W_t^2} = \mu\tag{V.64a}$$

e

$$\frac{\partial X_t}{\partial W_t} = 1.\tag{V.64b}$$

A segunda equação em (V.64) nos mostra que sua segunda derivada é nula; com isso, a primeira equação acima se torna igual a

$$\frac{\partial X_t}{\partial t} = \mu.\tag{V.65}$$

Solucionando-se ambas as equações, obtemos resultado idêntico a (V.63).

V.3.2 O Movimento Browniano Geométrico

O segundo caso mais elementar que as funções $\xi(t, X_t)$ e $\zeta(t, X_t)$ podem assumir ocorre quando estas são proporcionais ao próprio processo X_t :

$$\begin{aligned}\xi(t, X_t) &= \mu X_t, \\ \zeta(t, X_t) &= \sigma X_t.\end{aligned}\tag{V.66}$$

Identificando-se novamente as equações (V.57) e (V.58), e substituindo-se nestas as expressões (V.66), obtemos as seguintes equações diferenciais:

$$\frac{\partial X_t}{\partial t} + \frac{1}{2} \frac{\partial^2 X_t}{\partial W_t^2} = \mu X_t \quad \text{V.67a}$$

e

$$\frac{\partial X_t}{\partial W_t} = \sigma X_t. \quad \text{V.67b}$$

Perfazendo-se a separação de variáveis $X_t = \phi(t)\chi(W_t)$ e substituindo-a na equação (V.67b), obtemos a solução para $\chi(W_t)$:

$$\chi(W_t) = \kappa \exp\{\sigma W_t\}, \quad \text{V.68}$$

onde κ é uma constante. A derivada segunda desta expressão resulta em

$$\frac{\partial^2 \chi}{\partial W_t^2} = \sigma^2 \chi; \quad \text{V.69}$$

portanto, substituindo este resultado na equação (V.67a), e reagrupando os termos, obtemos a equação diferencial

$$\frac{d\phi}{dt} = \left(\mu - \frac{1}{2} \sigma^2 \right) \phi, \quad \text{V.70}$$

cuja solução é dada por

$$\phi(t) = \lambda \exp\left\{ \left(\mu - \frac{1}{2} \sigma^2 \right) t \right\}, \quad \text{V.71}$$

onde λ é uma constante. Usando-se a condição inicial $X_{t_0} = x_0$, obtemos como solução final das equações (V.67a – b):

$$X_t = x_0 \exp\left\{ \sigma(W_t - W_{t_0}) + \left(\mu - \frac{1}{2} \sigma^2 \right) (t - t_0) \right\}. \quad \text{V.72}$$

O processo (V.72) é denominado *movimento Browniano geométrico*. Tal processo estocástico desempenhará um importante papel em nossa discussão acerca da teoria de Black & Scholes.

CAPÍTULO VI

**APLICAÇÕES DO MOVIMENTO BROWNIANO: A
TEORIA DE BLACK & SCHOLES****VI.1 Introdução**

Desenvolvemos nos capítulos anteriores todo o fundamento do movimento Browniano, desde seu aspecto puramente físico, nas teorias de Einstein e Langevin, até seus aspectos matemáticos mais profundos, como na teoria de Itô. Ao longo do tempo, desde as primeiras especulações envolvendo fenômenos biológicos até os tempos atuais, o movimento Browniano tem demonstrado ser uma das principais peças do jogo matemático da natureza. Na área de finanças ([1], [17], [22], [30]), por exemplo, as aplicações do movimento Browniano têm sido tão extensas a ponto de revolucionar os conceitos de investimentos conhecidos até então. O mercado tornou-se tão complexo que, para uma boa compreensão de seu comportamento, o uso de ferramentas matemáticas avançadas, como a teoria dos Processos Estocásticos, tornou-se inevitável. Para entendermos o porquê da crescente complexidade que o mercado financeiro está atingindo, precisamos introduzir inicialmente alguns importantes conceitos da teoria financeira.

Basicamente, *investir* ([22], [30]) significa alocar uma certa quantidade de recursos de modo a obtermos benefícios futuros. Isto é, ao alocarmos, por exemplo, um certo montante de dinheiro em uma dada aplicação financeira, esperamos que, dentro de um certo período de tempo, este montante aumente com a operação financeira efetuada. Em outras palavras, esperamos que o *retorno do investimento* [22] seja maior que o montante inicial aplicado. No entanto, apesar de existirem investimentos onde o ganho é certo (como, por exemplo, a caderneta de poupança), há outros onde há uma *possibilidade de perda* ([22], [30]). Isto significa que, ao final de um determinado período, o montante aplicado pode ser perdido. Este é o caso, por exemplo, do investimento em ações (que discutiremos mais adiante), que permitem lucros altos e também perdas altas. É exatamente essa possibilidade de perda que denominamos de *risco*. Mais exatamente ainda, podemos quantificar o risco como sendo igual à *probabilidade de fracasso de um dado investimento* [30].

Os conceitos de risco e retorno de um investimento são os dois princípios básicos de toda a teoria financeira. Podemos dizer que o lema de qualquer profissional de mercado é: maximizar seu retorno e minimizar os riscos envolvidos. O controle do risco, aliás, tem sido uma das principais tendências da teoria financeira moderna. Isso porque os mercados têm, nas últimas décadas, apresentado um grau de risco cada vez maior. Mas, por que este fenômeno do crescimento dos riscos financeiros ocorre?

Retrocedamos um pouco no tempo [1]. Em 1944, foi firmado o acordo de Bretton Woods, onde fora criado o Fundo Monetário Internacional (FMI), e fora estipulada a política internacional monetária de câmbio fixo para o pós-guerra. De acordo com esta política, as principais moedas do mundo passavam a ter paridade fixa ao dólar, ao passo que este teria paridade fixa com o ouro. Este sistema de câmbio fixo dava ao mercado uma certa rigidez em suas transações. No entanto, em dezembro de 1971, os Estados Unidos, devido a fortes pressões internacionais, solicitaram uma revisão deste acordo, culminando com o acordo

firmado no Smithsonian Institute, em Washington. Neste acordo internacional foram estabelecidas novas regras em relação à paridade das moedas em relação ao dólar, o que conferiu uma maior liberdade o sistema de câmbio internacional. Com isso, grandes mudanças na paridade das moedas ocorreram, o que acarretou, por sua vez, fortes oscilações nas taxas de juros.

Aliado a este cenário internacional acrescentam-se as duas crises mundiais de petróleo das décadas de 70 e 80, o que contribuíram ainda mais para um ambiente internacional instável, propício a maiores riscos.

Portanto, em face desta conjuntura econômica instável, os bancos passaram a ser mais exigidos no sentido de oferecerem instrumentos financeiros que conferissem a seus clientes maior proteção em seus investimentos. Com isso, o *mercado de derivativos* tomou impulso, promovendo aos investidores um maior número de possibilidades de investimentos.

Podemos então explicar a crescente complexidade dos mercados através da crescente exposição destes ao risco. Tal exposição demandou não somente a criação de meios mais eficientes de investimento, como também um controle maior sobre os riscos de tais instrumentos. Como o conceito de risco está associado a fatores aleatórios (vide a crise do petróleo), o uso de ferramentas da teoria da Probabilidade passou a ser fundamental para quantificá-lo. Ademais, devido ao fato dos valores dos instrumentos financeiros (sejam eles ações, derivativos, etc.) variarem de forma aleatória com o tempo (em função da conjuntura macroeconômica global), tais grandezas passaram a ser modeladas via processos estocásticos, o que demandou no crescente emprego de técnicas matemáticas avançadas para a análise dos dados financeiros.

Uma vez esclarecidos os fatores que promovem a aproximação da Matemática com a Teoria das Finanças, podemos passar à discussão dos principais instrumentos financeiros em uso. Iniciaremos com a explanação de alguns elementos básicos da teoria financeira, fundamentais para podermos aplicar consistentemente as técnicas matemáticas adequadas.

VI.2 Finanças – Conceitos Básicos

VI.2.1 Taxas de Juros

O conceito de *juros* ([22], [30]) está tão presente em nossa vida diária, que, mesmo intuitivamente, o prevemos quando planejamos nossa economia doméstica. Levamos em conta os juros quando efetuamos uma compra a prazo, quando tomamos dinheiro emprestado ou mesmo quando investimos nosso dinheiro na caderneta de poupança. Poderíamos listar tantas outras operações financeiras usuais onde o conceito de juros apareceria normalmente. Com isso, vemos que os juros consistem em um dos conceitos mais importantes da Teoria de Finanças. Mas, do ponto de vista técnico, o que são os juros?

Antes de conceituarmos de forma mais técnica o conceito de juros, vamos analisar com um pouco mais de detalhe os três exemplos dados acima. Notemos primeiramente que nas três operações mencionadas está implícito o conceito de *tempo futuro*: quando compramos a prazo, pretendemos dividir o preço da mercadoria em parcelas a serem pagas no *futuro*, quando tomamos dinheiro emprestado temos a intenção de empregá-lo no presente e o devolvermos no *futuro*, e quando investimos em caderneta de poupança temos a intenção de resgatar um capital maior no *futuro*. Por conseguinte, o conceito de juros está intimamente relacionado ao *tempo* necessário para efetuarmos uma dada operação financeira. Em segundo lugar, é fácil observar que, no primeiro caso, os juros associados correspondem ao valor adicional ao preço que pagamos da mercadoria; no segundo, correspondem ao valor adicional

ao montante de dinheiro emprestado; e, no terceiro, correspondem ao valor adicional ganho pelo rendimento da caderneta de poupança. Portanto, nos três exemplos, os juros consistem em um *valor adicional* ao valor *inicial* investido ou a ser pago.

Com o discutido acima podemos agora introduzir a expressão matemática que define o conceito de juros. Para tanto, tomemos o primeiro exemplo dado acima; as generalizações posteriores serão evidentes. O preço à vista consiste, no *tempo presente*, o *valor inicial* no âmbito da discussão acima, o qual denominaremos de *valor presente*, P ; o preço à prazo, a ser pago no futuro, consiste no *valor futuro*, F ; e, os juros, consistem no *valor adicional* ao valor presente, denominado pela letra J . Portanto, a relação entre estas três variáveis é dada por

$$F = P + J . \quad \text{VI.1}$$

É fácil verificar que os conceitos de valor presente e valor futuro, bem como a fórmula (VI.1), se aplicam perfeitamente aos dois outros exemplos mencionados.

Uma vez tendo estabelecido melhor as variáveis que compõem os juros, podemos agora analisar as razões de sua existência. Para tanto, recorramos a um dos exemplos acima. Consideremos um comerciante que seja proprietário de uma livraria. Como as vendas podem ser feitas tanto à vista quanto à prazo, seu faturamento será composto pelos montantes de vendas à vista, e pelas parcelas de vendas à prazo que vencem no mês. Considerando-se as todas as parcelas de vendas à prazo que vencerão nos meses seguintes, suponhamos que os valores faturados em um dado mês seja:

Vendas à Vista :	10.490,90
Parcelas de Compras à Prazo recebidas no mês :	5.945,00
SUBTOTAL :	16.435,90
Parcelas de Compras à Prazo a ser recebida nos próximos meses :	18.500,90
TOTAL GERAL:	34.936,80

Tabela. VI.1 : Faturamento do mês e previsto para o exemplo da livraria.

Neste exemplo, vemos que no mês o faturamento do comerciante será de \$ 16.435,90; entretanto, caso todas as vendas fossem à vista, faturaria \$ 34.936,80, ou seja, 47% a mais! Com isso, o comerciante deixa de dispor \$18.500,90 neste mês; ou seja, parte de seu capital fica imobilizado, não podendo, por exemplo, investi-lo na compra de novos livros, na contratação de novos funcionários ou mesmo investi-lo no mercado financeiro. Não devemos nos esquecer que, em função da inflação, das variações da taxa do dólar, etc, o valor do montante a ser recebido pode se desvalorizar no futuro, acarretando em redução das margens de lucro e até mesmo em prejuízos. Eliminar as vendas a prazo não seria interessante, uma vez que estas representam parte considerável das vendas de livros. Como solucionar então o problema?

É neste contexto que entram os juros. Como o comerciante ficará com parte do capital imobilizado para *financiar* as compras dos clientes, o *custo* desta imobilização é repassado

aos clientes através da adição de valores aos preços de venda a prazo dos livros. São estes os valores adicionais que entendemos sob a forma de juros.

Com isso, os *juros consistem na remuneração a ser paga pelo uso do capital*. O cliente da livraria, ao financiar seus livros à prazo, paga com os juros pelo uso do capital imobilizado do comerciante. No caso da pessoa que solicita um empréstimo, esta paga em forma de juros pelo uso do capital da financiadora. No caso da caderneta de poupança, o banco remunera o investidor com os juros decorrentes da aplicação¹.

Retornemos agora à expressão matemática dos juros. Podemos expressar os juros em sua forma percentual; para tanto, basta dividirmos a equação (VI.1) por P :

$$\begin{aligned}\frac{F}{P} &= 1 + \frac{J}{P} \\ &= 1 + r,\end{aligned}\tag{VI.2}$$

onde definimos a *taxa de juros*:

$$r := \frac{J}{P}.\tag{VI.3}$$

Portanto:

$$F = P(1 + r).\tag{VI.4}$$

A equação (VI.4) demonstra então qual será o valor futuro de um investimento em um dado período. Esta equação representa um *regime de capitalização discreta*, pois estamos considerando períodos bem distintos. Podemos generalizar a equação acima para os casos onde os preços dos ativos variam continuamente; esta expressão facilitará consideravelmente os cálculos futuros. Consideremos que o valor do ativo seja uma função contínua do tempo, $X(t)$, cujo valor inicial é igual a $P = X(t_0)$; para facilitar, assumamos a taxa de juros (VI.3) como sendo constante. Tomando-se um intervalo de tempo Δt , podemos representar a equação (VI.3) como sendo igual a

$$r = \frac{X(t + \Delta t) - X(t)}{X(t)} := \frac{\Delta X}{X}.\tag{VI.5}$$

Tomando-se o limite $\Delta t \rightarrow 0$, obtemos uma equação diferencial cuja solução é dada por

$$X(t) = Pe^{rt}.\tag{VI.6}$$

Denominamos este regime de capitalização como sendo *contínuo*. Neste regime de capitalização podemos introduzir fatores que permitem uma análise de sistemas mais complexos. Por exemplo, podemos considerar um modelo onde a taxa de juros não é mais constante, mas sim uma função do tempo, e assim por diante.

¹ Neste caso, poderíamos nos perguntar qual seria então o ganho das instituições financeiras em captar recursos de terceiros e premiá-los em forma de juros. O ganho se consegue através da aplicação de grandes volumes de recursos de capitais de terceiros em ativos que remuneram à uma taxa de juros *maior* que a taxa paga pelo rendimento da caderneta de poupança.

VI.2.2 Arbitragem e a Hipótese do Mercado Eficiente

Para podermos operar em um mercado financeiro, é fundamental que disponhamos do maior número de informações possíveis a respeito dele. É absolutamente essencial a um operador de mercado que acompanhe de perto as variações dos índices financeiros (taxa do dólar, índices de bolsas de valores, etc) de modo que possa tomar as decisões mais eficientes. Na Era da Informática podemos obter estas informações de maneira mais fácil que em épocas anteriores, pois os intervalos de tempo entre a ocorrência de um dado fenômeno no mercado e sua repercussão é ínfimo, sendo praticamente instantâneo.

O fluxo instantâneo de informações nos permite verificar distorções entre mercados diferentes; para compreendermos melhor este tópico, utilizemo-nos de um exemplo. Suponhamos que um astuto operador de mercado obtenha informações simultâneas nas bolsas de valores de São Paulo e do Rio de Janeiro. Assumamos que o preço do café na bolsa de São Paulo seja de \$15,00 a saca, ao passo que a mesma saca tenha \$13,00 como preço na bolsa do Rio. Neste caso, se o operador comprar café no Rio e vender em São Paulo, obterá um lucro certo de \$2,00 por saca. Ou seja: o operador obteve um lucro simplesmente por haver uma distorção de preços entre dois mercados diferentes. É este fenômeno que denominamos por *arbitragem*. Notemos que, caso a informação desta distorção de preços fosse conhecida pelo mercado, as operações de compra e de venda de café entre as bolsas de São Paulo e do Rio levariam o preço do café a um equilíbrio. Portanto, um sistema de informações eficiente leva à quase inexistência de oportunidades de arbitragem².

Um outro fator que reforça a necessidade de um sistema de informações eficiente se relaciona à constante alteração dos preços. Conforme vimos anteriormente, diversos fatores macroeconômicos levam à variações constantes em índices financeiros, como valor de ações, preços de mercadorias, taxas de câmbio, etc. Quanto maiores forem as informações sobre o histórico passado dos preços de um ativo, maior será nossa capacidade de podermos prever seu comportamento futuro. Podemos caracterizar então um mercado como sendo eficiente caso assumamos duas hipóteses básicas:

- ✘ *Os mercados respondem imediatamente a qualquer nova informação a respeito de um ativo;*
- ✘ *O histórico passado reflete-se inteiramente no preço presente de um ativo.*

As duas colocações acima consistem na *Hipótese do Mercado Eficiente*. A primeira reflete o fato, como vimos, da quase inexistência de oportunidades de arbitragem; a segunda reflete o fato de a dinâmica dos preços de um ativo ser *aleatória*.

Exploraremos a hipótese do mercado eficiente quando introduzirmos a Teoria de Black & Scholes.

² No caso discutido na seção anterior, o da caderneta de poupança, vimos um caso típico de arbitragem: a instituição financeira usa-se de seu alto volume de capital para arbitrar com as taxas de juros. Neste caso é possível a arbitragem porque altos volumes negociados em ações de riscos levam a possibilidades de retorno maiores que o que se pode obter com o investimentos de pequenos recursos, como é o caso de pequenos investidores.

VI.3 O Conceito de Derivativos

Em nossa introdução a esse capítulo mencionamos a ascensão do mercado de derivativos ([1], [17], [22]) como sendo uma opção aos investidores de modo a poderem operar em conjunturas incertas com uma maior margem de segurança. Mas, o que é um derivativo?

Para ilustrarmos melhor este conceito, consideremos um exemplo. Suponhamos que um industrial da área alimentícia esteja planejando suas vendas futuras. Seu ramo de atuação é o de produtos derivados de carne bovina. De modo a se precaver da instabilidade do preço da arroba do boi, o industrial decide fazer o seguinte contrato: comprar hoje a quantia necessária de bois para a fabricação dos insumos e recebê-los daqui a seis meses, e pagar na data de recebimento dos bois. O preço da arroba é estipulado hoje entre as partes; tal preço é fixo, mesmo que o preço da arroba do boi no mercado oscile. Este contrato estipulado é o que denominamos contrato de *derivativo* ([1], [17], [22]) : o valor do contrato *deriva* do valor da *commodity* (no caso, arrobas de boi) negociada. Com isso, os derivativos são instrumentos financeiros referenciados sobre outros ativos – sejam eles commodities, taxas, ou um outro índice financeiro.

Voltando ao exemplo mencionado, analisemos quais foram as vantagens para ambas as partes do contrato. Para o industrial foi interessante porque ele conseguiu fixar o preço futuro das arrobas de boi; com isso, caso no futuro seu preço aumente, ele não perderá em valor. No entanto, poderíamos perguntar: e se o preço da arroba caísse? O produtor perderia? A resposta é: não, pois em seu produto os preços da matéria-prima estão embutidos. Este industrial, ao fazer esta operação, é classificado no mercado como sendo um *hedger*, ou seja, um investidor cauteloso, que se utiliza do mercado para se proteger dos riscos inerentes às transações. E de fato, se protegeu, conforme vimos.

