Determinação Entrópica do Preço Racional da Opção Européia Simples Ordinária sobre Ação e *Bond*:

uma Aplicação da Teoria da Informação em Finanças em Condição de Incerteza

José de Oliveira Siqueira

Tese apresentada ao Departamento de Administração da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo como requisito parcial para obtenção do título de Doutor em Administração

Área de Concentração: Métodos Quantitativos

Orientador: Prof. Dr. Abraham Sin Oih Yu

São Paulo

1999

Membros da Comissão Julgadora para Defesa da Tese

Presidente (orientador)

Prof. Dr. Abraham Sin Oih Yu

PhD pela Stanford University

Membros Externos à Faculdade de Economia, Administração e Contabilidade da USP (FEA-USP)

Prof. Dr. Max Henrique Machado Costa

PhD pela Stanford University

Professor da Universidade Estadual de Campinas (UNICAMP)

Prof. Richard Saito

PhD pela Stanford University

Professor da Escola de Administração de Empresas de São Paulo - Fundação Getúlio Vargas (EAESP-FGV)

Membros Internos

Prof. Dr. José Roberto Securato

Doutor pela FEA-USP

Prof. Dr. Joe Akira Yoshino

PhD pela University of Chicago

Reitor da Universidade de São Paulo

Prof. Dr. Jacques Marcovitch

Diretor da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. Eliseu Martins

Chefe do Departamento de Administração

Prof. Dr. Cláudio Felisoni de Angelo

Coordenador do Programa de Pós-Graduação em Administração

Prof. Dr. Ronaldo Zwicker

Ficha Catalográfica

Autor: José de Oliveira Siqueira

Título: Determinação entrópica do preço racional da opção européia simples ordinária sobre ação e *bond*: uma aplicação da Teoria da Informação em Finanças em condição de incerteza

Data: 17/12/1999

Unidade: FEA/USP

Departamento: Administração

Bibliografia: (vide anexo)

Paginação: 334 p.

Tese ou dissertação: tese de doutorado

Assuntos:

1. Determinação entrópica do preço racional da opção européia simples ordinária sobre ação e *bond*
2. Arbitragem, mercado completo, mundo com preço da incerteza neutralizado, linearidade de preços, medida de probabilidade neutralizadora do preço da incerteza, mundo martingale, movimento browniano, mundo informacionalmente eficiente e mundo entrópico
3. Teoria da Informação (incerteza, surpresa, informacional, entropia, princípio de máxima entropia e de mínima divergência, divergência de Kullback-Leibler, ganho informacional da previsão, taxa de entropia etc.) e Sistema de Distribuições de Pearson
4. Finanças em condição de incerteza, risco, derivativo e valor em risco (VaR)

Revisora do Português

Profa. Maria Nair Moreira Rebelo

Processador de texto

MS-Word 97 para Windows utilizando a fonte Euclid 12 do MathType 4.0

Processador de expressões matemáticas

MathType 4.0 para Windows utilizando a fonte TeX Look.eqp .

Processador matemático (simbólico, numérico e gráfico)

Mathematica 3.0 para Windows e MS-Excel 97 para Windows

DEDICATÓRIA

A Deus.

Dedico também esta tese a duas pessoas:

- Prof. Dr. Hélio Cardoso Chagas, meu mentor com quem tive o privilégio da convivência diária fortalecendo a comunhão intelectual construída ao longo deste processo de criação, pelo desvelo paternal de que tanto me beneficiei e por mostrar com seu exemplo o verdadeiro significado de estar intectualmente vivo.

- Sueli A. Gonçalves de Oliveira Siqueira, minha esposa, por ter arcado bravamente com o mundo real para que eu pudesse trabalhar no mundo das idéias.

AGRADECIMENTOS

Devo um agradecimento especial às seguintes pessoas e entidades:

- Prof. Dr. Abraham Sin Oih Yu (orientador): por ter aceito orientar uma tese téorica, pela amizade, pelo entusiasmo em difundir a Análise da Decisão em condição de incerteza na Administração, por ter indicado a leitura de teses de boa qualidade e pelo desejo de gerar conhecimento útil à Administração;
- Prof. Dr. Nicolau Reinhard: pelos conselhos oportunos, pelo apoio e, principalmente, pela confiança depositada;
- Prof. Dr. Adolpho Walter Pimazoni Canton: pelas inúmeras vezes que paciente e atenciosamente dirimiu minhas dúvidas estatísticas e me conduziu ao entendimento de meu papel como Estatístico e Pesquisador no Departamento de Administração da Faculdade de Economia, Administração e Contabilidade da USP;
- Prof. Dr. Fernando Carvalho de Almeida: pela amizade, por compartilhar da mesma cosmovisão, os mesmos ideais de vida e pelo estímulo à prática harmoniosa da atividade acadêmica e de consultoria/treinamento;
- Prof. Dr. José Roberto Securato: pelas inúmeras oportunidades proporcionadas na área de Finanças e pelo exemplo de ousadia, pioneirismo e entusiasmo acadêmico;
- Prof. Dr. Hélio Cardoso Chagas, Ricardo Sassatani (Mestre) e Fernando Hideki Kato (mestrando): pela co-autoria em diversas idéias contidas na tese, parceria em diversos estudos, artigos e discussões sobre os mais variados temas ligados a Finanças, Estatística e Teoria da Informação;
- Prof. Dr. Joe Akira Yoshino e Prof. Dr. Paulo Picchetti: membros da comissão do exame de qualificação pelas críticas e sugestões;
- Prof. Dr. Ronaldo Zwicker: por sua atuação efetiva na criação de uma autêntica atividade acadêmica em sua atuação na FEA-USP e pelo apoio financeiro para a reprodução e encadernação dos exemplares e para a revisão da tese;