Quanto ao pecuarista que cria os bois, este também lucrou com o contrato, uma vez que, além de garantir sua venda, o fez com o preço já fixado. Assim, se o preço do mercado cair, ele já garantiu o seu preço de venda; no entanto, se o preço subir, o pecuarista não sairá no prejuízo, pois já embutiu seus custos e sua margem de lucro ao negociar o preço com o industrial. Por este contrato podemos também considerar o pecuarista como sendo um *hedger*.

Deste modo, o mercado de derivativos passou a ser uma excelente opção de investimento no mercado financeiro, pois permitiu aos investidores a opção do hedge, ou de usar o mercado para não incorrer riscos. No entanto, o mercado de derivativos também dá a opção àqueles que desejem obter ganhos altos; ou seja, permite o que é chamado no mercado financeiro de *especulação*: o *especulador* usa o mercado para obter lucros maiores. É claro que, nesta modalidade de investidor, os riscos incorridos são muito mais altos que para um *hedger* ou um outro investidor mais cauteloso.

Em termos de derivativos, os contratos mais comuns são os contratos a termo (como o dado no exemplo acima), os contratos futuros, os *swaps* e opções. É claro que cada um destes derivativos apresenta algumas variantes, mas a classificação é, em linhas gerais, esta apresentada aqui.

VI.4 Opções

VI.4.1 Definições

Depois de termos introduzido o conceito de derivativos, passemos à discussão de um dos principais tipos de derivativos: as opções ([1], [17], [22]). Novamente, a fim de ilustrarmos o conceito de opções, tomaremos como base um exemplo. Consideremos o caso de uma empresa aérea que deseja renovar sua frota de aeronaves. O comprador desta companhia então entra em contato com uma empresa fabricante de aeronaves, de modo a firmarem um contrato. Devido às oscilações do mercado de aviação civil, a companhia aérea adquirirá dez aeronaves dentro de um ano e meio por um valor de US\$ 200 milhões. No entanto, ao invés de haver uma obrigação formal de aquisição das aeronaves, a empresa aérea poderá *ou* efetuar esta aquisição *ou* simplesmente poderá desistir dela. A empresa aérea terá um ano para pronunciar-se acerca de sua decisão. Caso esta seja favorável, as aeronaves serão entregues seis meses após o pronunciamento da empresa aérea.

A empresa aérea paga, na data de hoje, US\$ 5 mi pelo direito de aquisição das aeronaves; o valor de US\$ 200 mi será pago caso a empresa exerça o contrato no final de um ano, ou seja, adquira efetivamente as aeronaves. E estes valores são fixados no ato da celebração do contrato.

Analisemos agora as vantagens deste contrato. Vamos trabalhar com dois cenários possíveis em um ano a partir da data acordada:

- 1 – O preço das aeronaves no mercado *subiu* para US\$ 210 mi;
- 2 – O preço das aeronaves no mercado *caiu* para US\$ 190 mi.

No primeiro cenário é mais vantajoso que a empresa aérea exerça o contrato e adquira as aeronaves por US\$ 200 mi; isso pois, somado a esse valor o montante já pago de US\$ 5 mi, o valor total pago pela empresa aérea será de US\$ 205 mi. Com isso, esta empresa terá realizado uma economia de US\$ 5 mi, pois terá adquirido as aeronaves por um valor inferior ao do mercado.

No segundo cenário é mais vantajoso que a empresa aérea não exerça o contrato, pois assim poderia adquirir as aeronaves do mercado por US\$ 190 mi e assim ter desembolsado US\$ 195 mi. Deste modo, esta empresa aérea também estaria economizando os mesmos US\$ 5 mi nesta outra transação.

O contrato mencionado acima constitui o que denominamos de *opção*. As opções são instrumentos financeiros muito atrativos, pois desobrigam o comprador a exercer a opção. Isso traz uma flexibilidade que os outros contratos de derivativos não apresentam. Em nosso exemplo, a companhia aérea, caso realizasse um contrato a termo, conforme vimos na seção anterior, seria obrigada a adquirir as aeronaves na data de vencimento do contrato; com isso, se o preço de mercado fosse maior que o preço estipulado, então o contrato seria lucrativo para a empresa aérea. No entanto, se o preço fosse menor, a empresa teria desvantagens com o contrato, pois estaria adquirindo aeronaves por um preço superior ao do mercado. Por isso vemos que as opções são autênticas soluções para transações em um cenário macroeconômico instável, pois são instrumentos financeiros que dispõem de uma flexibilidade maior que a de outros tipos de contratos.

Bem, e no caso do fabricante de aeronaves? Quais são suas vantagens ao firmar uma opção? A vantagem principal consiste na possibilidade de uma venda futura com preço

definido. No caso de a empresa aérea não exercer a opção, a fabricante de aeronaves não terá perdas, uma vez que esta já recebera um valor determinado pela venda desta opção.

Com todo o cenário discutido acima podemos conceituar as opções da seguinte maneira ([1], [17], [22]) :

VI. 2. Definição : Opções são instrumentos financeiros pelos quais o adquirente tem o direito, *mas não a obrigação*, de comprar ou vender um ativo real ou financeiro por um preço prefixado dentro de um prazo predeterminado. Existem dois tipos de opções:

- ✘ **Opção de Compra (Call Option)** : *O adquirente tem o direito (mas não a obrigação) de comprar;*
- ✘ **Opção de Venda (Put Option)** : *O adquirente tem o direito (mas não a obrigação) de vender.*

Os agentes participantes no mercado de opções são os seguintes:

- ✘ **Titular da Opção** : *É o agente que compra a opção. Este tem o direito (mas não a obrigação) de comprar a opção (no caso de uma call), ou de vender a opção (no caso de uma put).*
- ✘ **Lançador da Opção** : *É o agente que vende a opção. Este tem a obrigação de vender (no caso de uma call), ou de comprar (no caso de uma put), se o titular exercer a opção.*

O ativo ou instrumento financeiro que o lançador se obriga a vender (no caso de uma *call*), ou comprar (no caso de uma *put*) denomina-se **objeto**.

O preço que o titular paga ao lançador para ter o direito de comprar (no caso de uma *call*), ou o direito de vender (no caso de uma *put*) recebe o nome de **prêmio**. Este valor, pago no ato da aquisição da opção, não retorna ao titular caso este, em data futura, decida não exercer a opção.

A data de expiração da opção recebe o nome de **vencimento**. O preço a ser pago pelo titular caso exerça seu direito de compra (no caso de uma *call*), ou de venda (no caso de uma *put*) recebe o nome de **preço de exercício**. O preço de exercício é fixado no ato da aquisição da opção pelo titular.

Se o titular puder exercer a opção somente **no** vencimento, dizemos que esta é uma **opção européia**; no caso de o titular poder exercer a opção **até** o vencimento, dizemos que esta é uma **opção americana**.

No exemplo introduzido acima, a opção negociada é uma *call*, e a empresa aérea é a *titular da opção*; a fabricante é a *lançadora da opção*. O *objeto* da opção consiste nas aeronaves negociadas. O *prêmio* pago foi a quantia de US\$ 5 mi, o *vencimento* da opção é 1 ano a contar da data da celebração do contato, e o *preço de exercício* da opção é US\$ 200 mi. A opção considerada é do tipo *européu*.

Em nossa discussão ficaram surgiram dois valores sem maiores detalhamentos: o valor do preço de exercício e o valor do prêmio. Em sendo o preço de exercício o valor que as aeronaves assumirão no vencimento, como podemos *hoje* avaliar um *preço futuro*? E como se chegou a um prêmio de US\$ 5 mi? Como podemos calcular o valor a ser pago por uma opção? São as estas duas questões que abordaremos na próxima seção.

VI.4.2 Precificando Opções

Um tema que desafiou os financistas desde o início das operações no mercado de derivativos se refere exatamente a *precificação das opções* ([1], [6], [17], [22], [31], [35]) – ou seja, qual o preço *justo* a ser pago por uma opção? Estamos mencionando um preço a ser pago *hoje*, mas que deve ser referenciado em termos do preço futuro da opção, conforme vimos. A dificuldade reside, portanto, na obtenção de um valor futuro – o *preço futuro do ativo-objeto*. Inúmeros eventos macro e microeconômicos podem ocorrer no tempo transcorrido entre a data de hoje e a data de maturação da opção que afetarão decisivamente o valor do preço do ativo. Alguns destes eventos podem ser previsíveis; outros, entretanto, sequer podemos imaginar: atentados terroristas, crises políticas, intempéries da Natureza, e assim por diante. Estes eventos imprevisíveis indubitavelmente impactam os mercados globais como um todo, podendo até, em um caso extremo, levar empresas e bolsas de valores à bancarrota.

Diante destas dificuldades, como poderemos, então, estimar valores de preços futuros em cenários tão incertos? Primeiramente, em desconsiderando-se estes eventos imprevisíveis, podemos assumir que o comportamento futuro de uma dada variável espelha seu comportamento passado; em segundo lugar, em considerando-se as informações do mercado, podemos delinear um cenário macroeconômico no curto prazo de modo a verificarmos certas *tendências* no mercado. Estes dois fatores, em conjunto, nos permitem efetuar uma análise geral que podem levar à resultados muito satisfatórios. Tão melhor serão estes resultados quanto melhor for o modelo matemático usado na análise dos dados históricos e a avaliação dos especialistas financeiros dos cenários macroeconômicos em questão.

Deste modo, para podermos obter resultados mais acurados, é fundamental desenvolvermos um modelo matemático adequado para analisarmos as variações dos parâmetros financeiros ao longo do tempo. Daremos o nome genérico de *parâmetro financeiro* à todas as variáveis financeiras que tenham alterações em seus valores ao longo do tempo: são elas os preços das ações, a cotação do dólar, os valores de índices financeiros³ como IBOVESPA⁴ e Dow Jones, e assim por diante. Portanto, a análise de tais parâmetros se resume à análise da *evolução temporal* das variáveis que as representam.

A evolução temporal destes parâmetros está sujeita a inúmeros fatores conjunturais da economia que, apesar de *previsíveis*, não podemos adotar como *certos*. Por exemplo: em situações onde há a fuga de capitais do mercado, é *previsível* que o governo adote como medida corretiva o aumento das taxas de juros para a economia; entretanto, não podemos dizer *com certeza* que esta é a medida que será efetivamente adotada. Ademais, mesmo que o aumento da taxa de juros seja iminente, somente podemos *estimar* quanto será este aumento e quando este se dará. Este é apenas um exemplo dentre tantos fatores que podemos levantar como sendo responsáveis pela alta ou pela baixa dos parâmetros financeiros. Como tais fatores não podem ser descritos ao longo do tempo com exatidão, podemos classificá-los

³ Segundo Julio Brandt, repórter do InvestShop.com, (para maiores detalhes, consultar o site <http://www.bertolo.pro.br/MatFin/HTML/Ibovespa.htm>), “O índice de uma bolsa de valores serve para dar parâmetros de variação de valores ao mercado, ou seja, serve para que o investidor possa saber se, naquela bolsa, os papéis estão valorizando ou desvalorizando”.

⁴ Segundo o mesmo autor (consultar o site da nota anterior), o IBOVESPA é um índice composto pelas 55 ações mais negociadas na BOVESPA (Bolsa de Valores de São Paulo), e “(...) sua carteira é formada pelas ações que, em conjunto, representam 80% do volume negociado nos 12 meses anteriores à formação da carteira. Além disso, a ação precisa apresentar participação, em volume, de pelo menos 0,1% e tem que ter tido 80% de presença nos pregões do mesmo período”. Enio Fernando, da corretora carioca Futuro, diz que “O IBOVESPA serve para dar parâmetros ao mercado do como está sendo a negociação em um determinado dia. Por ser o principal índice da BOVESPA, é usado por todo o mercado como base para as tomadas de decisões”.

então como *agentes aleatórios* que interferem na evolução temporal dos parâmetros financeiros. Deste modo, concluímos que, diante do exposto, os parâmetros financeiros se comportam como *processos estocásticos*, pois sua evolução temporal está submetida à agentes aleatórios.

Portanto, podemos dizer que a eficiência do método matemático de análise dos dados históricos relativos aos parâmetros financeiros reside na escolha correta do processo estocástico para modelar seu comportamento futuro. Iniciaremos esta discussão na próxima subseção escolhendo o processo estocástico mais simples: o Passeio Aleatório. Antes, entretanto, é necessário observarmos mais alguns importantes pontos relativos à precificação de opções.

Verifiquemos de forma qualitativa como se perfaz a valorização de uma opção na data de vencimento ([6], [17], [22], [31], [35]). Consideremos uma *call* referenciada em um dado ativo cujo preço atual seja igual a K . Suponhamos que, na data de vencimento da opção, o valor do ativo seja igual a S . Caso $K > S$, a opção deixa de ser interessante, pois é mais vantajoso comprar o ativo diretamente do mercado, por um preço menor; portanto, a opção não teria valor algum, ou seja, seu valor é igual a zero. Se ocorrer $S > K$, o titular exercerá a opção, pois assim adquirirá o ativo por um preço inferior ao do mercado, e poderá negociá-lo a um preço superior. Portanto, neste caso, o lucro do titular desta *call* seria igual a $S - K$. Assim, a *call* vale, nesta data, $S - K$. A expressão matemática destes eventos assume a forma simples

$$C = \max\{0, S - K\}. \quad \text{VI.7}$$

Para o caso de uma *put*, a análise é reversa: no caso de $K > S$, o titular da *put* vende o ativo por um preço superior ao do mercado, percebendo então um lucro de $K - S$; portanto, este exerce a *put*, e seu valor será igual ao lucro obtido: $K - S$. Já no caso de $S > K$, o titular não exercerá a opção, pois preferirá vender o ativo no mercado por um preço maior. Portanto, o valor da *put* será igual a zero, neste caso. Com isso, obtemos uma equação análoga à (VI.7):

$$P = \max\{0, K - S\}. \quad \text{VI.8}$$

Pelo o discutido acima, podemos observar que, na data de vencimento, há três possibilidades de valor para a diferença entre o preço de exercício, S , e o preço atual K . Cada uma destas possibilidades recebe um termo de mercado:

Classificação	Call	Put
Dentro do Dinheiro <i>In the money</i>	$S > K$	$S < K$
No Dinheiro <i>At the money</i>	$S = K$	$S = K$
Fora do Dinheiro <i>Out of the money</i>	$S < K$	$S > K$

Tabela. VI.2 : Terminologia empregada no mercado para se descrever as relações entre os preços atual e de exercício.

Na terminologia de mercado usada no Brasil, dizemos que a opção ‘*virou pó*’ quando ela está fora do dinheiro. A região próxima à região ‘no dinheiro’ é denominada de ‘*linha d’água*’, pois esta é uma região de alto risco.

VI.4.3 O Modelo Binomial de Precificação de Opções

Uma vez discutidas algumas características gerais das opções, iniciaremos nesta subseção a análise de como podemos efetuar sua precificação. Discutimos na seção anterior que podemos estudar a evolução temporal de parâmetros financeiros através do uso de processos estocásticos. O processo mais simples que podemos adotar é o Passeio Aleatório, que é discreto no tempo e no seu espaço de probabilidades. Após adquirirmos um melhor insight sobre a precificação via este modelo discreto, poderemos passar aos casos mais complexos, utilizando-se de processos estocásticos contínuos.

O Modelo Binomial foi utilizado pela primeira vez por William Sharpe, da Universidade de Stanford, tendo depois sido desenvolvido por John C. Cox, do MIT, e Mark Rubinstein, da Universidade de Berkeley ([1], [6], [17], [22], [31], [35]). Neste modelo é assumido que a dinâmica dos preços de um certo ativo X é representada por um Passeio Aleatório. No momento presente, $t = 0$, o valor do preço do ativo é igual a S_0 ; no período seguinte, $t = 1$ este valor pode tanto valorizar a uS_0 , com probabilidade p , ou desvalorizar a dS_0 com probabilidade $1 - p$. Assumimos que os fatores u e d obedecem a relação $u > d > 0$. Em termos de uma árvore binomial de um passo apenas, o preço assumirá então os valores

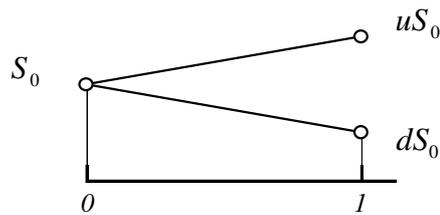


Fig. VI.1 : Dinâmica de preços do ativo X em um período.

Suponhamos agora que exista uma opção de compra sobre este ativo, cujo preço de exercício seja igual a K . Como a opção deriva do ativo X , sua evolução temporal também apresentará uma estrutura de Passeio Aleatório. Utilizando a equação (VI.7), a árvore binomial assume a seguinte forma para o preço da *call* :

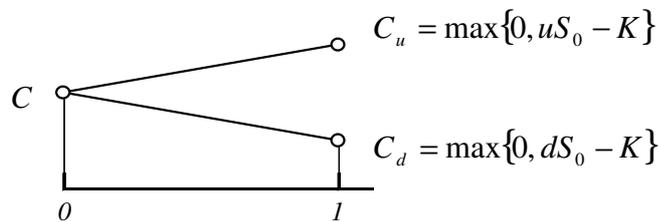


Fig. VI.2 : Dinâmica de preços da *call* em um período.

Para encontrarmos o valor da *call*, C , podemos utilizar uma técnica muito difundida em Física: a aproximação de funções. Esta técnica consiste na aproximação sucessiva de uma dada função desconhecida $f(t)$ em termos de combinações do tipo $\alpha_1 g_1(t) + \alpha_2 g_2(t) + \dots$ onde $g_1(t), g_2(t), \dots$ são funções conhecidas. Em nosso caso presente, $C(t)$ é a função desconhecida; cabe-nos encontrar as funções cujo comportamento se aproximem de $C(t)$. Adotemos a combinação mais simples, ou seja, da forma $\alpha g(t) + \beta$, onde β é uma constante; podemos tomar, sem perda de generalidade, $g(t)$ como sendo a função que representa o comportamento do próprio ativo-objeto da *call*, ou seja, X . Portanto, $C(t)$ e $\alpha g(t) + \beta$ apresentarão evoluções temporais semelhantes. Cabe-nos agora encontrar o valor de β . Em Finanças, há um tipo de ativo, denominado *ativo livre de risco*, que desempenhará o papel do parâmetro β . Basicamente, um ativo livre de risco é um instrumento financeiro que remunera o investimento realizado a uma taxa de juros r . Com isto, ao investir \$1 em ativos deste tipo, o investidor receberá após um período o valor $\$(1+r)$ (ver equação (VI.4)). Este ativo é livre de risco porque seu retorno é certo. Portanto a soma $\alpha g(t) + \beta$, onde β representa um ativo livre de risco, apresenta, evolução temporal idêntica à de $C(t)$.

Para comprovarmos a afirmação acima, elaboremos uma *carteira*⁵ de ativos de acordo com a seguinte estratégia:

⁵ Investir em uma *carteira* significa investir em um conjunto de ativos financeiros escolhidos de modo a fornecer a melhor relação retorno do investimento \times risco de perdas.

Evento em $t = 0$	Resultado em $t = 1$
I. Investir \$ α no ativo X ;	αu , no caso de uma alta; αd , no caso de uma baixa;
II. Investir \$ β em ativos livres de risco;	$(1+r)\beta$

A evolução temporal desta carteira é representável, conforme o esquema acima, via uma árvore binomial:

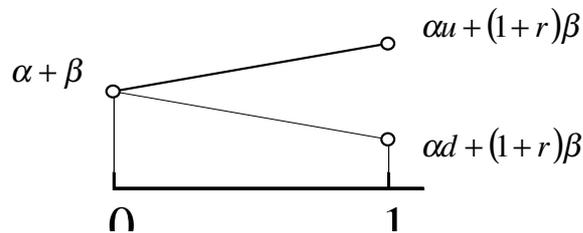


Fig. VI.3 : Dinâmica de preços da carteira replicante (ver texto).

Portanto, a evolução temporal da carteira é idêntica à da *call*, com a vantagem de todas as variáveis serem conhecidas. Notemos que, em ajustando-se os parâmetros α e β , podemos replicar qualquer outra carteira regida por um Passeio Aleatório; com isso, podemos então ajustá-los de modo que as seguintes expressões

$$\alpha u + (1+r)\beta = C_u, \text{ e}$$

$$\alpha d + (1+r)\beta = C_d, \tag{VI.9}$$

sejam verdadeiras. Resolvendo-se este sistema de equações, obtemos os seguintes valores para os montantes α e β :

$$\alpha = \frac{C_u - C_d}{u - d} \tag{VI.10}$$

$$\beta = \frac{uC_d - dC_u}{(1+r)(u-d)}. \tag{VI.11}$$

Com estes dados, a soma do montante investido hoje será igual a

$$\alpha + \beta = \frac{1}{1+r} \left\{ \frac{1+r-d}{u-d} C_u + \frac{u-1-r}{u-d} C_d \right\}. \tag{VI.12}$$

É fundamental que inexistas possibilidades de arbitragem no mercado em questão para efetuarmos as identificações (VI.9). Isso porque, caso haja arbitragem, sempre será possível comprar ativos onde estes estão em melhor preço para vendê-los em mercados onde o preço é maior. Com isso, as identificações feitas perdem o sentido. Portanto, em nossa discussão, consideraremos um mercado eficiente, sem oportunidades de arbitragem. Neste caso, podemos mostrar que $u > 1+r > d$.

Definamos agora a variável

$$q := \frac{1+r-d}{u-d}. \quad \text{VI.13}$$

Fazendo-se a identificação $C = \alpha + \beta$, segue-se de (VI.12) a seguinte expressão:

$$C = \frac{1}{1+r} \{qC_u + (1-q)C_d\}. \quad \text{VI.14}$$

A equação (VI.14) representa então o preço a ser pago pela *call* em um modelo binomial de um período. A generalização para mais de um período é imediata; em observando-se a árvore de dois períodos,

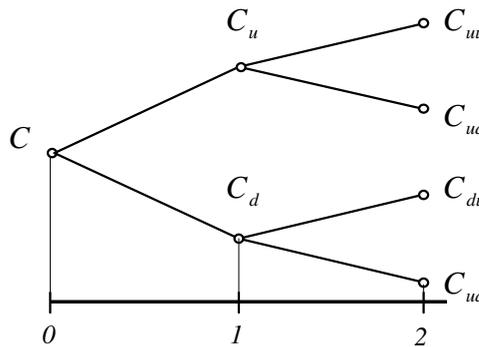


Fig. VI.4 : Dinâmica da *call* em dois períodos.

é fácil verificar que cada novo nó representa uma árvore de um período. Portanto, pela equação (VI.14), em $t=1$ temos os seguintes valores para a *call*:

$$C_u = \frac{1}{1+r} \{qC_{uu} + (1-q)C_{ud}\}$$

$$C_d = \frac{1}{1+r} \{qC_{du} + (1-q)C_{dd}\}. \quad \text{VI.15}$$

A generalização das equações acima é imediata para os períodos seguintes.

Façamos agora algumas observações acerca da variável q definida em (VI.13). Devido à relação $u > 1+r > d$, segue-se que $0 < q < 1$; portanto, podemos encarar q como sendo uma medida de probabilidade. Em termos desta medida de probabilidade, o lado esquerdo da equação (VI.14) representa, a menos de um fator de proporcionalidade, o valor esperado de C em relação à medida q . Como C é um processo estocástico, representemo-lo pela seguinte expressão:

$$C = C_t(\omega),$$

onde ω é um ponto amostral de um passeio aleatório. Efetuando-se as identificações

$$\begin{aligned} C_1(H) &= C_u, & C_1(T) &= C_d, \\ C_2(HH) &= C_{uu}, & C_2(HT) &= C_{ud}, & C_2(TH) &= C_{du} & \text{e} & C_2(TT) &= C_{dd}, \end{aligned}$$

VI.16

obtemos, via equações (VI. 14) e (VI.15),

$$C = \frac{1}{1+r} \{qC_1(H) + (1-q)C_1(T)\},$$

VI.17a

e

$$C_1(H) = \frac{1}{1+r} \{qC_2(HH) + (1-q)C_2(HT)\}$$

VI.17b

$$C_1(T) = \frac{1}{1+r} \{qC_2(TH) + (1-q)C_2(TT)\}.$$

VI.17c

Definindo-se o valor esperado em relação à medida q ,

$$E_q C_t := qC_{t+1}(\omega_1) + (1-q)C_{t+1}(\omega_2),$$

VI.18

segue-se que as equações (VI.17a – c) podem ser expressas pela fórmula geral

$$C_t = \frac{1}{1+r} E_q C_{t+1}.$$

VI.19

Denominamos a medida de probabilidade q por *medida livre de risco*, pois esta é independente das probabilidades de subida ou descida do valor do ativo.

Vamos ilustrar o presente modelo de Cox-Rubinstein através de um exemplo. Consideremos uma opção de compra de café, cujo vencimento seja em dois meses. Suponhamos que o valor de duzentos quilos de café seja igual a \$300, e seu preço de exercício seja \$290. A volatilidade é igual a 0,25, e a taxa de juros anual é igual a 10%. Escolhendo-se os parâmetros u e d por meio das equações (II.24), segue-se que

$$u = 1,0748,$$

$$d = 0,9304,$$

$$1+r = 1,0083,$$

onde a taxa de juros anual é convertida para uma taxa mensal em dividindo-a por 12. A probabilidade livre de risco é então igual a

$$q = 0,5396.$$

A árvore binomial do preço do ativo, e a do valor da *call*, são

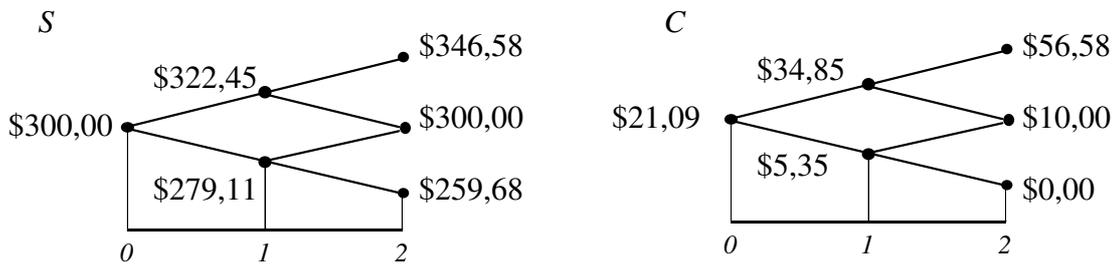


Fig. VI.5 : Árvores binomiais do preço do ativo e do valor da *call*.

Portanto, o valor da *call* é \$21,09.

O modelo de Cox-Rubinstein, conforme apresentado aqui, consiste na base sobre a qual edificaremos o mais importante modelo de precificação adotado hoje: o modelo de Black & Scholes.

VI.5 A Teoria de Black & Scholes

VI.5.1 A Equação de Black & Scholes

Conforme mencionamos na seção anterior, a eficiência do método matemático de análise dos dados históricos relativos aos parâmetros financeiros reside na escolha correta do processo estocástico para modelar seu comportamento futuro. Apresentamos na subseção anterior o modelo discreto de Cox-Rubinstein, baseado na escolha de um Passeio Aleatório como representante da dinâmica temporal dos preços. O próximo processo estocástico que escolheremos é o Movimento Browniano, que, conforme vimos, é o mais empregado na modelagem da evolução dinâmica de séries temporais. O modelo de Cox-Rubinstein, conforme discutido na seção anterior, nos servirá como uma preciosa fonte de insights para o desenvolvimento da teoria de Black & Scholes ([6], [17], [22], [31], [35]).

No capítulo anterior, discutimos as propriedades matemáticas gerais de um movimento Browniano. Estudamos o caso geral de um processo de Itô, representado pela equação (V.58),

$$dX_t = \xi(t, X_t)dt + \zeta(t, X_t)dW_t, \quad \text{V.58}$$

onde W_t é um Movimento Browniano e $\xi(t, X_t)$, $\zeta(t, X_t)$ são funções estocásticas. A equação acima apresenta um caso particular já conhecido em Finanças. Para obtê-lo, observemos que, caso $\zeta(t, X_t) = 0$, a equação (V.58) resulta em

$$dX_t = \xi(t, X_t)dt. \quad \text{VI.20}$$

Se $\xi(t, X_t) = \mu X_t$, a equação (VI.20) teria como solução

$$X_t = Pe^{\mu t}, \quad \text{VI.21}$$

onde $P = X_0$. Notemos que a solução (VI.21) é idêntica a (VI.6) no caso de μ ser uma taxa de juros; notemos também que, a cada \$1 investido à taxa μ , obteremos e^μ como retorno no tempo $t = 1$. Portanto, neste presente caso, X_t é um ativo livre de risco, conforme vimos anteriormente, pois seu retorno é garantido. Deste modo, vemos que um caso particular de um processo de Itô pode modelar um instrumento financeiro por nós já conhecido: o ativo livre de risco ou *bônus*. Com isso, para introduzirmos efeitos aleatórios na teoria financeira, é fundamental que o termo $\zeta(t, X_t)$ seja não-nulo na equação (V.58). No capítulo anterior discutimos um processo estocástico que apresentava $\xi(t, X_t) = \mu X_t$ e $\zeta(t, X_t) \neq 0$: o Movimento Browniano Geométrico (seção V.3.2). É este processo estocástico que tomaremos como base para modelar nossa dinâmica de preços do ativo-objeto X_t . Portanto, a equação obedecida pelo ativo-objeto é então

$$dX_t = \mu X_t dt + \sigma X_t dW_t. \quad \text{VI.22}$$

Uma vez estabelecida a equação que rege a dinâmica dos preços do ativo-objeto, passemos à análise de uma *call* referenciada sobre ele. Como a opção deriva seu valor do preço do ativo-objeto, devemos considerar que esta é uma função do processo estocástico X_t

$$C = C(X, t). \quad \text{VI.23}$$

A equação satisfeita por este processo estocástico é dada pelo lema de Itô (V.61):

$$dC[X_t] = \left\{ \frac{\partial C}{\partial t} + C'[X_t] \mu X_t + \frac{1}{2} C''[X_t] \sigma^2 X_t^2 \right\} dt + C'[X_t] \sigma X_t dW_t. \quad \text{VI.24}$$

Vamos utilizar neste ponto de nossa discussão algumas técnicas usadas para o desenvolvimento do modelo de Cox-Rubinstein. Vimos que o uso de uma carteira replicante foi fundamental para obtermos a equação (VI.14), e todas as outras subseqüentes. O cerne da estratégia da carteira replicante reside no fato de compormos uma carteira com um montante do ativo-objeto X_t e um montante de ativos livres de risco ou *bônus*, ajustados de forma a proporcionar valores iguais, a cada passo, entre o preço da *call* e o valor total da carteira. No caso contínuo, adotaremos a mesma estratégia; isto é, comporemos uma carteira Z_t de tal modo que

$$C_t = Z_t$$

a cada instante t . Adotando-se esta estratégia, a equação (VI.24) assumirá uma forma muito mais fácil de se trabalhar.

Assumamos então que a carteira Z_t seja composta por uma quantidade α do ativo X_t e por uma quantidade β de um *bônus* B_t que remunera a uma taxa r :

$$Z_t = \alpha X_t + \beta B_t. \quad \text{VI.25}$$

A equação diferencial satisfeita por Z_t é então igual a

$$dZ_t = \alpha dX_t + \beta dB_t. \quad \text{VI.26}$$

A equação diferencial satisfeita por um bônus é (VI.6), ao passo que a satisfeita pelo ativo X_t é (VI.22). Substituindo-se tais equações em (VI.26), obtemos então

$$\begin{aligned} dZ_t &= \alpha(\mu X_t + \sigma X_t dW_t) + r\beta B_t dt \\ &= (\alpha\mu X_t + r\beta B_t)dt + \alpha\sigma X_t dW_t. \end{aligned} \quad \text{VI.27}$$

Igualando-se as equações (VI.24) e (VI.27), obtemos

$$\left\{ \frac{\partial C}{\partial t} + C'[X_t] \mu X_t + \frac{1}{2} C''[X_t] \sigma^2 X_t^2 \right\} dt + C'[X_t] \sigma X_t dW_t = (\alpha\mu X_t + r\beta B_t) dt + \alpha\sigma X_t dW_t, \quad \text{VI.28}$$

Escolhendo-se

$$\alpha = C'[X_t], \quad \text{VI.29}$$

a equação (VI.25) nos fornece

$$\beta = \frac{1}{B_t} \left\{ C_t - X_t \frac{\partial C_t}{\partial X_t} \right\}. \quad \text{VI.30}$$

Substituindo-se as expressões (VI.29) e (VI.30) em (VI.28), segue-se que

$$\frac{\partial C_t}{\partial t} + \frac{1}{2} C_t'' \sigma^2 X_t^2 = rC_t - rX_t C_t',$$

ou

$$\frac{\partial C_t}{\partial t} + rX_t \frac{\partial C_t}{\partial X_t} + \frac{1}{2} \sigma^2 X_t^2 \frac{\partial^2 C_t}{\partial X_t^2} = rC_t. \quad \text{VI.31}$$

A expressão (VI.31) é a famosa equação de Black & Scholes, cujas implicações no escopo da teoria financeira são profundas. Esta equação estabelece que, em um mercado eficiente e perfeito, a função estocástica que representa o preço da opção é governada pela equação diferencial estocástica (VI.31). Se, de modo inverso, fixarmos C_t como a função que governa a dinâmica dos preços da opção, e escolhermo-la de modo que não satisfaça a equação de Black & Scholes, então haverá oportunidades de arbitragem no mercado para estes instrumentos financeiros. Isto porque o princípio fundamental no qual se baseia a equação de Black & Scholes reside na Hipótese do Mercado Eficiente.

Na dedução da equação de Black & Scholes, assumimos algumas condições de mercado para que esta representa de forma mais adequada os parâmetros financeiros; entretanto, há algumas outras condições que, embora importantes, foram omitidas. Entretanto, de forma geral, as condições necessárias para que as soluções da equação de Black & Scholes representem a dinâmica de preços de derivativos são dadas na seguinte

VI.3. Proposição ([6], [17], [22], [31], [35]) : Consideremos um mercado sem oportunidades de arbitragem, onde, a uma taxa livre de risco r , um dado ativo seja negociado cujo preço, a cada instante, seja representado por um Movimento Browniano Geométrico X_t . A volatilidade associada a este ativo é σ , e assumimos que não há custos de transação, e que o ativo não paga dividendos durante o período de validade da opção. Tais ativos podem ser negociados continuamente, em qualquer quantidade. Diante destas condições, o preço de uma opção sobre o ativo X_t é dado pela função estocástica $f(t, X_t)$ que satisfaça a **equação de Black & Scholes**:

$$\frac{\partial f}{\partial t} + rX_t \frac{\partial f}{\partial X_t} + \frac{1}{2} \sigma^2 X_t^2 \frac{\partial^2 f}{\partial X_t^2} = rf. \quad \text{VI.32}$$

É importante ressaltar um aspecto fundamental associado à equação de Black & Scholes: ela descreve a dinâmica de *qualquer* derivativo cujo preço dependa *somente* dos valores do ativo X_t e do tempo t . Para verificarmos este ponto, basta introduzirmos as condições de contorno que propiciarão as soluções da equação (VI.32). Na próxima subseção, verificaremos algumas soluções da equação de Black & Scholes, principalmente a solução desta para uma opção de compra européia.

VI.5.2 A Solução da Equação de Black & Scholes para uma Opção Européia

Sabemos da teoria das equações diferenciais que as condições iniciais e de contorno são os fatores que concedem às soluções das equações sua forma final e a tornam única. Portanto, ao estabelecermos estas condições, estaremos fixando o tipo de derivativo descrito pela equação de Black & Scholes. Como exemplo, suponhamos que $f(t, X_t)$ descreva um derivativo que satisfaça à equação de Black & Scholes sob as condições de contorno $f(T, X_T) = X_T$. A solução deste problema é dada pela função $f(t, X_t) = X_t$, o que demonstra o fato de a dinâmica do próprio ativo ser também governada pela equação de Black & Scholes.

Um outro exemplo que podemos dar se refere a um bônus: tomando-se a condição de contorno $f(T, X_T) = e^{rT}$, a solução da equação de Black & Scholes nos será dada pela função $f(t, X_t) = e^{rt}$, que representa um bônus. Portanto, vemos que as soluções da equação de Black & Scholes determinam a própria natureza financeira do ativo representado.