- Marcelo Guterman: pela prática da autêntica amizade e pelas constantes discussões sobre o mundo das Finanças;
- Prof. Dr. Luiz João Corrar e Prof. Dr. Geraldo Barbieri: pelo empenho na aplicação dos métodos quantitativos nas Ciências Contábeis e pelas oportunidades proporcionadas nos cursos para executivos;
- Todos os professores, especialmente aos da área de Métodos Quantitativos e Informática e de Finanças do Departamento de Administração da FEA-USP: que contribuíram para minha formação e para o sucesso deste trabalho;
- Professores do Instituto de Matemática e Estatística da USP (IME-USP), especialmente aos do bacharelado em Estatística: pela dedicação e empenho em proporcionar uma excelente formação matemática, estatística e computacional;
- Colegas de pós-graduação André Luiz Oda e Gustavo Aleixo pelo compartilhamento dos momentos de descobertas em Finanças;
- Profa. Dra. Celma de Oliveira Ribeiro, Dr. André Cavalcanti Rocha Martins e Cibele Dunder (mestre) pela parceria no estudo de Finanças;
- Vera Lúcia Hemiko Sakamoto Obara e Lícia Mutsuko Abe: pelo apoio e dedicação constantes;
- Funcionários da FEA - especialmente aos da Seção de Pós-graduação sob o competente comando de Valéria Lourenção e aos do Serviço de Biblioteca e Documentação e da Seção de Atendimento aos Usuários da FEA sob a competente direção de Dulcineia Dilva Jacomini - Diretora Técnica do Serviço de Biblioteca da FEA-USP;
- Funcionários do Serviço de Biblioteca da biblioteca Prof. Carlos Benjamin de Lyra do IME-USP pela eficiência e atenção;
- Funcionários da Livraria Atlas, especialmente ao Humberto Sampaio Lima, pelos excelentes livros indicados de Estatística, Matemática, Informática e Métodos Quantitativos aplicados em Finanças e Economia que atualizaram-me intelectual e profissionalmente e pelo serviço personalizado e eficiente;

- Opus Dei: pela excelente formação doutrinal e ascética católica, teológica e filosófica aos seus membros e cooperadores e pela santidade que semeia e cultiva no mundo do trabalho;
- Johann Sebastian Bach (1685-1750): pela sua música, obra de uma vida dedicada a Deus e à busca da perfeição, que foi meu refrigério nas longas horas de árduo estudo;
- Agradeço, finalmente, a todos aqueles que, direta ou indiretamente, colaboraram para o sucesso deste trabalho.

Curriculum vitae

www.fea.usp.br/professores/siqueira.htm
siqueira@usp.br
0xx11-8185861

Professor em regime de dedicação integral à pesquisa e à docência (RDIPD) da área de Métodos Quantitativos e Informática do Departamento de Administração da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (FEA-USP) (1990-).

Ex-graduando do curso de Engenharia Elétrica da Escola Politécnica da Universidade de São Paulo (1984-7).

Bacharel em Estatística pelo Instituto de Matemática e Estatística da Universidade de São Paulo (IME-USP) (1985-9).

Mestre em Administração pela Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (FEA-USP) tendo desenvolvido uma dissertação sobre aplicação de *Conjoint Analysis* na área de Marketing (1990-5).

Doutor em Administração pela Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (FEA-USP) tendo desenvolvido uma tese sobre aplicação da Teoria da Informação em Finanças em condição de incerteza (1996-9).

Tem larga experiência com ferramentas de base, tais como programação orientada para objeto em C++, tendo livros publicados na área de informática pelas editoras Atlas e Pioneira.

Membro dos seguintes conselhos e associações:

- Associação Brasileira de Estatística (ABE)
- Região Brasileira da Sociedade Internacional de Biometria (RBRAS)
- Sociedade Brasileira para o Progresso da Ciência (SBPC)
- Sociedade Brasileira de Matemática (SBM)

- Sociedade Brasileira de Econometria (SBE)
- Conselho Regional de Estatística (CONRE – 3ª Região)
- Conselho Diretor de Informática da Livraria Pioneira Editora

Tem se dedicado à pesquisa de técnicas de tratamento e análise de dados, com diversas publicações nesta área.