Consideremos agora a solução da equação de Black & Scholes para um tipo de opção extremamente importante: *opções européias* [35]. Suponhamos que $C_t[X_t]$ represente o preço de uma *call* européia com preço de exercício K e data de vencimento T . No vencimento, o valor da *call* será igual, conforme vimos na subseção VI.4.2, à equação (VI.7):

$$C_T[X_T] = \max\{0, X_T - K\}. \quad \text{VI.33}$$

No caso de $X = 0$, sabemos que a *call* não tem valor; portanto,

$$C_t[0] = 0. \quad \text{VI.34}$$

É fácil observar também que

$$C_t[X_t] \rightarrow \infty, \quad \text{se} \quad X_t \rightarrow \infty. \quad \text{VI.35}$$

Uma vez estabelecidas as condições de contorno, passaremos agora à análise da equação (VI.32). Conhecemos as soluções de uma equação de difusão, conforme desenvolvemos no Capítulo IV; entretanto, a equação de Black & Scholes não se apresenta sob a forma de uma equação de difusão. Entretanto, em escolhendo-se adequadamente os parâmetros que comparecem na equação (VI.32), podemos facilmente reduzi-la a uma equação de difusão. O primeiro passo consiste em se eliminar os termos proporcionais a X e a X^2 em (VI.32). Como os expoentes de X e de X^2 coincidem com a ordem das derivadas parciais às quais estão multiplicados, podemos representar X através do *ansatz*

$$X_t = \alpha e^\xi, \quad \text{VI.36}$$

onde α é uma constante de proporcionalidade. De fato, em substituindo-se (VI.34) nos termos $X \partial/\partial X$ e $X^2 \partial^2/\partial X^2$, segue-se que

$$X \frac{\partial}{\partial X} = \frac{\partial}{\partial \xi},$$

$$X^2 \frac{\partial^2}{\partial X^2} = \frac{\partial^2}{\partial \xi^2} - \frac{\partial}{\partial \xi}.$$

Com isso, a equação (VI.32) para a *call* C_t será igual a

$$\frac{\partial C_t}{\partial t} + \left(r - \frac{1}{2} \sigma^2 \right) \frac{\partial C_t}{\partial \xi} + \frac{1}{2} \sigma^2 \frac{\partial^2 C_t}{\partial \xi^2} = r C_t.$$

Dividindo-se ambos os membros da equação acima por $\sigma^2/2$, obtemos

$$-\frac{2}{\sigma^2} \frac{\partial C_t}{\partial t} = (\kappa - 1) \frac{\partial C_t}{\partial \xi} + \frac{\partial^2 C_t}{\partial \xi^2} - \kappa C_t, \quad \text{VI.37}$$

onde $\kappa = 2r/\sigma^2$. O fator multiplicativo na derivada temporal acima pode ser eliminado através de uma mudança de variáveis para o tempo. Seja τ esta nova variável, relacionada à t por meio da expressão $t = \beta\tau + \gamma$, sendo β e γ constantes. Portanto,

$$\frac{\partial}{\partial t} = \frac{1}{\beta} \frac{\partial}{\partial \tau},$$

o que nos sugere a identificação $\beta = -\sigma^2/2$. Como o parâmetro γ tem dimensão de tempo, escolhamos $\gamma = T$, para fins de simplificação. Deste modo, obtemos

$$t = -\frac{1}{2} \sigma^2 \tau + T. \quad \text{VI.38}$$

Portanto, por meio das relações (VI.36) e (VI.38), obtemos

$$-\infty < \xi < \infty, \quad \tau \geq 0. \quad \text{VI.39}$$

Analisemos agora como a *call* C_t se comporta em termos das novas variáveis (VI.36) e (VI.38). A condição de contorno (VI.32) torna-se

$$C_T [X_t] = \max\{0, \alpha e^{\xi(t)} - K\},$$

ou

$$C_T [X_t] = K \max\{0, e^{\xi(t)} - 1\},$$

em escolhendo-se $\alpha = K$. Como $t = T$ corresponde a $\tau = 0$, definamos

$$\phi(\xi, 0) = \max\{0, e^\xi - 1\};$$

com isso, podemos escrever a função estocástica C_t como

$$C_t = K\phi(\xi, \tau). \quad \text{VI.40}$$

A equação (VI.37), em termos destas novas variáveis, será então igual a

$$\frac{\partial \phi}{\partial \tau} = (\kappa - 1) \frac{\partial \phi}{\partial \xi} + \frac{\partial^2 \phi}{\partial \xi^2} - \kappa \phi. \quad \text{VI.41}$$

Esta expressão é semelhante a uma equação de difusão, excetuando-se os termos linear na função ϕ e sua primeira derivada. Podemos convertê-la em uma equação de difusão, por meio de uma mudança de variáveis; introduzamos, para tanto, a relação

$$\phi(\xi, \tau) = \psi(\xi, \tau)\chi(\xi, \tau); \quad \text{VI.42}$$

substituindo-se a relação (VI.42) em (VI.41), segue-se que

$$\left\{ \frac{\partial \chi}{\partial \tau} - \frac{\partial^2 \chi}{\partial \xi^2} \right\} \psi = -\chi \frac{\partial \psi}{\partial \tau} + 2 \frac{\partial \psi}{\partial \xi} \frac{\partial \chi}{\partial \xi} + \chi \frac{\partial^2 \psi}{\partial \xi^2} + (\kappa - 1) \frac{\partial \chi}{\partial \xi} \psi + (\kappa - 1) \frac{\partial \psi}{\partial \xi} \chi - \kappa \chi \psi. \quad \text{VI.43}$$

Supondo-se que a função $\chi(\xi, \tau)$ satisfaça uma equação de difusão

$$\frac{\partial}{\partial \tau} \chi(\xi, \tau) = \frac{\partial^2}{\partial \xi^2} \chi(\xi, \tau), \quad \text{VI.44}$$

encontramos então a forma da função $\psi(\xi, \tau)$ por meio da solução da equação

$$\left\{ -\frac{\partial \psi}{\partial \tau} + \frac{\partial^2 \psi}{\partial \xi^2} + (\kappa - 1) \frac{\partial \psi}{\partial \xi} - \kappa \psi \right\} \chi + \left\{ 2 \frac{\partial \psi}{\partial \xi} + (\kappa - 1) \psi \right\} \frac{\partial \chi}{\partial \xi} = 0. \quad \text{VI.45}$$

Para que a combinação linear acima seja linearmente independente, segue-se que os coeficientes acima devem se anular; por conseguinte,

$$2 \frac{\partial \psi}{\partial \xi} + (\kappa - 1) \psi = 0, \quad \text{VI.46}$$

e

$$-\frac{\partial \psi}{\partial \tau} + \frac{\partial^2 \psi}{\partial \xi^2} + (\kappa - 1) \frac{\partial \psi}{\partial \xi} - \kappa \psi = 0. \quad \text{VI.47}$$

A primeira equação pode ser solucionada por meio da separação de variáveis $\psi(\xi, \tau) = \theta(\xi)\vartheta(\tau)$, resultando na seguinte solução para $\theta(\xi)$:

$$\theta(\xi) = \lambda e^{-\frac{1}{2}(\kappa-1)\xi}, \quad \text{VI.48}$$

onde λ é uma constante. Substituindo-se $\psi(\xi, \tau) = \theta(\xi)\vartheta(\tau)$, com a expressão (VI.48), em (VI.47), obtemos, para $\vartheta(\tau)$,

$$\vartheta(\tau) = \nu e^{-\frac{1}{4}(\kappa+1)^2 \tau}, \quad \text{VI.49}$$

onde ν também é uma constante. Portanto, a função $\phi(\xi, \tau)$, solução da equação de Black & Scholes (VI.41), pode ser expressa em termos de uma solução de uma equação de difusão $\chi(\xi, \tau)$ por meio da relação

$$\phi(\xi, \tau) = e^{-\frac{1}{2}(\kappa-1)\xi - \frac{1}{4}(\kappa+1)^2 \tau} \chi(\xi, \tau). \quad \text{VI.50}$$

O produto $\lambda\nu$, oriundo da multiplicação dos termos (VI.48) e (VI.49), foi tomado como sendo igual a 1, sem perda de generalidade.

Retornemos agora à solução da equação de difusão (VI.44). A condição de contorno $\phi(\xi, 0) = \max\{0, e^\xi - 1\}$, juntamente com a expressão (VI.50), nos fornece

$$\chi(\xi, 0) = \max\left\{0, e^{\frac{1}{2}(\kappa+1)\xi} - e^{\frac{1}{2}(\kappa-1)\xi}\right\}; \quad \text{VI.51}$$

a condição (VI.34), por sua vez, se torna

$$\chi(0, \tau) = 0. \quad \text{VI.52}$$

Uma vez colocadas as condições iniciais, cabe-nos agora resolver a equação de difusão (VI.44), tendo como condição de contorno $\chi(\xi, 0) = \chi_0(\xi)$. No capítulo IV solucionamos a equação de difusão (IV.30) para um movimento Browniano utilizando como condição de contorno a expressão

$$\int_{-\infty}^{+\infty} f(x, t) dx = 1;$$

o que demonstra a relação entre a função densidade de probabilidade do movimento Browniano e uma função delta de Dirac, pois

$$\begin{aligned} \delta(x - x_0) &= \lim_{t \rightarrow 0} \frac{1}{\sqrt{4\pi t}} \exp\left\{-\frac{(x - x_0)^2}{4t}\right\} \\ &= \lim_{t \rightarrow 0} f(x - x_0, t). \end{aligned} \quad \text{VI.53}$$

Soluções do tipo

$$\chi_0(\zeta) f(\xi - \zeta, \tau),$$

com ζ fixo, satisfazem claramente a equação de difusão, assim como combinações lineares destes termos. Portanto, a função

$$\chi(\xi, \tau) = \int_{-\infty}^{+\infty} \chi_0(\zeta) f(\xi - \zeta, \tau) d\zeta \quad \text{VI.54}$$

satisfaz a equação de difusão (VI.44) sujeita à condição de contorno (VI.51), pois, em usando-se (VI.53),

$$\begin{aligned} \chi_0(\xi) &= \lim_{\tau \rightarrow 0} \chi(\xi, \tau) \\ &= \int_{-\infty}^{+\infty} \chi_0(\zeta) \delta(\xi - \zeta) d\zeta \\ &\equiv \chi_0(\xi). \end{aligned}$$

Com estas relações, podemos finalmente encontrar a solução para $\chi(\xi, \tau)$. Notemos primeiramente que

$$\begin{aligned} \chi_0(\zeta) &= \max\left\{0, e^{\frac{1}{2}(\kappa+1)\zeta} - e^{\frac{1}{2}(\kappa-1)\zeta}\right\} \\ &= e^{\frac{1}{2}\kappa\zeta} \max\left\{0, e^{\frac{1}{2}\zeta} - e^{-\frac{1}{2}\zeta}\right\} \\ &= e^{\frac{1}{2}(\kappa+1)\zeta} - e^{\frac{1}{2}(\kappa-1)\zeta}, \quad \text{para } \zeta > 0, \end{aligned} \quad \text{VI.55}$$

e sendo igual a zero, no caso $\zeta \leq 0$. Substituindo-se então (VI.55) e a expressão de $f(\xi - \zeta, \tau)$ ((IV.35), com $D = 1$ e $x = \xi - \zeta$) em (VI.54), obtemos

$$\begin{aligned} \chi(\xi, \tau) &= \frac{1}{\sqrt{4\pi\tau}} \int_0^{+\infty} \exp\left\{\frac{1}{2}(\kappa+1)\zeta\right\} \exp\left\{-\frac{(\xi-\zeta)^2}{4\tau}\right\} d\zeta \\ &\quad - \frac{1}{\sqrt{4\pi\tau}} \int_0^{+\infty} \exp\left\{\frac{1}{2}(\kappa-1)\zeta\right\} \exp\left\{-\frac{(\xi-\zeta)^2}{4\tau}\right\} d\zeta. \end{aligned} \quad \text{VI.56}$$

Para solucionarmos estas integrais, façamos a mudança de variáveis $\eta = (\zeta - \xi)/\sqrt{2\tau}$, o que fornece $-\xi/\sqrt{2\tau} \leq \eta < +\infty$; portanto, a primeira integral acima fornece

$$\begin{aligned} I_1 &= \frac{1}{\sqrt{2\pi}} \int_{-\xi/\sqrt{2\tau}}^{+\infty} \exp\left\{\frac{1}{2}(\kappa+1)(\sqrt{2\tau}\eta + \xi)\right\} e^{-\frac{1}{2}\eta^2} d\eta \\ &= \frac{e^{\frac{1}{2}(\kappa+1)\xi}}{\sqrt{2\pi}} \int_{-\xi/\sqrt{2\tau}}^{+\infty} \exp\left\{\frac{1}{2}[(\kappa+1)\sqrt{2\tau}\eta - \eta^2]\right\} d\eta \\ &= \frac{e^{\frac{1}{2}(\kappa+1)\xi + \frac{1}{4}(\kappa+1)^2\tau}}{\sqrt{2\pi}} \int_{-\xi/\sqrt{2\tau}}^{+\infty} \exp\left\{-\frac{1}{2}\left[\eta^2 - (\kappa+1)\sqrt{2\tau}\eta + \frac{1}{2}(\kappa+1)^2\tau\right]\right\} d\eta \\ &= \frac{e^{\frac{1}{2}(\kappa+1)\xi + \frac{1}{4}(\kappa+1)^2\tau}}{\sqrt{2\pi}} \int_{-\xi/\sqrt{2\tau}}^{+\infty} \exp\left\{-\frac{1}{2}\left[\eta - \sqrt{\frac{\tau}{2}}(\kappa+1)\right]^2\right\} d\eta \\ &= \frac{e^{\frac{1}{2}(\kappa+1)\xi + \frac{1}{4}(\kappa+1)^2\tau}}{\sqrt{2\pi}} \int_{-\frac{\xi}{\sqrt{2\tau}} - \sqrt{\frac{\tau}{2}}(\kappa+1)}^{+\infty} e^{-\frac{1}{2}\eta^2} d\eta. \end{aligned} \quad \text{VI.57}$$

Definindo-se

$$d_1 := \frac{\xi}{\sqrt{2\tau}} + \sqrt{\frac{\tau}{2}}(\kappa+1), \quad \text{VI.58}$$

segue-se que a equação (VI.57) será igual a

$$I_1 = \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{1}{2}(\kappa+1)\xi + \frac{1}{4}(\kappa+1)^2\tau\right\} N(d_1), \quad \text{VI.59}$$

onde $N(x)$ é a função distribuição gaussiana

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\eta^2} d\eta.$$

Notemos que a diferença entre as duas integrais do lado direito de (VI.56) se refere ao termo $(\kappa - 1)$; portanto, em substituindo-se $(\kappa + 1)$ por $(\kappa - 1)$ em (VI.59), obtemos o valor da segunda integral em (VI.56):

$$I_2 = \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{1}{2}(\kappa - 1)\xi + \frac{1}{4}(\kappa - 1)^2 \tau\right\} N(d_2), \quad \text{VI.60}$$

onde

$$d_2 := \frac{\xi}{\sqrt{2\tau}} + \sqrt{\frac{\tau}{2}}(\kappa - 1). \quad \text{VI.61}$$

Uma vez descoberta a forma final de $\chi(\xi, \tau)$, calculamos facilmente a forma final da função estocástica $C_t[X_t]$; para tanto, substituindo-se (VI.59) e (VI.60) em (VI.56), por sua vez (VI.56) em (VI.50), (VI.50) em (VI.40), e fazendo-se as identificações

$$\xi = \ln\left(\frac{X}{K}\right), \quad \tau = \frac{1}{2}\sigma^2(T - t) \quad \text{e} \quad \kappa = \frac{2r}{\sigma^2},$$

chega-se à solução final

$$C_t[X_t] = X_t N(d_1) - Ke^{-r(T-t)} N(d_2), \quad \text{VI.62}$$

onde

$$d_1 = \frac{\ln\left(\frac{X_t}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T - t)}{\sigma\sqrt{T - t}}$$

$$d_2 = \frac{\ln\left(\frac{X_t}{K}\right) + \left(r - \frac{1}{2}\sigma^2\right)(T - t)}{\sigma\sqrt{T - t}}. \quad \text{VI.63}$$

Esta é a solução final da equação de Black & Scholes para uma opção européia. Vamos retomar nosso exemplo mencionado na subseção onde discutimos o modelo de Cox-Rubinstein. Neste, um contrato de duzentos quilos de café seria negociado, com vencimento em dois meses, com preço igual a \$300, preço de exercício igual a \$290, a uma taxa de juros de 10% ao ano, e volatilidade de 0,25. Obtemos os seguintes parâmetros em utilizando-se as equações (VI.63):

$$d_1 = 0,546496,$$

$$d_2 = 0,444434,$$

o que fornece

$$N(d_1) = 0,70763,$$

$$N(d_2) = 0,67163.$$

Portanto, o valor da call será igual a

$$C = \$20,74,$$

ou seja, um valor próximo ao obtido pelo modelo de Cox-Rubinstein, onde $C = \$21,09$. Comparações mais detalhadas dos modelos de Cox-Rubinstein e de Black & Scholes podem ser encontradas na bibliografia especializada ([6], [17], [22], [31], [35]).

CAPÍTULO VII

GENERALIZAÇÕES DO MOVIMENTO BROWNIANO E SUAS APLICAÇÕES ÀS SÉRIES FINANCEIRAS

VII.1 Processos de Lévy

No capítulo anterior discutimos como podemos aplicar a Teoria dos Processos Estocásticos na precificação de opções. Observamos a aplicação de um passeio aleatório como modelo base para a evolução temporal dos preços do ativo-objeto, culminando no modelo de Cox-Rubinstein; posteriormente, analisamos como modelo base para a dinâmica de preços um Movimento Browniano Geométrico, o que resultou no modelo de Black & Scholes. A escolha do processo estocástico determina, assim, todo o processo posterior de análise. Entretanto, poderíamos nos questionar: a dinâmica de preços dos ativos se comportam realmente como um Passeio Aleatório, ou como um Movimento Browniano Geométrico? Podemos responder esta pergunta somente de forma empírica; portanto, é necessário desenvolvermos métodos de análise que nos permitam avaliar quão bom é um processo estocástico para se modelar um sistema dinâmico [24].

No capítulo III discutimos a Teoria de Lévy, que consiste em uma classe de processos estocásticos mais abrangentes, contendo, inclusive, o movimento Browniano como um caso particular. Dentre estes processos estocásticos analisamos os processos estáveis, caracterizados por um expoente de estabilidade, α , que englobam tanto os processos gaussianos ($\alpha = 2$) quanto processos de Cauchy ($\alpha = 1$). Por sua generalidade, podemos supor que, caso a dinâmica de preços não seja um Movimento Browniano Geométrico, seja um processo estável com $\alpha \neq 2$. Portanto, tomaremos como base em nossa análise da dinâmica de preços processos estáveis.

Retornemos então ao capítulo III. A função característica mais geral para um processo estável é dada pelo Teorema III.15,

$$\ln \phi(\xi) = i\gamma\xi - \beta|\xi|^\alpha \left\{ 1 + i\delta \frac{\xi}{|\xi|} \Phi(\xi, \alpha) \right\}, \quad \text{VII.1}$$

onde a função $\Phi(\xi, \alpha)$ é definida por

$$\Phi(\xi, \alpha) = \begin{cases} \tan \frac{\pi\alpha}{2}, & \text{se } \alpha \neq 1 \\ \frac{2}{\pi} \ln|\xi|, & \text{se } \alpha = 1 \end{cases} \quad \text{VII.2}$$

Consideremos, para efeitos de análise, que os processos estocásticos em consideração apresentem distribuição de probabilidades simétrica, o que nos traria $\delta = 0$; também consideraremos que o valor esperado de tais processos sejam iguais a zero, o que nos traria,

por sua vez, $\gamma = 0$ na equação (VII.1). Estas suposições não restringem a aplicabilidade dos métodos a serem desenvolvidos, pois, conforme já vimos, é sempre possível escolhermos, a partir de um processo estocástico com $\delta \neq 0$ e $\gamma \neq 0$, um processo estocástico equivalente que apresente $\delta = 0$ e $\gamma = 0$, sem perda de generalidade. Vimos, no capítulo I, que uma função característica $\phi(\xi)$ se relaciona à função de distribuição de uma variável aleatória através da equação

$$F(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(\xi) e^{-i\xi x} d\xi; \quad \text{VII.3}$$

portanto, em substituindo-se (VII.1) em (VII.3), com $\delta = 0$ e $\gamma = 0$, segue-se que

$$L_{\alpha,\beta}(x) = \frac{1}{\pi} \int_0^{\infty} \exp\{-\beta\xi^\alpha\} \cos(\xi x) d\xi \quad \text{VII.4}$$

que consiste em uma *distribuição de Lévy*. Portanto, as distribuições de probabilidade reais dos preços de um dado ativo devem assumir a forma acima, cada uma caracterizada por seu expoente de estabilidade e fator de escala próprios. Serão, portanto, *processos de Lévy*.

A equação que define uma distribuição de Lévy apresenta uma simetria muito importante para nossos objetivos: a *simetria de escala* [24]. Consideremos que haja uma variável \tilde{X} que se relacione com X através da relação $\tilde{X} = \varepsilon X$, onde ε é um dado parâmetro. A função de distribuição para esta variável aleatória será dada por

$$\begin{aligned} \tilde{L}_{\alpha,\beta}(x) &= L_{\alpha,\beta}(x) \\ &= \frac{1}{\pi} \int_0^{\infty} \exp\{-\beta\xi^\alpha\} \cos(\varepsilon\xi x) d\xi. \end{aligned}$$

Fazendo-se a mudança de variáveis $\zeta = \varepsilon\xi$, e perfazendo-se os cálculos triviais, segue-se que

$$\tilde{L}_{\alpha,\beta}(x) = \frac{1}{\pi\varepsilon} \int_0^{\infty} \exp\{-\eta\zeta^\alpha\} \cos(\zeta x) d\zeta, \quad \text{VII.5}$$

onde $\eta := \beta\varepsilon^{-\alpha}$. Observamos, então, que, a menos do fator ε^{-1} , a função de distribuição (VII.5) para uma variável εX é igual à de X , tendo sido alterado apenas o *fator de escala*. Portanto, em dilatando-se os processos estocásticos X obtemos processos estocásticos de *mesma distribuição*, porém com *fatores de escala diferentes*. Esta propriedade das distribuições de Lévy recebe o nome de *auto-similaridade*. É através da propriedade de auto-similaridade dos processos de Lévy que desenvolveremos o método para verificarmos se uma dada distribuição de probabilidades é normal ou não.

Podemos obter infinitos processos através da dilatação de um processo estocástico. Para nossos propósitos, introduzamos o parâmetro de dilatação ε como sendo $\varepsilon = \tau^{-1/\alpha}$; com isso, a função de distribuição (VII.5) assume a forma

$$\tilde{L}_{\alpha,\beta}(x) = \frac{\tau^{1/\alpha}}{\pi} \int_0^{\infty} \exp\{-\beta\tau\xi^\alpha\} \cos(\xi x) d\xi. \quad \text{VII.6}$$

Definindo-se

$$L_{\alpha,\beta}(x, \tau) = \frac{1}{\pi} \int_0^{\infty} \exp\{-\beta\tau\xi^\alpha\} \cos(\xi x) d\xi, \quad \text{VII.7}$$

segue-se a relação

$$L_{\alpha,\beta}(x, \tau) = \frac{L_{\alpha,\beta}(\tilde{x}, 1)}{\tau^{1/\alpha}}, \quad \tilde{x} = \tau^{-1/\alpha} x. \quad \text{VII.8}$$

VII.2 Cálculo do Expoente de Estabilidade para Processos Estocásticos Simulados

Uma vez introduzidos os processos estocásticos de Lévy, investiguemos como podemos obter empiricamente o expoente de estabilidade para um dado processo estocástico. Calculemos, inicialmente, a probabilidade de ocorrência do retorno do processo à origem, $P[X = 0]$ [24]. Sabemos, de nossas discussões anteriores, que a medida de probabilidade e sua função de distribuição se relacionam por meio da equação

$$P[X \leq x] = F(x).$$

No caso de variáveis aleatórias cuja distribuição de probabilidades é de Lévy, segue-se que

$$P[X \leq x] = L_{\alpha,\beta}(x, \tau). \quad \text{VII.9}$$

Portanto, para variáveis de Lévy, segue-se que $P[X = 0] = L_{\alpha,\beta}(0, \tau)$, o que resulta na expressão

$$P[X = 0] = \frac{\Gamma(1/\alpha)}{\pi\alpha(\beta\tau)^{1/\alpha}}, \quad \text{VII.10}$$

onde usamos a equação (VII.7). Deste modo, a probabilidade que, em um dado instante, o valor do processo estocástico seja igual a zero é função do parâmetro de dilatação τ ; reescrevendo-se $P[X = 0]$ como $p(\tau) = P[X = 0]$, obtemos uma função linear em tomando-se os logaritmos de ambos os lados da equação (VII.10):

$$\log p(\tau) = -\frac{1}{\alpha} \log \tau + \log \left[\frac{\Gamma(1/\alpha)}{\pi\alpha\beta^{1/\alpha}} \right]. \quad \text{VII.11}$$

Esta função linear nos permite de forma imediata a obtenção do expoente de estabilidade e do fator de escala para uma dada série temporal. Para tanto, basta seguirmos os seguintes passos [24]:

- I. Obter processos estocásticos equivalentes ao original X_t por meio da dilatação $\tilde{X}_t = \tau^{-1/\alpha} X_t$;
- II. Calcular a distribuição de probabilidades empíricas para cada um destes, obtendo-se posteriormente os valores de $p(\tau)$;
- III. Traçar o gráfico $\log p \times \log \tau$;
- IV. Ajustar uma reta aos pontos deste gráfico, calculando-se os coeficientes angular e linear desta reta;

Sendo $q(x) = a_1 x + a_0$ a equação da reta ajustada aos pontos empíricos, segue-se, a partir de (VII.11), as seguintes expressões:

$$\alpha = -\frac{1}{a_1};$$

$$\log \beta = \alpha \left\{ \log \left[\frac{\Gamma(1/\alpha)}{\pi \alpha} \right] - a_0 \right\}; \quad \text{VII.12}$$

Verificaremos o emprego deste método através da análise de dois processos estocásticos com expoente de estabilidade conhecidos: o Movimento Browniano e um Processo de Cauchy.

VII.2.1 Cálculo do Expoente de Estabilidade para um Movimento Browniano Simulado

Como primeira aplicação do método consideremos um movimento Browniano. Simulamos em MATLAB um movimento Browniano com valor esperado zero e variância igual a 1; empregamos, para tanto, o processo de Wiener

$$X_t = X_{t-\tau} + \xi \sqrt{\tau},$$

onde ξ é um número aleatório sacado de uma distribuição normal de média zero e variância um. Usamos o gerador de números aleatórios com distribuição normal do MATLAB para obter a série de números $\{\xi\}$. Neste processo, escolhemos $X_0 = 0$, $\tau = 1$ e geramos uma série de 25.000 pontos, onde então obtivemos o seguinte gráfico:

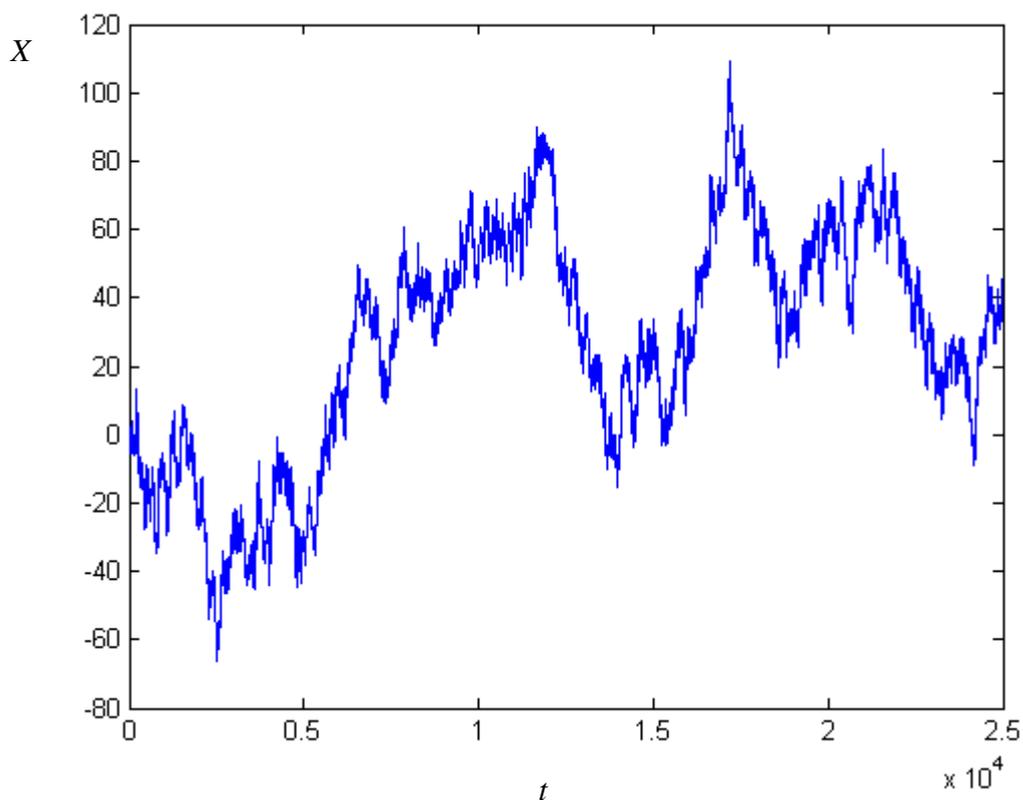


Fig. VII.1 – Movimento Browniano simulado com $\tau = 1$

Tomando-se os incrementos $W_t = X_{t+1} - X_t$, obtemos os valores empíricos

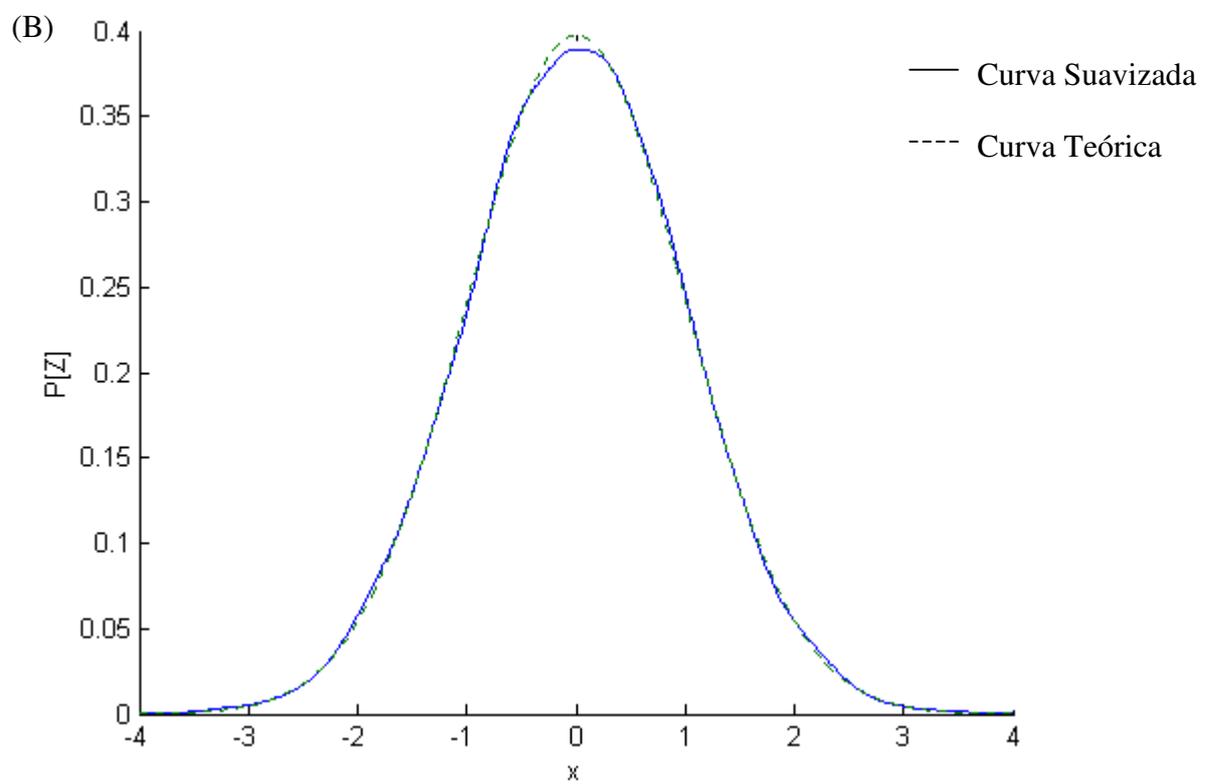
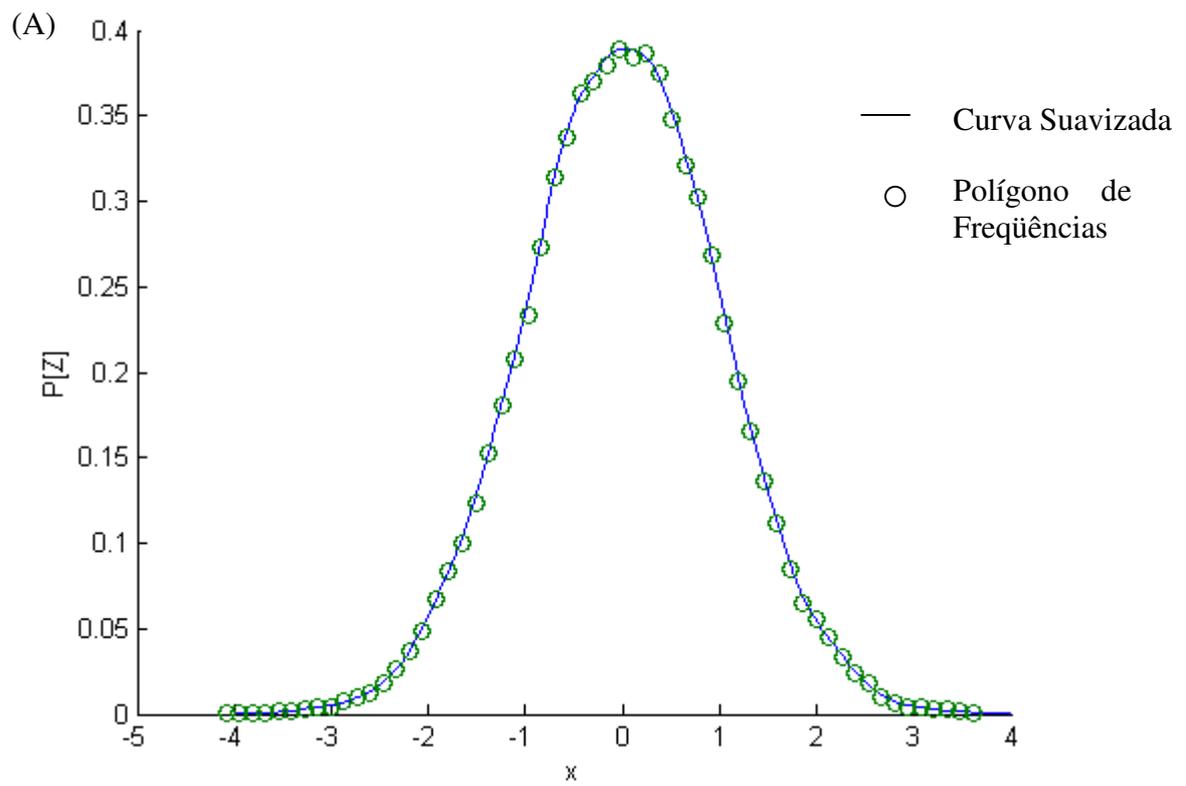
$$\mu = 0.002$$

$$\sigma = 1,003 .$$

VII.13

A distribuição de probabilidades do processo estocástico W_t pode ser obtida por meio do polígono de freqüência dos dados empíricos¹; entretanto, esta distribuição é, de certo modo, grosseira, não nos permitindo avaliar a probabilidade de retorno do processo a zero. Deste modo, precisamos obter uma curva lisa, baseada em dados empíricos, que melhor represente a distribuição de probabilidades deste Movimento Browniano; empregamos, para tanto, um método estatístico não-paramétrico [14], para o qual desenvolvemos um programa em MATLAB. Na figura VII.2 (A) comparamos o polígono de freqüências dos dados empíricos com sua distribuição suavizada. Comparamos também, na figura seguinte, VII.2 (B), a distribuição teórica, dada pela equação que define uma distribuição gaussiana (com os parâmetros (VII.13)), com a distribuição suavizada pelo método não-paramétrico. Por fim, na figura VII.2 (C), comparamos as três curvas.

¹ O polígono de freqüências corresponde à curva da distribuição de probabilidades de um dado experimento. Como tal, será uma curva grosseira e, conseqüentemente, não-diferenciável em diversos pontos. Para maiores detalhes, ver as referências [14], [26] e [32] da bibliografia.



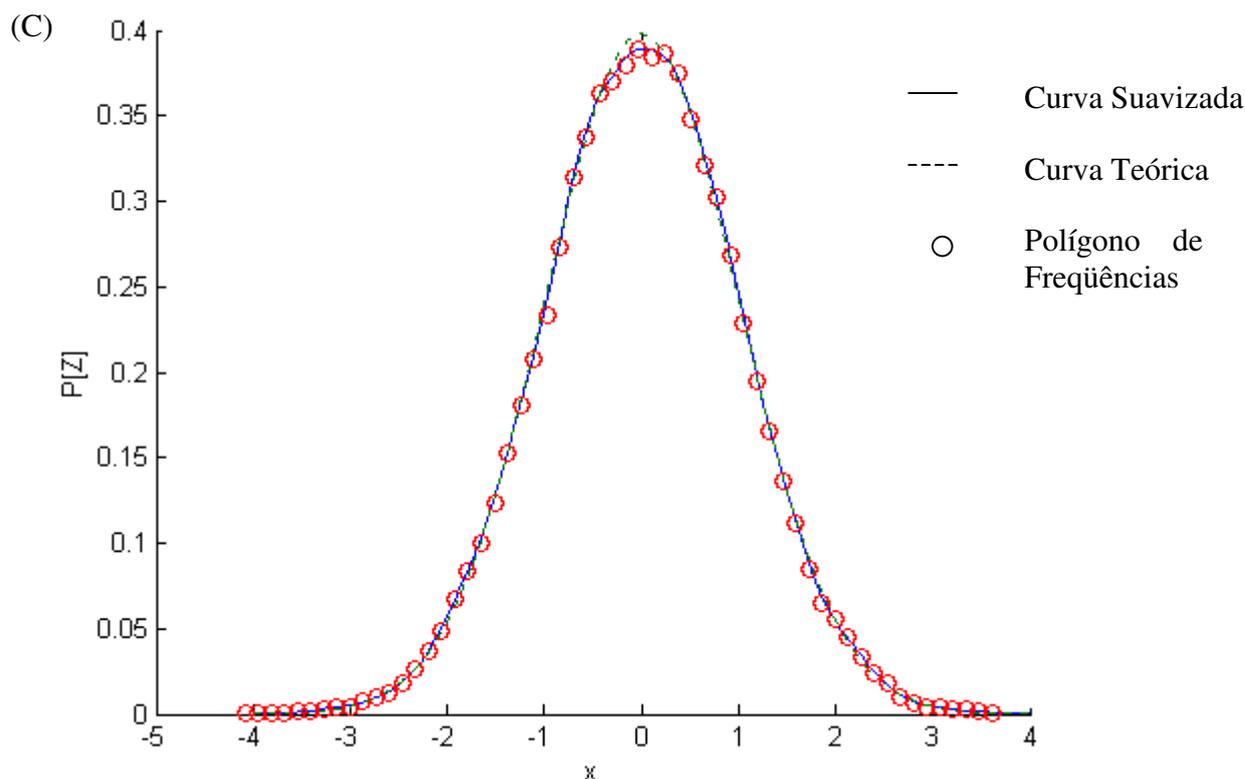


Fig VII.2 – Comparações entre (A) a distribuição suavizada com o polígono de freqüências; (B) a distribuição suavizada e a distribuição teórica e (C), as três curvas.