Possui grande conhecimento sobre a família de softwares estatísticos da SPSS.

Sua preocupação profissional e de pesquisa tem sido a de explorar ferramentas de análise de dados que proporcionam um aumento de informação e de conhecimento do decisor a partir da exploração de bases de dados.

Tem trabalhos publicados de aplicações de estatística em Marketing.

Desenvolveu trabalhos de avaliação de cadastro e de crédito em empresas como ELETROPAULO (1994-95) e CEF (1996-97).

No Banco ITAÚ participou de um projeto de transferência de tecnologia de técnicas de *Data Mining* (árvore de classificação e regressão e redes neurais) juntamente com os Profs. Dr. Fernando Carvalho de Almeida e Dr. Hélio C. Chagas em 1998.

Em 1998-99 ministrou cursos de técnicas emergentes de *data mining* para os estatísticos e analistas do UNIBANCO e Cartão UNIBANCO juntamente com os Profs. Dr. Fernando Carvalho de Almeida e Dr. Hélio C. Chagas.

Tem explorado a avaliação de risco de crédito através de técnicas estatísticas e de redes neurais, tendo publicado trabalhos a este respeito no Brasil e no exterior.

Desenvolveu uma aplicação de Controle Estatístico de Processos - CEP - na área de Recursos Humanos da Souza Cruz para monitorização das fontes de pagamento aos funcionários (1998-99).

"*É por isso que não desfalecemos. Ainda que exteriormente se desconjunte nosso homem exterior, nosso interior renova-se de dia para dia. A nossa presente tribulação, momentânea e ligeira, nos proporcionará um peso de glória incomensurável. Porque não miramos as coisas que se vêem, mas sim as que não se vêem. Pois as coisas que se vêem são temporais e as que não se vêem são eternas.*"

Paulo de Tarso (São Paulo). Segunda Epístola aos Coríntios, cap. 4, versículos 16-8. *Bíblia Sagrada*. 116ª ed. Tradução dos originais mediante a versão dos Monges Maredsous (Bélgica) pelo Centro Bíblico Católico. São Paulo: Ave-Maria.

"*Toda universidade só tem o nome de 'universidade' se for um lugar de pesquisa, criação de conhecimento. Quando você dá aula, se é um pesquisador, suas aulas são diferentes das do outro, que não é pesquisador. Porque o que não é pesquisador vai procurar nos livros poeirentos o que foi, e o pesquisador dá conhecimentos antigos, clássicos, mas como ele pesquisa, apresenta de uma maneira sempre nova.*"

José Leite Lopes (1918-)
Entrevista em setembro de 1999 para a
Revista ADUSP publicada no número 18 (outubro).

Ideograma chinês para entropia

Resumo

Esta tese promove uma integração entre Finanças e Teoria de Informação para criação de um ambiente alternativo para a determinação do preço racional da opção européia simples ordinária sobre ação e ativo de renda fixa (*bond*). Uma das características deste novo ambiente de determinação de preço racional é poder continuar utilizando o cálculo newtoniano em vez do estocástico. Cria uma notação matemática precisa e completa para a Teoria da Informação e a integra com a teoria de Finanças em condições de incerteza. Integra as abordagens entrópicas de determinação do preço racional da opção européia simples ordinária de Gulko (1998 e 1998a) e de Yang (1997).

Define precisamente o mundo com preço da incerteza neutralizado (*risk-neutral world*), o mundo martingale, o mundo informacionalmente eficiente e o mundo entrópico e suas implicações para a Ciência do Investimento e, mais especificamente, para a determinação do preço racional de ativos básicos e derivativos.

Demonstra detalhadamente a fórmula do preço racional da opção européia simples ordinária de Black-Scholes-Merton, melhorando a notação matemática, simplificando (eliminando a abordagem martingale) e complementando a demonstração feita por Baxter & Rennie (1998).

Interrompe uma sucessão de trabalhos que estabelecem uma forma equivocada de calcular o preço da opção européia simples ordinária. Esse erro teve sua origem, muito provavelmente, numa edição de Brealey & Myers, que equivocadamente utilizou um resultado de Cox & Rubinstein (1985); esse resultado facilitava o cálculo do preço racional da opção européia simples ordinária por meio de uma tabela que evita o uso direto da fórmula de Black-Scholes-Merton. Brealey & Myers (desde a quarta edição de 1991), Luehrman (nos seus dois artigos da HBR de 1998 e um caso de 1995 pela HBS) e Edleson (caso publicado em 1994 pela HBS) ensinam que o valor percentual encontrado nessa tabela deve ser multiplicado pelo preço do valor mobiliário, quando deveria ser multiplicado pelo valor presente do preço de exercício.