Verificamos em (A) que há uma boa concordância entre os pontos do polígono de freqüências e a distribuição suavizada. Em (B), vemos que há uma leve discrepância entre as probabilidades próximas aos pontos zero nas distribuições teórica e suavizada; entretanto, conforme verificaremos, esta diferença é pequena, não invalidando portanto nosso modelo não-paramétrico. Em (C) comparamos as três curvas, de modo a observar as diferenças entre os dois modelos empíricos e o modelo teórico.

Uma vez introduzida a distribuição de probabilidades empírica, nosso próximo passo consiste em encontrar os processos equivalentes a nosso movimento Browniano simulado. A estratégia básica consiste no seguinte: tomamos, dentro da série original X_t , os pontos espaçados pelo intervalo τ , e tomamos seus incrementos; o movimento Browniano correspondente terá variância igual a τ . Na figura VII.2, mostramos o gráfico correspondente ao movimento Browniano gerado por X_t ao escolher $\tau = 2$, e verificamos que este oscila muito mais do que o movimento Browniano original. O desvio padrão calculado para este processo é igual a

$$\sigma = 1,408,$$

que é um valor próximo a $\sqrt{2} = 1,414$, o desvio original. Na figura VII.4 podemos observar os gráficos da distribuição teórica, do polígono de freqüências, e da distribuição suavizada para este processo:

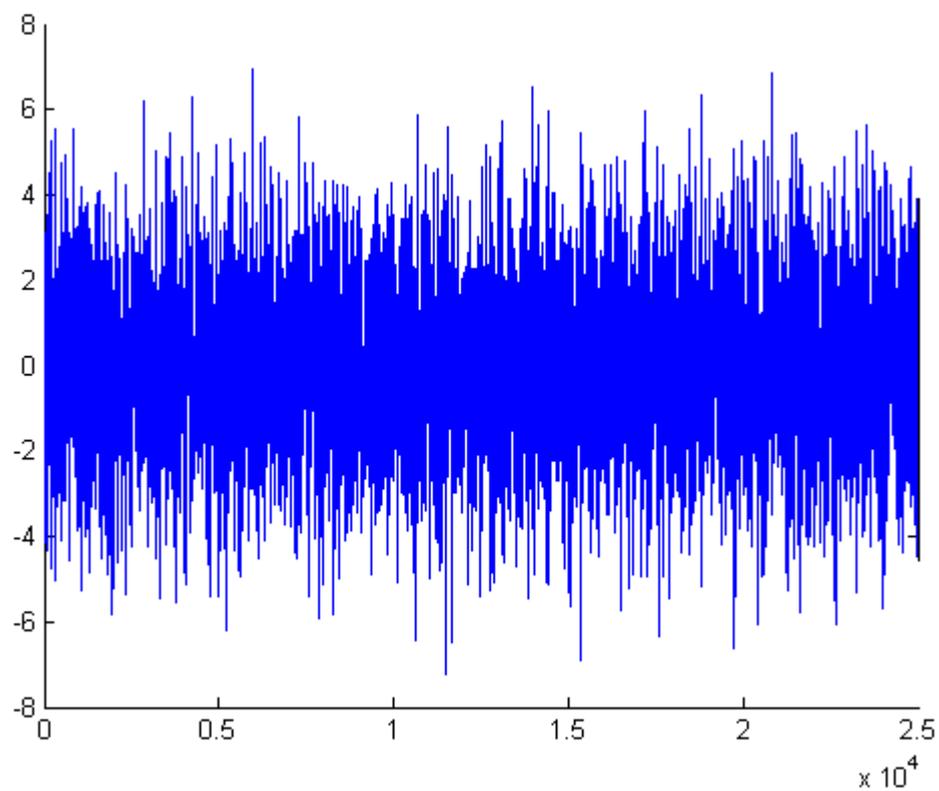


Fig. VII.3 – Movimento Browniano simulado com $\tau = 2$, obtido a partir da série original.

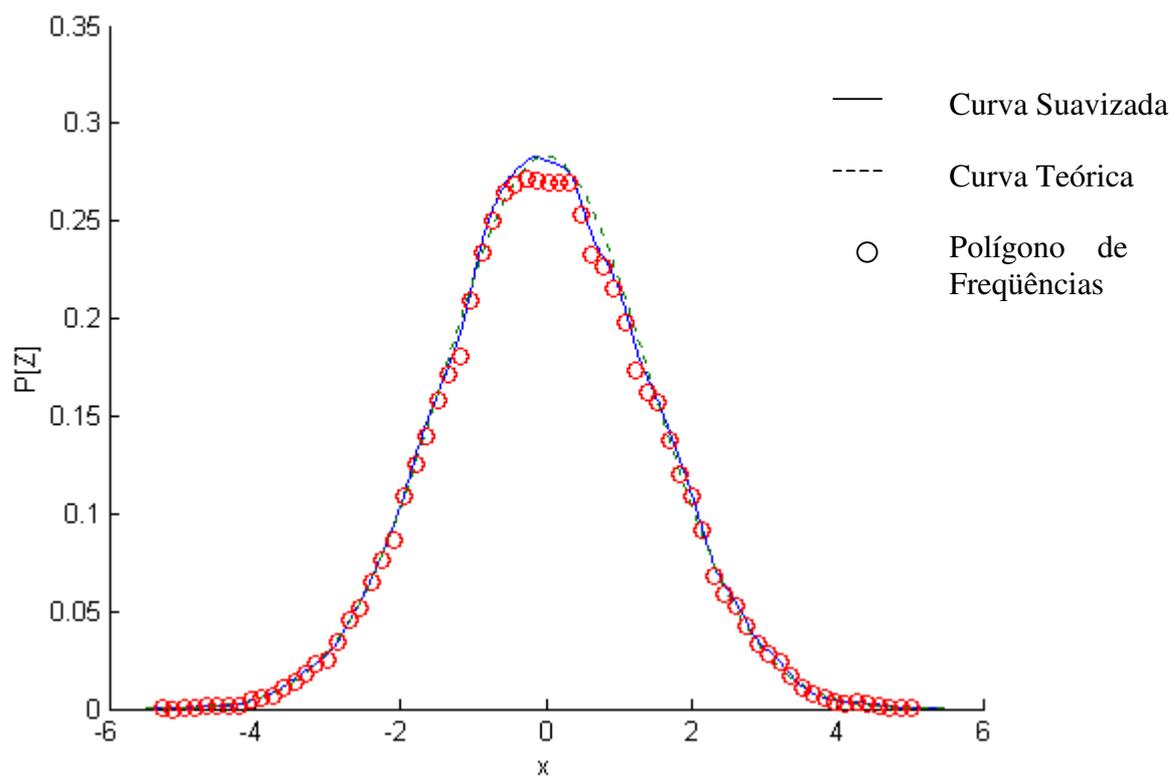


Fig. VII.4 – Comparação entre as distribuições teórica, suavizada e polígono de frequências

Procedemos como acima para obtermos outros processos estocásticos equivalentes. Uma vez obtidos os processos, calculamos os valores de $p(\tau)$, e traçamos assim o gráfico $\log p \times \log \tau$ com a reta ajustada:

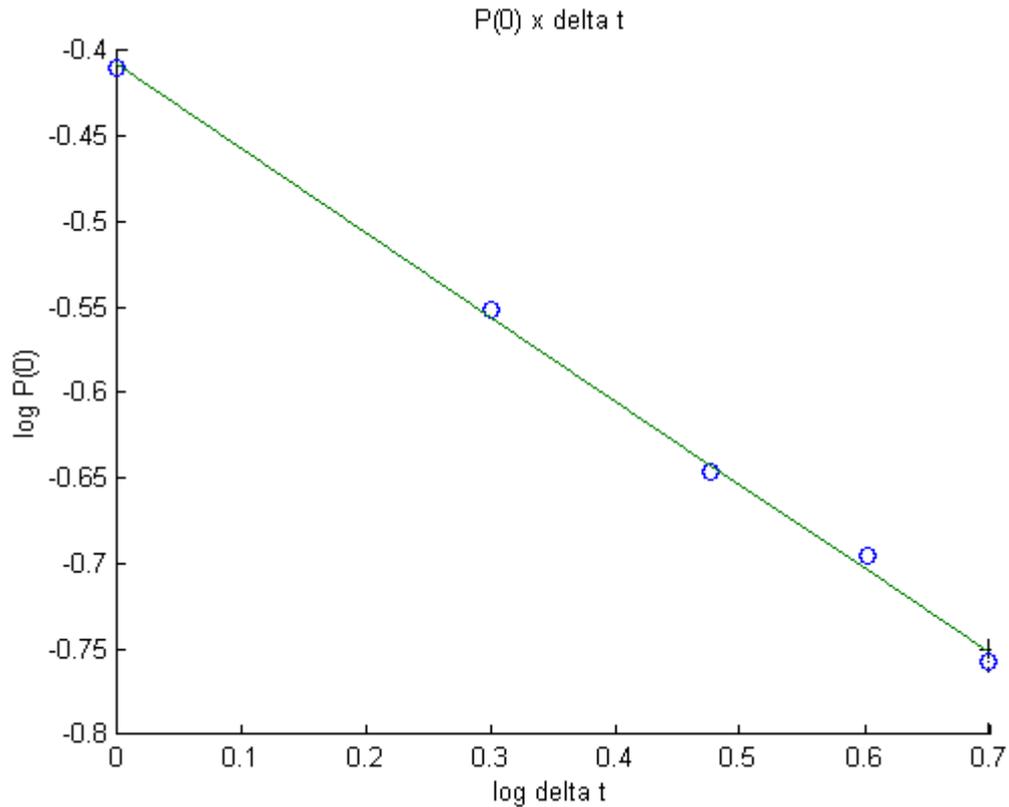


Fig. VII.5 – Gráfico $\log p \times \log \tau$ para o cálculo dos coeficientes angular e linear.

Após o ajuste, calculamos os coeficientes angular e linear da reta ajustada; em empregando-se a equação (VII.12), obtemos finalmente o expoente de estabilidade α e o fator de escala, β :

$$\alpha = 2.003 \pm 0.002 ,$$

$$\beta = 0.514 \pm 0.005 . \quad \text{VII.14}$$

O valor do expoente de estabilidade está de pleno acordo com o previsto pela teoria, $\alpha = 2$; verificaremos, na próxima subseção, se o modelo proposto também fornece valores compatíveis com a teoria para um processo simulado de Cauchy.

Para finalizarmos nossa análise do movimento Browniano simulado, desenhemos o gráfico da distribuição de Lévy (VII.4) associada com os parâmetros (VII.14), e o comparemos ao polígono de frequências da distribuição e sua distribuição teórica, gaussiana:

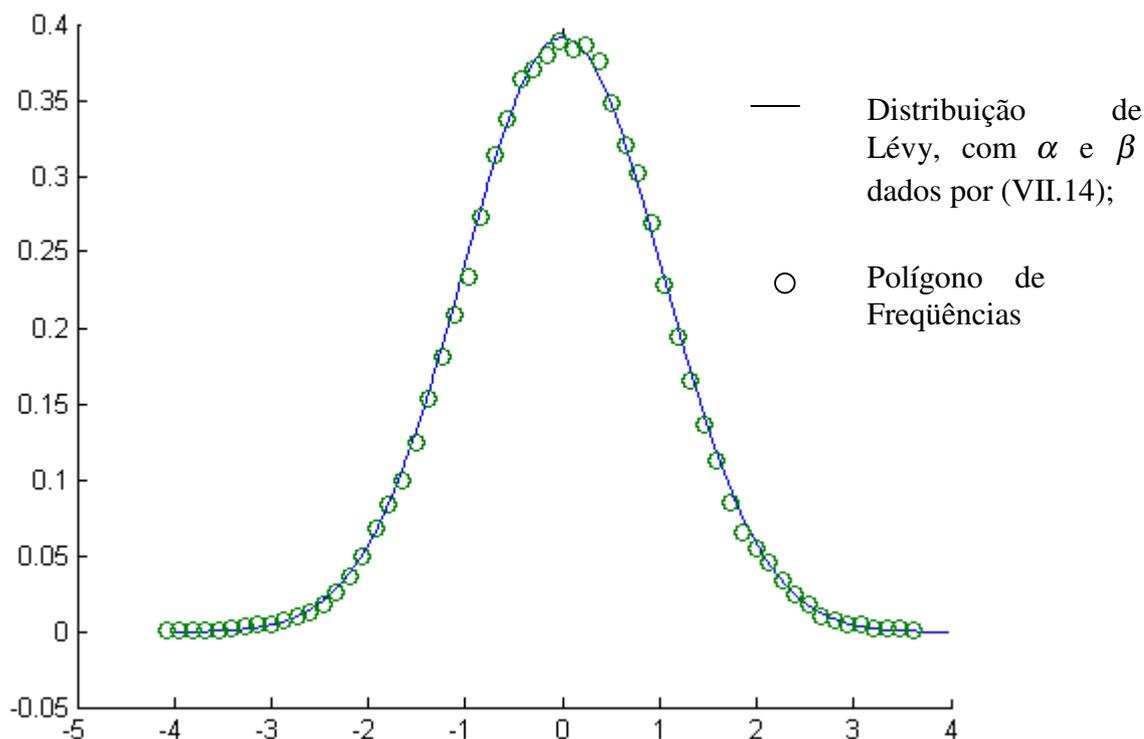
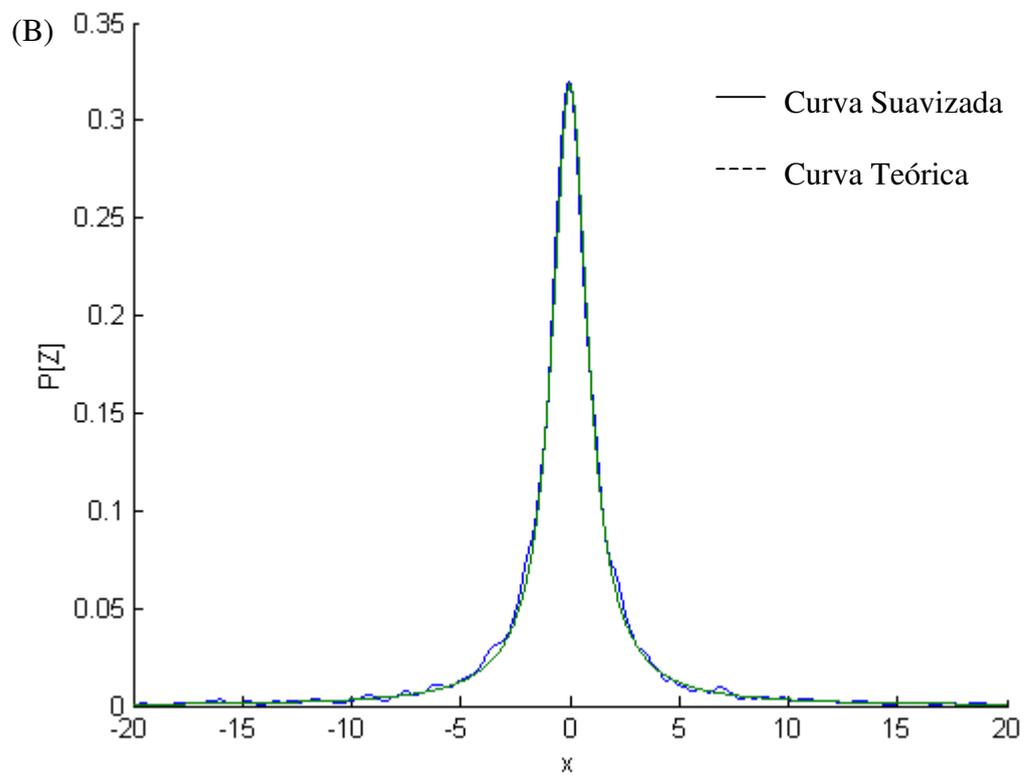
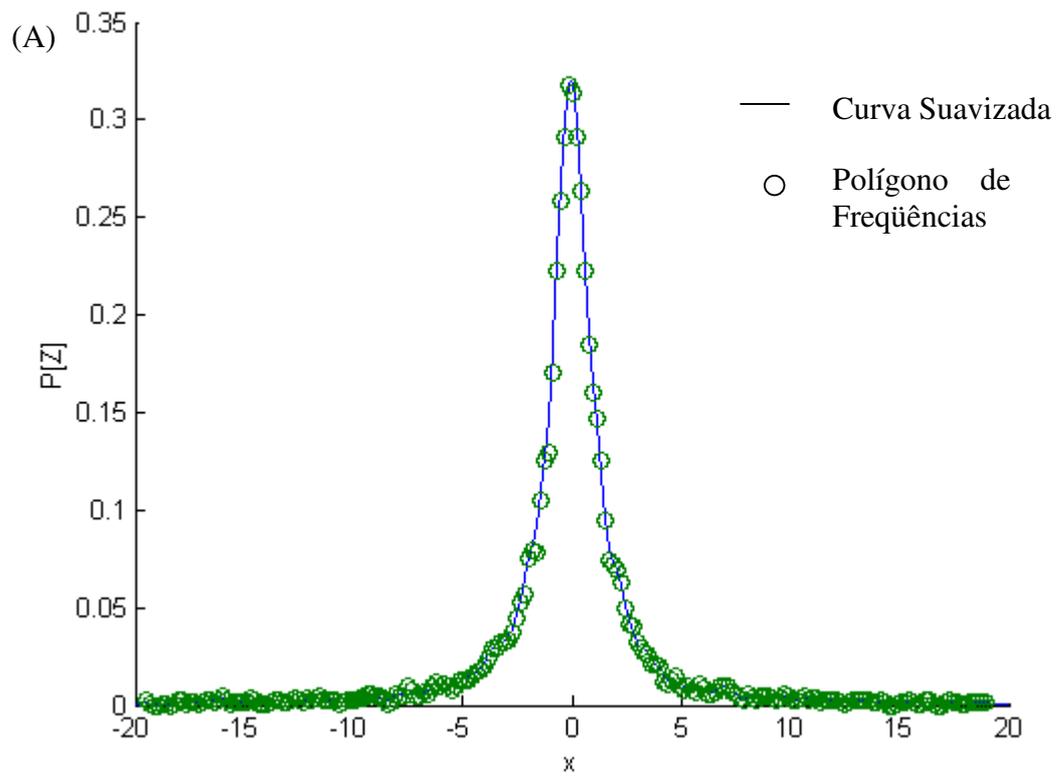


Fig. VII.6 – Comparação entre os gráficos de uma distribuição de Lévy (VII.4), com parâmetros dados por (VII.14), e o polígono de frequências do movimento Browniano simulado.

Pelo gráfico acima percebemos claramente que o modelo acima é confiável, pois o gráfico da distribuição de Lévy, com os parâmetros α e β calculados, se sobrepõe de forma precisa ao polígono de frequências empírico.

VII.2.2 Cálculo do Expoente de Estabilidade para um Processo de Cauchy

Nosso objetivo agora é o de obter o expoente de estabilidade e o fator de escala para um processo de Cauchy. Simulamos 5.000 pontos de uma série temporal com distribuição de Cauchy empregando o gerador de números aleatórios do Mathematica, e então elaboramos seu polígono de frequências, sua distribuição suavizada, e as comparamos com a distribuição teórica. As figuras VII.7 (A), (B), e (C) ilustram estas comparações.



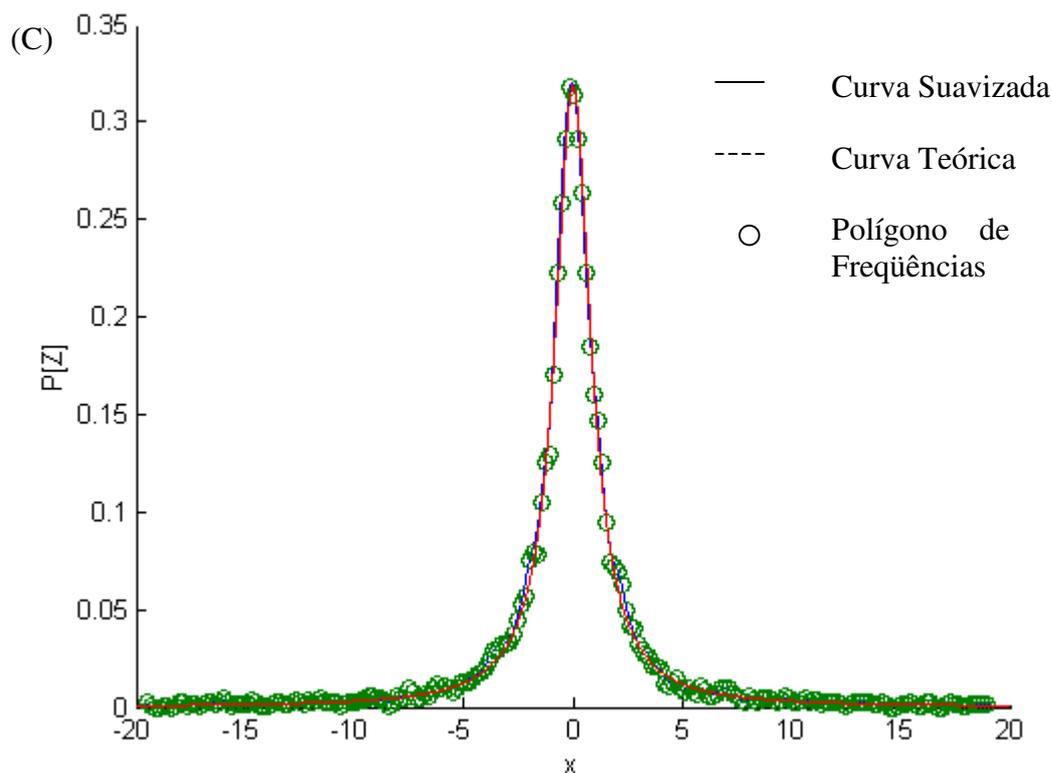


Fig VII.7 – Comparações entre (A) a distribuição suavizada com o polígono de frequências; (B) a distribuição suavizada e a distribuição teórica e (C), as três curvas.

Podemos observar por meio dos gráficos acima que a concordância entre os dados empíricos e a teoria é grande, e podemos até mesmo dizer que é melhor que a concordância obtida no caso do Movimento Browniano. Notamos, entretanto, que os pontos se acumulam nas caudas das distribuições, o que demonstra a tendência dos processos não-gaussianos a melhor descreverem os eventos raros.

Nosso próximo passo consiste em calcular o expoente de estabilidade e o fator de escala. Procedendo-se de modo análogo ao caso do Movimento Browniano, traçamos na figura VII.8 o gráfico $\log p \times \log \tau$ para a distribuição de Cauchy com a reta ajustada aos pontos; usando-se as equações (VII.12), obtemos os seguintes valores para o expoente de estabilidade e o fator de escala de uma distribuição de Cauchy

$$\alpha = 0.989 \pm 0.002 ,$$

$$\beta = 1.013 \pm 0.004 .$$

VII.15

Observemos que estes valores correspondem ao predito pela teoria. Substituindo tais parâmetros na expressão (VII.7) que define a distribuição de probabilidades de Lévy, obtemos a figura VII.9:

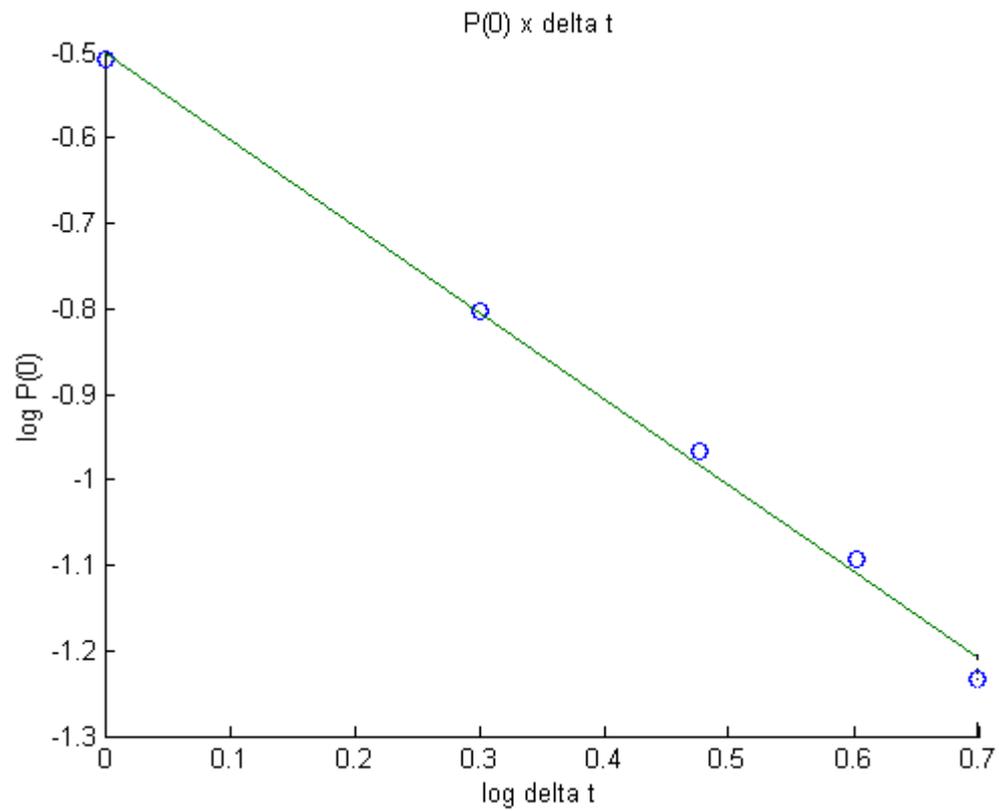


Fig VII.8 – Gráfico $\log p \times \log \tau$ para o cálculo dos coeficientes angular e linear.

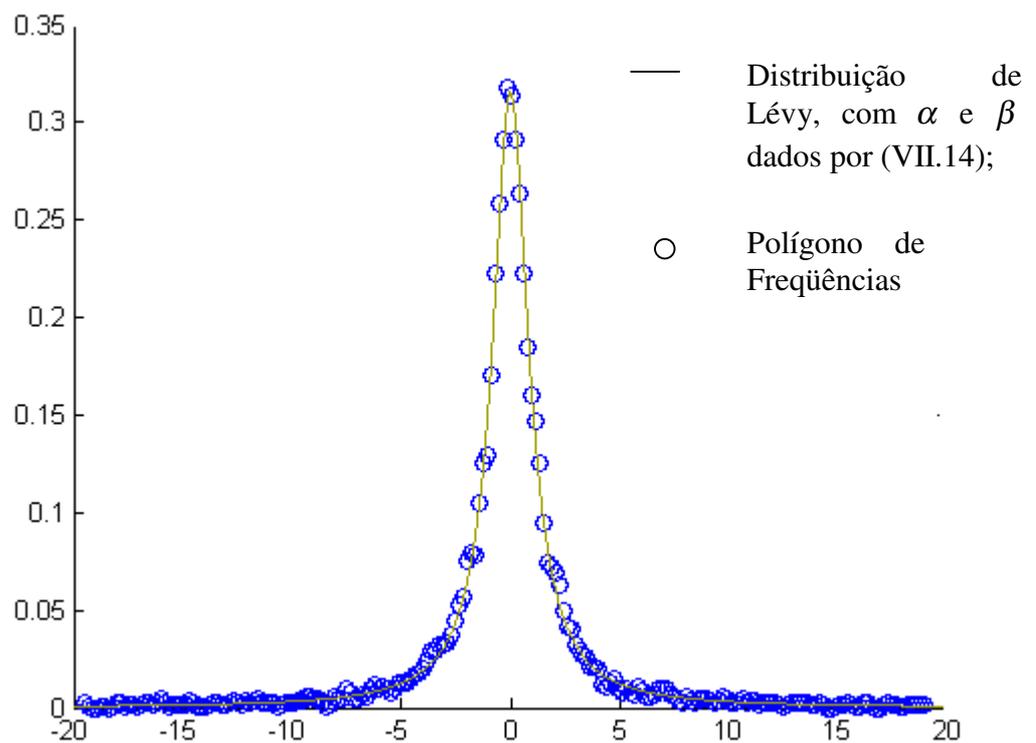


Fig. VII.9 – Comparação entre os gráficos de uma distribuição de Lévy (VII.4), com parâmetros dados por (VII.14), e o polígono de frequências de um processo de Cauchy simulado.

Novamente observamos a plena concordância dos modelos empírico e teórico, o que valida o método adotado para se efetuar a análise de processos estocásticos cuja distribuição não é conhecida *a priori*. É este ferramental que ilustramos nesta seção que empregaremos para nos responder à questão colocada no início deste capítulo: a dinâmica de preços dos ativos se comportam realmente como um Passeio Aleatório, ou como um Movimento Browniano Geométrico? É o que investigaremos na próxima seção.

VII.3 Cálculo do Expoente de Estabilidade para Séries Financeiras

Uma vez desenvolvidos os conceitos teóricos dos processos de Lévy, bem como os algoritmos computacionais para análise de dados, estamos aptos a investigar o expoente de estabilidade para séries financeiras. Utilizamos como dados a pontuação de fechamento do Índice IBOVESPA² entre os dias 15 de fevereiro de 1996 e 24 de abril de 1999. Os dados correspondem a medidas efetuadas a cada 15 segundos, perfazendo uma base de dados com um total de 14.684 observações. Sendo X_t um ponto qualquer desta série temporal, definiremos uma família $\{Z_t(\Delta t)\}$ de processos estocásticos definidos por

$$Z_t(\Delta t) = \log \frac{X_{t+\Delta t}}{X_t}, \quad \text{VII.16}$$

onde Δt corresponde a um dado intervalo de tempo. Escolhemos a família de processos (VII.16) como sendo o logaritmo natural do quociente entre duas observações pelo fato de ser esta uma boa alternativa ao uso de fatores de deflação ou de desconto.

Definida então a família de processos-base, vamos empregar o método desenvolvido na subseção anterior para o cálculo do expoente de estabilidade. A partir da série de dados, traçamos o gráfico dos logaritmos de $P(0)$ contra t (Figura VII.10), que nos forneceu os seguintes valores:

$$\alpha = 1.68 \pm 0.01,$$

$$\beta = 0.0365 \pm 0.0004. \quad \text{VII.17}$$

Vemos, portanto, que a distribuição não é gaussiana, visto que $\alpha \neq 2$. Com base nos dados experimentais, traçamos, para $\Delta t = 1$, a curva da função densidade de probabilidade suavizada, e a comparamos com a curva teórica da função densidade de probabilidade de Lévy para os valores (VII.17); incluímos também nesta comparação a função densidade de probabilidade normal para esta série de dados, em usando-se os valores empíricos $\mu = 0.0030893$ e $\sigma = 0.14603$ (Figura VII.11):

² Para maiores detalhes sobre o índice IBOVESPA, consultar a página 135, notas 3 e 4.

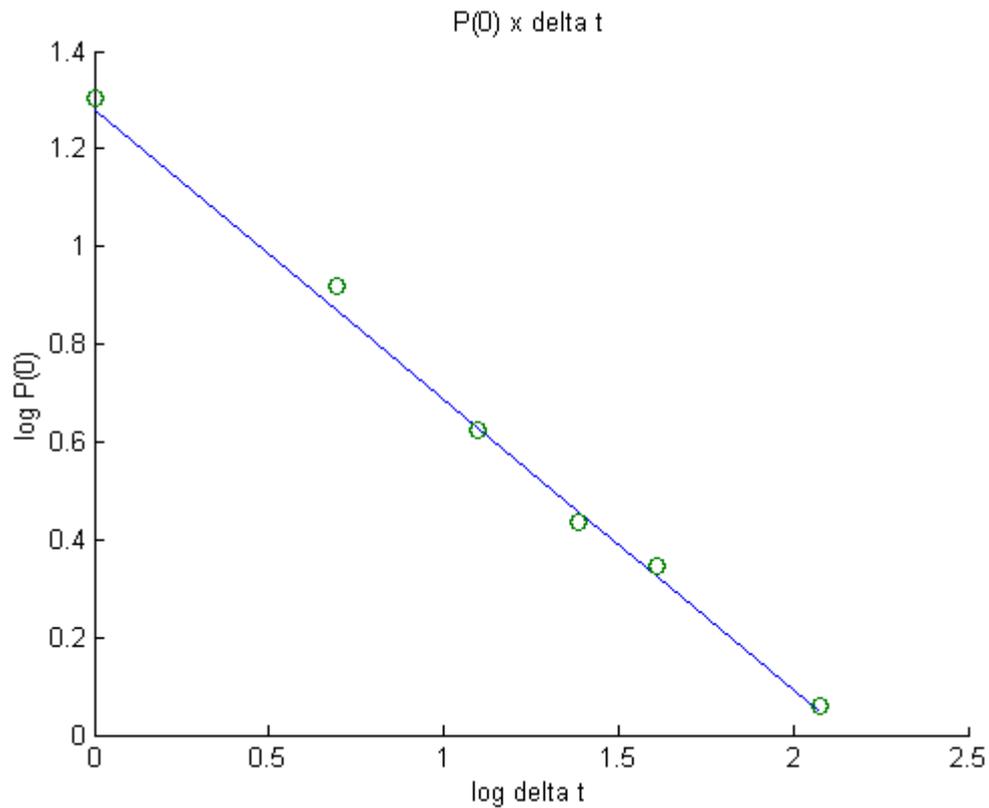


Fig VII.10 – Gráfico $\log p \times \log \tau$ para o cálculo dos coeficientes angular e linear.

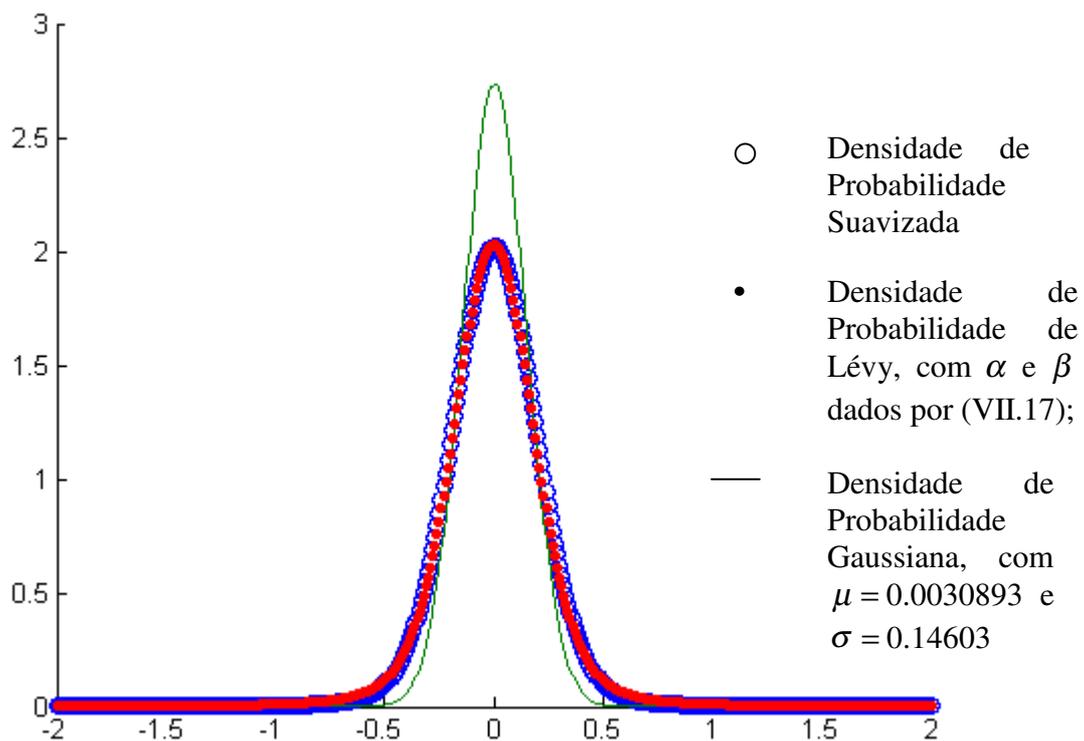


Fig VII.11 – Comparações entre as funções densidade de probabilidade empírica, de Lévy e gaussiana.