Os resultados mais importantes desta tese para Finanças são: (i) desenvolvimento de um método alternativo, robusto e parcimonioso, baseado no princípio da máxima entropia da Teoria da Informação e do Sistema de Distribuições de Pearson para obtenção de uma única medida de probabilidade neutralizadora do preço da incerteza (*risk-neutral probability*), (ii) obtenção de fórmula prática para a determinação do preço racional da opção européia simples ordinária para ação, (iii) validação da fórmula de Black-Scholes-Merton para ação, (iv) obtenção de uma fórmula adequada para a determinação do preço racional da opção européia simples ordinária sobre um título de renda fixa (*bond*), (v) estimação da volatilidade implícita entrópica do preço do valor mobiliário e (vi) definição e estimação do valor em risco (*value at risk*) entrópico.

Há ainda dois resultados importantes para a Teoria da Informação e Economia: (i) distinção mais precisa entre incerteza e risco e (ii) desenvolvimento da medida de ganho informacional da previsão aprimorando o resultado de Theil (1967) e Benish (1999) pela utilização do conceito de divergência de Kullback-Leibler.

Abstract

This thesis integrates Finance and Information Theory in order to create an alternative environment to the calculation of the rational price of the simple ordinary European option over stocks and bonds. One of the features of this new environment is to allow us to continue using the Newtonian calculus instead of the stochastic one. It creates a precise and complete mathematical notation for the Information Theory and integrates it with the Finance Theory under uncertainty conditions. It integrates Gulko's (1998 and 1998a) and Yang's (1997) entropic approaches to the calculation of the rational price of the simple ordinary European option.

It precisely defines the uncertainty-price-neutral world (risk-neutral world), the martingale world, the informationally efficient world and the entropic world and their implications to the Investment Science and, more specifically, to the calculation of the rational price of ordinary assets and derivatives.

It demonstrates with details the Black-Scholes-Merton formula of the rational price of the simple ordinary European option, improves the mathematical notation, simplifies it (by eliminating the martingale approach) and completes the demonstration done by Baxter & Rennie (1998).

It breaks a succession of works that established a mistaken way to calculate the price of the simple ordinary European option. This mistake had its origin, much probably, in an edition of Brealey & Myers, who erroneously used a result from Cox & Rubinstein (1985). This result facilitates the calculation of the rational price of the simple ordinary European option by using a table that avoids the direct usage of the Black-Scholes-Merton formula. Brealey & Myers (since the 1991 fourth edition), Luehrman (in his two 1998 articles in HBR and in a 1995 case in HBS) and Edleson (1994 case published in HBS) teach that the percentage value found in this table must be multiplied by the price of the asset, when in reality it should have been multiplied by the present value of the strike price.

The most important results of this thesis for Finance are: (i) development of a robust and economic alternative method, based on the maximum-entropy principle of the Information Theory and on Pearson's Distribution System, to the calculation of a unique uncertainty-price-neutral probability measure (risk-neutral probability), (ii) achievement of a practical formula to the calculation of the rational price of the simple ordinary European option on stocks, (iii) validation of the Black-Scholes-Merton formula on stocks, (iv) achievement of an adequate formula to the calculation of the rational price of the simple ordinary European option on bonds, (v) estimation of the implied entropic volatility of the price of an asset and (vi) definition and estimation of the entropic value-at-risk.

There are still two important results to the Information Theory and to Economics: (i) a more precise distinction between uncertainty and risk and (ii) development of the forecast informational gain, an enhancement of the result of Theil (1967) and Benish (1999) by using the Kullback-Leibler divergence concept.

Quadros

QUADRO 1: EXEMPLOS DE OPÇÕES 7
QUADRO 2: VARIÂNCIA E ENTROPIA DE ALGUMAS DISTRIBUIÇÕES DISCRETAS 175
QUADRO 3: EFICIÊNCIA INFORMACIONAL RELATIVA DA PREVISÃO 194
QUADRO 4: EFICIÊNCIA INFORMACIONAL RELATIVA DA PREVISÃO - II 198
QUADRO 5: ESTIMATIVAS DAS PROBABILIDADES NEUTRALIZADORAS DO PREÇO DA INCERTEZA NO MERCADO COMPLETO E ENTRÓPICO 249
QUADRO 6: PREÇO RACIONAL DA OPÇÃO NO MERCADO COMPLETO E NO ENTRÓPICO 249
QUADRO 7: VOLATILIDADE IMPLÍCITA ENTRÓPICA DO PREÇO DA AÇÃO RCTB RPN PARA 21 DIAS ÚTEIS 270
QUADRO 8: VOLATILIDADE DE BLACK-SCHOLES-MERTON DO PREÇO DA AÇÃO RCTB RPN PARA 21 DIAS ÚTEIS 272
QUADRO 9: RESUMO DAS CONDIÇÕES PARA NECESSÁRIAS PARA A OBTENÇÃO DOS MODELOS DE BLACK-SCHOLES-MERTON E ENTRÓPICO 280