A comparação efetuada na Figura VII.11 nos demonstra claramente a não-normalidade da distribuição empírica de probabilidades do IBOVESPA. Observamos a grande divergência que ocorre nos picos da densidade gaussiana em relação à densidade empírica e à de Lévy:

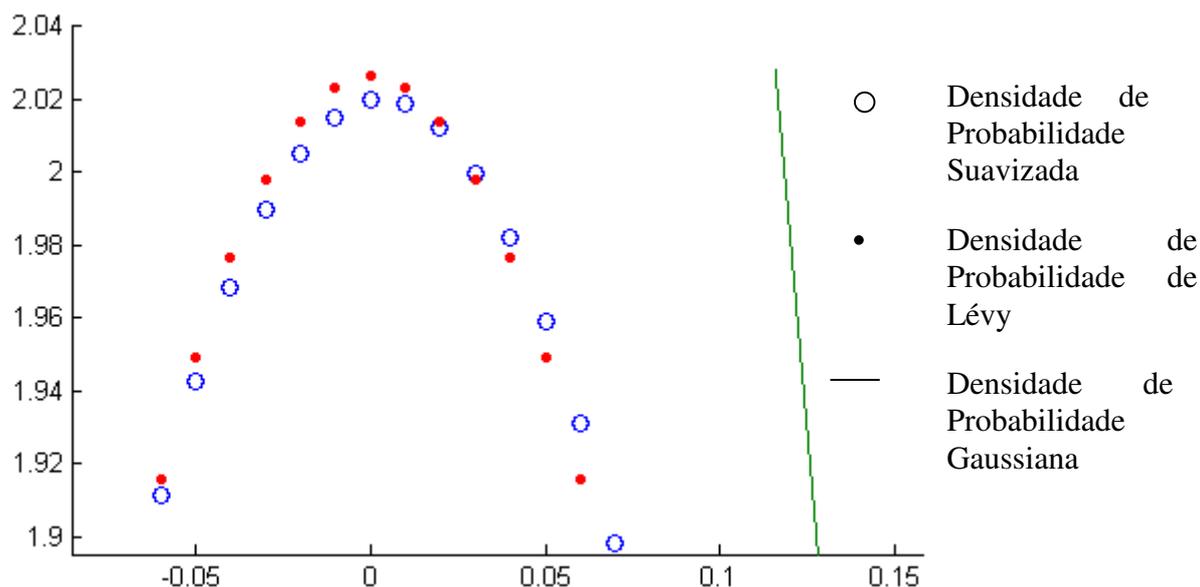


Fig VII.12 – Comparações entre os picos das densidades apresentadas na Figura VII.11

Para uma região abaixo do pico, a aproximação da densidade empírica por uma de Lévy também apresentou melhores resultados:

Para as caudas verificamos a grande divergência da densidade gaussiana para a empírica e a de Lévy:

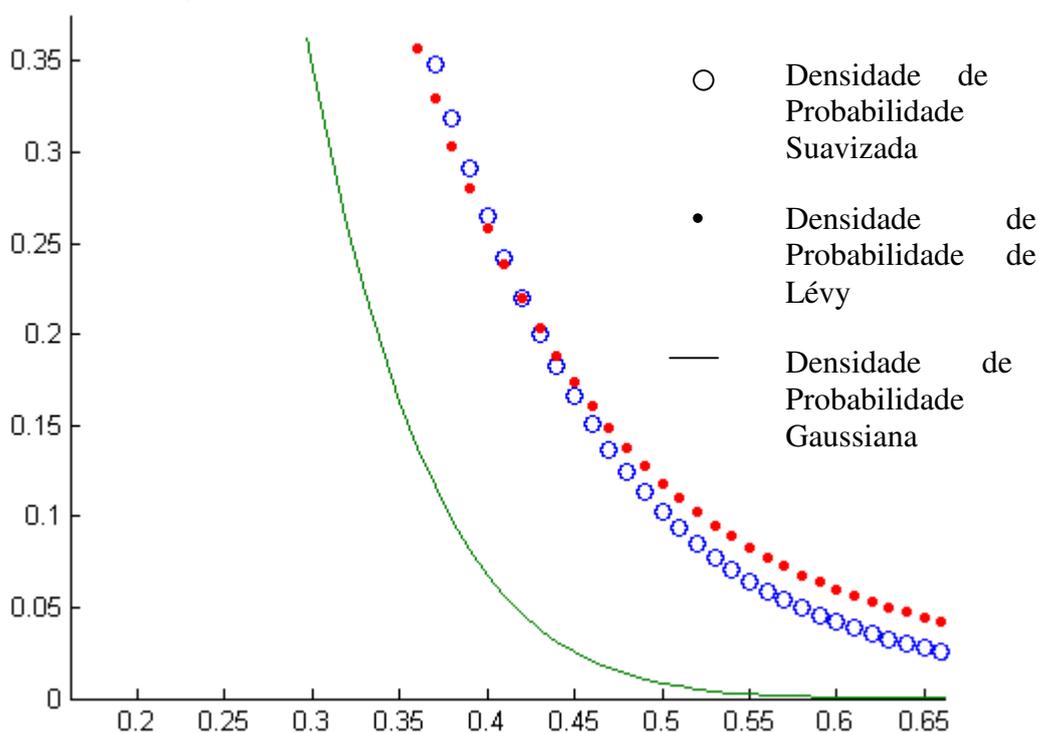


Fig VII.14 – Comparações entre as caudas das densidades apresentadas na Figura VII.11

Portanto, conforme discutimos no capítulo III e neste, as distribuições gaussianas não descrevem com boa aproximação os eventos raros. A densidade de Lévy tem um decaimento muito mais próximo ao da densidade empírica; estes só passarão a divergir significativamente para os valores de Z mais altos:

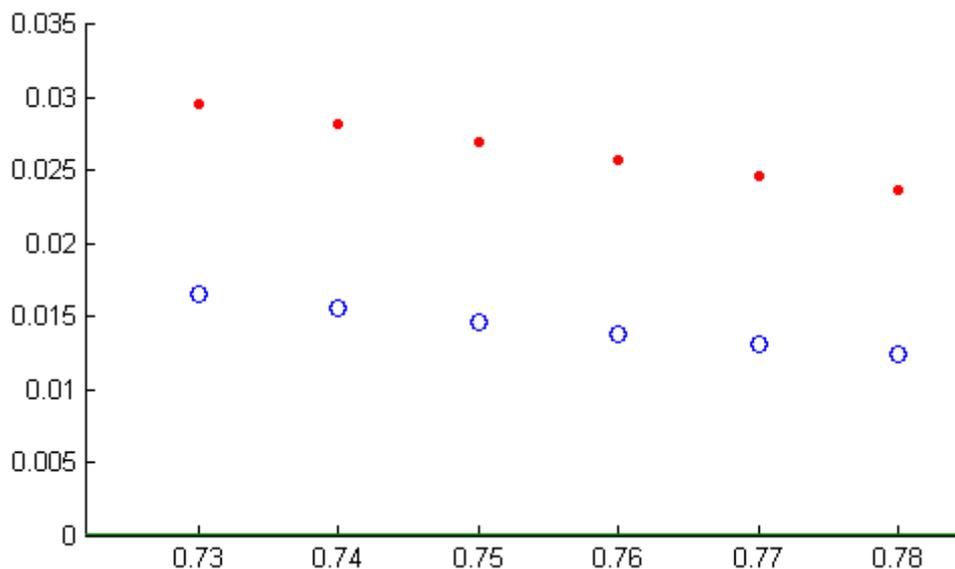


Fig. VII.14 – Comparações entre as caudas das densidades empírica e de Lévy

É notória a melhor aproximação propiciada pela densidade de Lévy com parâmetros (VII.17) do que a propiciada pela densidade gaussiana. Podemos concluir então que a densidade de probabilidade da carteira de ações que compõem o índice IBOVESPA é de Lévy com expoente de estabilidade igual a 1.68.

CONCLUSÕES

O tema fundamental deste trabalho se concentrou na Teoria do Movimento Browniano, tanto em sua formulação proveniente das leis da Física (Capítulo IV), quanto em sua formulação matemática (Capítulo V). Discutimos uma aplicação direta da Teoria do Movimento Browniano para a resolução do problema da precificação de opções (Capítulo VI), o que culminou na Teoria de Black & Scholes.

Analizamos a generalização do Movimento Browniano (Capítulo III) ao abordar a Teoria de Lévy, onde introduzimos os conceitos essenciais para a análise dos dados do IBOVESPA (Capítulo VII).

Para efetuar a análise de dados desenvolvemos um programa em MATLAB que calcula o expoente de estabilidade de uma dada série temporal. Partimos da expressão teórica de uma distribuição de Lévy, dada pela equação (VII.4), e pela propriedade de simetria de escala (VII.8); esta última permite calcular a probabilidade de retorno nulo para qualquer série temporal caracterizada por um parâmetro de simetria τ (equação VII.10). Com esta expressão em mãos, empregamos a prescrição descrita na página 156, baseada no artigo [24].

Uma dificuldade enfrentada logo de início relacionou-se com a obtenção de uma curva ‘lisa’ para a distribuição empírica de probabilidades das séries analisadas. Adotamos um procedimento não-paramétrico desenvolvido em [14].

Uma vez desenvolvido o algoritmo, testamo-lo para duas séries com expoentes de estabilidade conhecidos: uma série temporal com distribuição normal e outra com distribuição de Cauchy. Simulamos, então, duas séries de dados com estas distribuições, onde usamos os geradores de números aleatórios do MATLAB e do Mathematica. No caso da distribuição gaussiana, geramos uma série cujo expoente de estabilidade – calculado por meio do programa que desenvolvemos – forneceu o valor $\alpha = 2.003 \pm 0.002$, estando de acordo com o valor teórico $\alpha = 2$ característico desta distribuição. Do mesmo modo, o expoente de estabilidade calculado para a série com distribuição de Cauchy apresentou valor de $\alpha = 0.989 \pm 0.002$, estando também de acordo com o valor teórico $\alpha = 1$ esperado para este tipo de distribuição. Em ambas situações, obtivemos a curva de Lévy teórica para os expoentes determinados empiricamente, e comparamos todas as distribuições – empírica (polígono de frequências suavizado), teórica (Gauss e Cauchy, respectivamente) e a de Lévy (com o expoente calculado computacionalmente). Os gráficos comparativos encontram-se nas figuras VII.6 e VII.9.

Analizamos o expoente de estabilidade destas duas distribuições com α conhecido para exatamente testar o algoritmo desenvolvido. Tendo obtido resultados satisfatórios, passamos ao cálculo do expoente de estabilidade da série IBOVESPA.

O valor obtido para o expoente de estabilidade da série analisada, $\alpha = 1,68 \pm 0,01$, mostra que a mesma não apresenta uma estrutura temporal como Movimento Browniano, o que invalidaria a precificação de um ativo que tomasse esta série como base segundo a fórmula de Black & Scholes, equação VI.62. A análise gráfica das funções densidade de probabilidade – empírica, gaussiana para a média e desvio padrão calculados, e a curva teórica de Lévy para os parâmetros α e β encontrados (figura VII.11) – mostra que a função densidade de probabilidade gaussiana não se ajusta sobremaneira à densidade empírica, confirmando a constatação de que a série IBOVESPA apresenta um comportamento não-gaussiano.

A análise gráfica realizada na figura VII.11 e pelas figuras VII.12 e VII.13 – comparação dos picos e das caudas, respectivamente, das funções densidade de probabilidade empírica, gaussiana e a curva teórica de Lévy – demonstram claramente que o modelo mais adequado para a descrição da dinâmica de preços da série IBOVESPA é um processo de Lévy dotado de expoente de estabilidade $\alpha = 1,68 \pm 0,01$.

BIBLIOGRAFIA

- [1] – L. de Araújo Silva Neto, *Derivativos*, Ed. Atlas.
- [2] – P. Billingsley, *Probability and Measure* – John Wiley & Sons.
- [3] – Jean-Philippe Bouchaud et Didier Sornette – *The Black & Scholes option pricing problem in mathematical finance: generalization and extensions for a large class of stochastic processes* – J. Phys. I France, 4 (1994) 863.
- [4] – L. E. Clarke, *Random Variables* – Longman.
- [5] – D. Cohn, *Measure Theory* – Birkhäuser.
- [6] – J. Cowan, *Mathematics of Option Pricing*, Lecture Notes, University of Chicago.
- [7] – R. Durrett, *Stochastic Calculus*, MIR, Moscou.
- [8] – A. Einstein, *Investigations on the Theory of Brownian Motion* – Dover, New York.
- [9] – W. Feller - *An Introduction to Probability Theory and Its Applications* – John Wiley and Sons.
- [10] – P. Fernandez, *Teoria da Medida e Integração* – IMPA.
- [11] – I. Gihman, A. Skorohod, *Vvedeniye v Teorii Stokhasticheskikh Protsesov*, MIR, Moscou (Introdução à Teoria dos Processos Estocásticos).
- [12] – B. V. Gnedenko, *Theory of Probability* – Chelsea.
- [13] – B. V. Gnedenko, A. A. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables* – Addison Wesley.
- [14] – C. Gouriéroux and A. Monfort, *Time Series and Dynamic Models* – Cambridge University Press.
- [15] – P. Halmos, *Measure Theory* – Van Nostrand.
- [16] – T. Hida, *Brownian Motion*, Springer-Verlag, Berlim.
- [17] – J. Hull, *Options, Futures and other derivative securities*, Englewood Cliffs: Prentice-Hall.
- [18] – H. Itô, *An Introduction to Probability Theory* – Cambridge University Press.
- [19] – I. Karatzas, S. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, Berlim.

- [20] – Joseph Klafter, Michael Shlesinger, G. Zumofen – *Beyond Brownian Motion* – Physics Today, February 1996.
- [21] – P. Lévy, *Théorie de l'Addition des Variables Aléatoires*, Gauthier-Villars.
- [22] – D. Luenberger, *Investment Science*, Oxford.
- [23] – L. Lukács, *Fonctions Caractéristiques*, Gauthier-Villars.
- [24] – Rosario Mantegna and H. Eugene Stanley – *Scaling Behaviour in the Dynamics of an Economic Index* – Nature, 376 (1995).
- [25] – *Dynamical Theories of Brownian Motion* – Princeton University Press, Princeton.
- [26] – A.Pagan and A.Ullam, *Nonparametric Econometrics* – Cambridge University Press.
- [27] – D. Revuz, M. Yor, *Continuous Martingales and Brownian Motion*, Springer-Verlag, [Berlim.
- [28] – H. Risken, *The Fokker-Planck Equation* – Springer-Verlag, Berlim.
- [29] – G. Samorodnitsky, M. Taqqu, *Stable non-Gaussian Random Processes* – Chapman e Hall.
- [30] – J. Securato, *Decisões Financeiras em Condições de Risco*, Ed. Atlas.
- [31] – S. Shreve, *Stochastic Calculus and Finance*, Lecture Notes.
- [32] – Aris Spanos - *Probability Theory and Statistical Inference* – Cambridge University Press.
- [33] – Uspensky, *Introduction to Probability Theory* – Mc Graw-Hill.
- [34] – N. Wax, *Stochastic Processes and Noise* – Dover, New York.
- [35] – P. Wilmott, S. Howison, J. Dewynne, *The Mathematics of Financial Derivatives*, Cambridge University Press.

APÊNDICE

APLICAÇÕES DA TEORIA DE LÉVY À PRECIFICAÇÃO DE OPÇÕES

Vimos na seção anterior que a densidade de probabilidade de uma série financeira não é gaussiana, o que compromete os modelos desenvolvidos até então que assumem um processo gaussiano como processo estocástico base para a modelagem da dinâmica dos parâmetros financeiros. Faz-se necessária, portanto, uma nova abordagem da teoria financeira, em especial da teoria de Black & Scholes, uma vez que esta fundamenta-se totalmente em termos de um movimento Browniano.

A investigação de modelos que permitem uma estatística não-gaussiana é uma vasta área ainda em aberto, área esta em que físicos e matemáticos tem contribuído imensamente para os avanços da teoria financeira ([3], [20]). Discutiremos abaixo um modelo proposto por Bouchaud e Sornette [3] como alternativa à teoria de Black & Scholes, que permite o uso de um processo de Lévy para descrever a dinâmica de preços de um ativo financeiro. A aplicação destes conceitos para o mercado brasileiro, conforme fizemos na seção anterior, será estudada em um trabalho posterior; nos limitaremos, nesta dissertação, a somente discutir os princípios por trás deste modelo generalizado.

Consideremos o patrimônio de um banco formado por uma dada *call* C (exercício T , preço de exercício K), e seu ativo-objeto, cujo preço a cada instante t é dado pelo processo estocástico X_t . Suponhamos que realização do processo estocástico X_t seja dada pelo conjunto de números reais $\{x(t)\}$, e que o banco possua uma quantidade $\phi(x, t)$ do ativo-objeto. A variação total do patrimônio ΔW do banco neste contexto será dada então pelas seguintes contribuições:

- ✘ o valor da call em T , $C(x_0, K, T)$;
- ✘ a provável perda, igual a $-(x(T) - K)$ no caso de $x(T) > K$, e igual a zero, caso contrário;
- ✘ a perda ou o ganho obtido devido à variação dos preços do ativo dentro do período em consideração.

A segunda contribuição acima, dada pela provável perda, decorre do fato de o banco ter de emitir em T uma ação no caso de $x(T) > K$. A terceira contribuição decorre do fato de o preço do ativo variar no intervalo considerado, o que impacta o valor de seu patrimônio; então, a contribuição do valor dos ativos no patrimônio do banco será dado por

$$\int_0^T \phi(x, t) \frac{\partial x}{\partial t} dt . \quad \text{A.1}$$

Portanto, a variação do patrimônio do banco no intervalo de tempo $0 \leq t \leq T$ em termos da variação dos preços dos ativos, representada pela realização do processo estocástico $\{X_t\}$, será dada por

$$\Delta W = C(x_0, K, T) - \{x(T) - K\} \Theta(x(T) - K) + \int_0^T \phi(x, t) \frac{\partial x}{\partial t} dt, \quad A.2$$

onde $\Theta(x(T) - K)$ é uma função de Heaviside.

Se tomarmos o valor esperado de ΔW como sendo igual a zero, estaremos estabelecendo a condição onde nem o banco nem o investidor estejam ganhando; ou seja, esta é a condição do jogo justo. Neste caso, obtemos da equação (A.2) a seguinte expressão:

$$C(x_0, K, T) - E\{\{x(T) - K\} \Theta(x(T) - K)\} + \int_0^T \phi(x, t) E\left[\frac{\partial x}{\partial t}\right] dt = 0. \quad A.3$$

Consideremos que as inclinações $\frac{\partial x}{\partial t}$ sejam independentes para diferentes valores de t , ou seja, $E\frac{\partial x}{\partial t} = 0$; nesta situação simplificada, onde há a transição do valor do preço do ativo de x_0 em $t=0$, para $x(T)$ em $t=T$, suponhamos que a probabilidade de transição deste evento seja igual a $P[x, T|x_0, 0]$; portanto, da equação (A.3), segue-se que

$$C(x_0, K, T) = \int_K^\infty (x' - K) P[x', T|x_0, 0] dx'. \quad A.4$$

A equação (A.4) expressa o valor da opção na data de vencimento $t=T$, em termos de uma distribuição de probabilidades arbitrária P ; e é esta arbitrariedade que empregaremos para desenvolver a generalização da equação de Black & Scholes para uma distribuição de Lévy. Podemos mostrar que a equação (A.4) conduz a uma solução da equação de Black & Scholes no caso de P ser igual a uma distribuição log-normal. No caso de uma distribuição de Lévy do estilo

$$P[x, T|x_0, 0] = \frac{1}{(ZT)^{1/\alpha}} L_{\alpha, \beta} \left(\frac{x - x_0}{(ZT)^{1/\alpha}} \right), \quad A.5$$

onde Z é a generalização do conceito de volatilidade. O termo $(ZT)^{1/\alpha}$ desempenha o mesmo papel que o termo $(DT)^{1/2}$ no caso de um movimento Browniano. A função acima decai para o valor

$$L_{\alpha, \beta}(\xi) \sim \frac{\kappa(\alpha, \beta)}{\xi^{1+\alpha}}, \quad A.6$$

no limite $\xi \rightarrow \infty$. $\kappa(\alpha, \beta)$ é uma constante cujo valor depende de α e β . O limite acima pode ser representado pela condição $x - x_0 \gg (ZT)^{1/\alpha}$, o que implica no fato de a diferença

entre o preço atual e inicial ser muito superior à flutuação de mercado, medida pelo fator $(ZT)^{1/\alpha}$. Obtemos então, neste limite, a expressão

$$C(x_0, K, T) = \frac{\kappa(\alpha, \beta) ZT}{\alpha} \int_K^{\infty} \frac{d\xi}{\xi^\alpha}. \quad \text{A.7}$$

Resta-nos agora investigar a forma final do preço da opção em função dos valores de α e β . O caso mais relevante se relaciona a valores de α entre 1 e 2, conforme verificamos em nossa análise empírica do índice IBOVESPA. A equação (A.7) resulta então em

$$C(x_0, K, T) = \frac{\kappa(\alpha, \beta) ZT}{(\alpha - 1)\alpha} K^{1-\alpha}. \quad \text{A.8}$$

Vemos assim que é plenamente possível precificar opções por meio de um processo de Lévy, obtendo-se um equivalente da equação de Black & Scholes. A maior dificuldade na utilização de processos não-gaussianos reside no fato de as soluções não serem analíticas, devendo-se assumir determinadas condições no limite da função de distribuição para se obter o resultado.

Analisaremos em futuro trabalho as implicações da equação (A.8) na precificação de opções do mercado brasileiro, conforme fizemos com o índice IBOVESPA. Nosso intuito em acrescentar esta exposição aqui foi de ilustrar como seria possível a introdução de processos de Lévy na precificação de opções.