Figuras

FIGURA 1: ÁRVORE BINÁRIA 126
FIGURA 2: ÁRVORE BINÁRIA RECOMBINANTE 127
FIGURA 3: SIMULAÇÃO DE UMA TRAJETÓRIA DO MOVIMENTO BROWNIANO 143
FIGURA 4: PROCESSO ESTOCÁSTICO { Z_t ; $t \geq 0$ } 143
FIGURA 5: FUNÇÕES DENSIDADE DE PROBABILIDADE DO PROCESSO ESTOCÁSTICO { Z_t ; $t \geq 0$ } 144
FIGURA 6: LOGNORMAL ($\mu = 0,0009$; $\sigma = 0,025$; $s_0 = 147$; t ; Z_t) 147
FIGURA 7: RELAÇÕES ENTRE AS ENTROPIAS E DIVERGÊNCIAS DOS ESTADOS INFORMACIONAIS 190
FIGURA 8: GRÁFICO DA SUPERFÍCIE DA DIVERGÊNCIA 191
FIGURA 9: GRÁFICO DE CURVAS DE NÍVEL DA DIVERGÊNCIA 191
FIGURA 10: GRÁFICO DE SUPERFÍCIE DA DIVERGÊNCIA("+") QUANDO SENSIBILIDADE=ESPECIFICIDADE 195
FIGURA 11: GRÁFICO DE DENSIDADE DA DIVERGÊNCIA("+") QUANDO SENSIBILIDADE=ESPECIFICIDADE 195
FIGURA 12: GRÁFICO DA DIVERGÊNCIA TOTAL MÁXIMA QUANDO SENSIBILIDADE=ESPECIFICIDADE=1 196
FIGURA 13: GRÁFICO DE SUPERFÍCIE DA DIVERGÊNCIA TOTAL QUANDO SENSIBILIDADE=ESPECIFICIDADE 196
FIGURA 14: GRÁFICO DE DENSIDADE DA DIVERGÊNCIA TOTAL QUANDO SENSIBILIDADE=ESPECIFICIDADE 197
FIGURA 15: REPRESENTAÇÃO GRÁFICA EM FORMA DE BARRAS DAS ENTROPIAS ENTRE DUAS VARIÁVEIS ALEATÓRIAS DEPENDENTES 213
FIGURA 16: REPRESENTAÇÃO GRÁFICA EM FORMA DE BARRAS DAS ENTROPIAS ENTRE DUAS VARIÁVEIS ALEATÓRIAS INDEPENDENTES 216
FIGURA 17: REPRESENTAÇÃO GRÁFICA EM FORMA DE BARRAS DAS ENTROPIAS DE UM PROCESSO ESTOCÁSTICO MARKOVIANO 217
FIGURA 18: GRÁFICO DE SUPERFÍCIE DA TAXA DE ENTROPIA DE UM PROCESSO MARKOVIANO ESTACIONÁRIO DE DOIS ESTADOS 219
FIGURA 19: GRÁFICO DE CURVAS DE NÍVEL DA TAXA DE ENTROPIA DE UM PROCESSO MARKOVIANO ESTACIONÁRIO DE DOIS ESTADOS 220
FIGURA 20: ENTROPIA *VERSUS* VOLATILIDADE [GULKO (1998, P. 1-8)] 245
FIGURA 21: DISTRIBUIÇÃO DE F_τ 249
FIGURA 22: GRÁFICO DA FUNÇÃO DE DENSIDADE DE PROBABILIDADE GAMA(10;6) 265
FIGURA 23: GRÁFICO DA FUNÇÃO DE DENSIDADE DE PROBABILIDADE GAMA(1;2) 265
FIGURA 24: PREÇO RACIONAL DA OPÇÃO DE COMPRA *VERSUS* A VOLATILIDADE IMPLÍCITA ENTRÓPICA DO PREÇO DA AÇÃO 269
FIGURA 25: VOLATILIDADE IMPLÍCITA ENTRÓPICA DO PREÇO DA AÇÃO *VERSUS* O PREÇO DE EXERCÍCIO 270
FIGURA 26: VOLATILIDADE IMPLÍCITA ENTRÓPICA E DE B-S-M DO PREÇO DA AÇÃO *VERSUS* O PREÇO DE EXERCÍCIO 272
FIGURA 27: FUNÇÃO DE DENSIDADE DE PROBABILIDADE NEUTRALIZADORA DO PREÇO DA INCERTEZA ENTRÓPICA DO PREÇO DA AÇÃO PARA A VARIÁVEL DE ESTADO TAXA DE RETORNO (LOGNORMAL - REPRESENTADA PELA LINHA CONTÍNUA) VERSUS A DA VARIÁVEL DE ESTADO PREÇO DA AÇÃO (GAMA - REPRESENTADA PELA LINHA TRACEJADA) COM VOLATILIDADE IMPLÍCITA ENTRÓPICA IGUAL A 11 REAIS 277
FIGURA 28: CAUDAS AMPLIADAS DA FUNÇÃO DE DENSIDADE DE PROBABILIDADE NEUTRALIZADORA DO PREÇO DA INCERTEZA ENTRÓPICA DO PREÇO DA AÇÃO PARA A VARIÁVEL DE ESTADO TAXA DE RETORNO (LOGNORMAL - REPRESENTADA PELA LINHA CONTÍNUA) E DA VARIÁVEL DE ESTADO PREÇO DA AÇÃO (GAMA - REPRESENTADA PELA LINHA TRACEJADA) COM VOLATILIDADE IMPLÍCITA ENTRÓPICA IGUAL A 11 REAIS 278

SUMÁRIO

1 INTRODUÇÃO 1

1.1 JUSTIFICATIVA DO TEMA E MOTIVAÇÃO 2

1.1.1 A importância e o preço da incerteza na Economia 2

1.1.2 O conceito geral de opção 3

1.1.3 Mundo com o preço da incerteza neutralizado e a medida de probabilidade neutralizadora do preço da incerteza 8

1.1.4 Teoria da Informação: uma abordagem alternativa para quantificar a incerteza e determinar preço 9

1.2 FORMULAÇÃO DA QUESTÃO DE PESQUISA 10

1.3 OBJETIVO DA TESE 10

1.4 METODOLOGIA 11

1.5 CONTRIBUIÇÕES ORIGINAIS E RESULTADOS PRINCIPAIS DA TESE 15

1.6 ESCOPO DA TESE 18

1.7 ESTRUTURA DA TESE 20

1.8 NOTAÇÃO MATEMÁTICA UTILIZADA NA TESE E TERMOS TÉCNICOS 20

2 FUNDAMENTAÇÃO TÉORICA 23

2.1 A CIÊNCIA DO INVESTIMENTO 23

2.1.1 Entes do mercado financeiro 24

2.1.1.1 Pessoas 24

2.1.1.2 Ativos e instrumentos 25

2.1.1.3 Opções 28

2.1.1.3.1 Um pouco de história 31

2.1.1.4 Preços 31

2.1.1.4.1 Um pouco de história 33

2.1.1.5 Mercados 34

2.1.2 Os princípios fundamentais da Ciência do Investimento 35

2.1.2.1 Princípio da insaciabilidade 36

2.1.2.1.1 Exemplo paradigmático de aplicação do princípio da insaciabilidade 37

2.1.2.2 Princípio da comparação 37

2.1.2.2.1 Exemplo paradigmático de aplicação do princípio da comparação 37

2.1.2.3 Princípio de não-arbitragem 38

2.1.2.3.1 Exemplo paradigmático de aplicação do princípio de não-arbitragem 39

2.1.2.3.2 Exemplo paradigmático da aposta entre o circunspeto professor de Economia e o matuto .. 39

2.1.2.3.3 Exemplo paradigmático de determinação de preço de opção e da aplicação do teorema da arbitragem 41

2.1.2.4 Princípio da aversão à incerteza 45

2.1.2.4.1 Quantificação da incerteza 47

2.1.2.4.2 Incerteza técnica e de mercado 48

2.1.2.4.3 Um pouco de história 49

2.1.2.4.4 Relação entre risco financeiro e incerteza 50

2.1.2.4.4.1 Gestão do risco 54

2.1.2.4.4.1.1 Exemplo paradigmático de gestão de risco 54

2.1.2.4.4.2 Um pouco de história 57

2.1.3 Hipóteses normalmente utilizadas para determinação do preço racional de opção européia simples ordinária 59

2.1.3.1 Hipótese de mercado perfeito 59

2.1.3.1.1 Banco ideal constante 60

2.1.3.2 Hipótese de mercado competitivo, livre e aberto como vários investidores igualmente importantes 61

2.1.3.3 Hipótese do equilíbrio competitivo do mercado de capitais 63

2.1.3.4 Hipótese de mercado completo 63

2.1.3.5 Hipótese de mercado eficiente 64

2.1.3.6 Hipótese da atitude do decisor diante da incerteza 67

2.1.3.7 Hipótese do mundo com preço da incerteza neutralizado 68

2.1.3.7.1 Mudança de medida de probabilidade ... 75
2.1.3.7.1.1 Exemplo de mudança de medida de probabilidade de variável aleatória discreta ... 75
2.1.3.7.2 Processo estocástico martingale ... 79
2.1.3.7.3 Integral de Lebesgue: esperança e variância ... 81
2.1.3.7.4 Esperança parcial ... 81
2.1.3.7.5 Função de perda linear ... 81
2.1.4 Problemas de investimento típicos ... 82
2.1.4.1 Determinação de preço ... 82
2.1.4.1.1 Exemplo paradigmático de determinação de preço em condição de certeza ... 83
2.1.4.1.2 Exemplos paradigmáticos de determinação de preço em condição de incerteza de mercado 83
2.1.4.1.2.1 Contrato a termo (forward) ... 83
2.1.4.1.2.2 Contrato de opção européia simples ordinária sobre ação ... 85
2.1.4.2 Estratégia de proteção contra flutuações do preço do valor mobiliário (hedging) ... 86
2.1.4.2.1 Exemplo paradigmático de hedging ... 86
2.1.4.3 Investimento puro ... 87
2.1.4.4 Outros problemas ... 87
2.1.5 Relação entre o mundo com preço da incerteza neutralizado e a atitude do decisor diante da incerteza ... 87
2.1.6 Relação entre a hipótese de mercado completo, opções e a hipótese do mundo com preço da incerteza neutralizado ... 89
2.1.6.1 Mercado completo uniperíodo ... 89
2.1.6.1.1 Existência e unicidade da medida de probabilidade neutralizadora do preço da incerteza em mercado completo uniperíodo ... 89
2.1.6.2 Utilidade do mercado completo ... 95
2.1.6.2.1 A contribuição das opções ... 96
2.1.7 Relação entre hedging perfeito e a hipótese do mundo com preço da incerteza neutralizado 98
2.1.7.1 Determinação de rateios (odds) por um bookmaker para hedgear uma aposta (estratégia de seguro dos resultados) ... 99
2.1.7.2 Cálculo do preço de uma opção de abandono de projeto com cobertura completa de risco por meio de seguro governamental ... 102
2.1.7.3 Exemplo paradigmático do apreçamento racional de uma aposta digital ... 107
2.2 MODELO DE BLACK-SCHOLES-MERTON PARA DETERMINAÇÃO DO PREÇO RACIONAL DA OPÇÃO EUROPÉIA SIMPLES ORDINÁRIA SOBRE AÇÃO ... 110
2.2.1 A opção e suas propriedades ... 110
2.2.1.1 Condição de contorno ... 111
2.2.2 História das fórmulas de determinação do preço racional da opção européia simples ordinária sobre ação ... 112
2.2.2.1 Modelos pré-Black-Scholes-Merton (1900-1964) ... 112
2.2.2.1.1 Modelo de Bachelier (1900) ... 112
2.2.2.1.2 Modelo de Sprenkle (1961) ... 113
2.2.2.1.3 Modelo de Boness (1964) ... 113
2.2.2.2 Modelo de Black-Scholes-Merton (1973) ... 114
2.2.2.3 Modelos pós-Black-Scholes-Merton (até 1998) ... 116
2.2.3 Demonstração da fórmula de Black-Scholes-Merton ... 119
2.2.3.1.1 Mercado completo multiperíodo ou dinamicamente completo ... 120
2.2.3.1.1.1 Ramo binário em um processo discreto ... 122
2.2.3.1.1.2 Construção de derivativos para um período a partir de uma carteira formada por ações e bonds 123
2.2.3.1.1.3 A medida de probabilidade neutralizadora do preço da incerteza $\mathbb{Q}$ ... 124
2.2.3.1.1.4 Propriedades da medida de probabilidade neutralizadora do preço da incerteza $\mathbb{Q}$ ... 125
2.2.3.1.1.5 A árvore binária ... 125
2.2.3.1.1.6 A árvore binária recombinante ... 126
2.2.3.1.1.7 Indução de trás para diante ... 127
2.2.3.1.1.8 Carteira replicante auto-financiável ... 128
2.2.3.1.1.9 Comportamento do preço do ativo em uma árvore recombinante: a abertura para modelos contínuos 130
2.2.3.1.1.10 Criando a medida de probabilidade neutralizadora do preço da incerteza $\mathbb{Q}$ ... 134
2.2.3.1.1.11 Fórmula do movimento do preço do ativo-base ... 135

2.2.3.1.1.12 Fórmula de Black-Scholes-Merton de determinação do preço racional da opção de compra européia simples ordinária ... 136
2.2.3.1.1.13 Simplificação do cálculo do preço da opção ... 138
2.2.3.1.2 Abordagem contínua para o cálculo do preço de opções européias ... 139
2.2.3.1.2.1 Introdução ... 139
2.2.3.1.2.2 Processo de Wiener ... 141
2.2.3.1.2.3 Fórmula do preço do ativo ... 146
2.2.3.1.2.4 Lema de Ito e sua aplicação sobre o diferencial total estocástico do ativo ... 148
2.3 ELEMENTOS DA TEORIA DA INFORMAÇÃO ... 152
2.3.1 Um pouco de história ... 153
2.3.2 O conceito físico de entropia discreta de Boltzmann ... 153
2.3.3 O conceito lógico de entropia discreta de Shannon ... 160
2.3.3.1 Quantificação da surpresa no exemplo da eleição para governador do Distrito Federal ... 160
2.3.3.2 Auto-informação ou informacional de um evento ... 161
2.3.3.3 Dedução axiomática do informacional ... 163
2.3.3.4 Entropia ... 168
2.3.3.5 Exemplo paradigmático do jogo de adivinhar o número pensado por outra pessoa ... 171
2.3.3.6 Entropia de algumas variáveis aleatórias discretas ... 175
2.3.4 Divergência de Kullback-Leibler ... 177
2.3.4.1 Mensuração do ganho informacional da previsão e da eficiência informacional relativa da previsão usando o conceito de divergência: uma extensão do análise bayesiana ... 177
2.3.4.1.1 Exemplo paradigmático da compra de um carro usado ... 177
2.3.4.2 Estados informacionais do decisor ... 178
2.3.4.3 Análise informacional da decisão em condição de incerteza ... 184
2.3.4.4 Eficiência informacional relativa da previsão ... 192
2.3.4.5 Medida entrópica de ganho de informação aplicada em Economia ... 198
2.3.5 Entropia diferencial ... 206
2.3.5.1 Entropia de uma variável aleatória uniforme contínua ... 207
2.3.5.2 Entropia de uma variável aleatória normal contínua ... 208
2.3.6 Entropia diferencial como limitante da entropia discreta ... 209
2.3.7 Entropia mútua ou informação mútua ... 209
2.3.8 Interpretação gráfica das medidas de entropia ... 213
2.3.9 Versões contínuas das medidas de entropia ... 213
2.3.10 Relação entre entropia mútua e divergência ... 214
2.3.11 Entropia mútua da normal bivariada e sua relação com o coeficiente de correlação e a independência Estatística ... 214
2.3.12 Entropia de um processo estocástico markoviano ou taxa de entropia ... 215
2.3.13 Entropia mútua total ... 220
2.3.14 Relação entre a entropia de Shannon e a de Boltzmann ... 221
2.3.15 Entropia de ordem α e a distinção entre processos caóticos e aleatórios ... 222
2.3.16 Divergência de Amari e a função utilidade HARA ... 223
2.3.17 Processamento ótimo de informação incompleta ou estimação robusta com dados limitados utilizando o princípio da máxima entropia ou da mínima divergência ... 224
2.3.17.1 Exemplo paradigmático do princípio de máxima entropia aplicado a um problema com indeterminação ... 226
2.3.17.1.1 Qual a função densidade de probabilidade que maximiza a entropia quando o domínio da variável aleatória é finito? ... 229
2.3.17.1.2 Qual a função densidade de probabilidade que maximiza a entropia quando o domínio da variável aleatória é positivo e sua esperança é conhecida? ... 230
2.3.17.1.3 Qual a função densidade de probabilidade que maximiza a entropia quando o domínio da variável aleatória é infinito e sua variância é conhecida? ... 232
2.3.17.2 Princípio de mínima divergência ... 233
2.3.17.3 Axiomas de consistência ... 234

3 DETERMINAÇÃO ENTRÓPICA DO PREÇO RACIONAL DA OPÇÃO EUROPÉIA SIMPLES ORDINÁRIA SOBRE AÇÃO E *BOND* ... 236

3.1 INTRODUÇÃO ... 236
3.2 O MUNDO INFORMACIONALMENTE EFICIENTE É ENTRÓPICO ... 238
3.3 EXPECTATIVA DE MERCADO ... 239

3.4 EXEMPLO PARADIGMÁTICO DA DETERMINAÇÃO DO PREÇO RACIONAL DE UMA OPÇÃO RACIONAL PELO PRINCÍPIO DA MÁXIMA ENTROPIA 241
3.4.1 Propriedades empíricas da medida de probabilidade neutralizadora do preço da incerteza 242
3.4.1.1 Propriedade empírica 1: a distribuição da medida de probabilidade neutralizadora do preço da incerteza é em forma de sino (unimodal) 242
3.4.1.2 Propriedade empírica 2: o desvio-padrão na medida de probabilidade neutralizadora do preço da incerteza de F_T é praticamente igual à volatilidade real de F_T 243
3.4.1.3 Propriedade empírica 3: nenhuma variação unilateral de um preço de opção por um tick (0,01) aumenta a entropia acima de $h_{\mathbb{Q}}(F_T)$ dada a restrição de não poder aumentar o desvio-padrão acima de $\sigma_{\mathbb{Q}}(F_T)$. 244
3.6 DETERMINAÇÃO DA MEDIDA DE PROBABILIDADE NEUTRALIZADORA DO PREÇO DA INCERTEZA DE MÁXIMA ENTROPIA 252
3.6.1 Sistema de Distribuições de Pearson 255
3.6.2 Relação entre o Sistema de Distribuições de Pearson e o princípio da máxima entropia 256
3.6.3 Incorporação do sistema de Pearson no princípio de máxima entropia 258
3.7 DETERMINAÇÃO DA MEDIDA DE PROBABILIDADE NEUTRALIZADORA DO PREÇO DA INCERTEZA DE MÁXIMA ENTROPIA DO PREÇO DO BOND 259
3.8 DETERMINAÇÃO DO PREÇO ÚNICO É LIVRE DE ARBITRAGEM DA OPÇÃO DE COMPRA EUROPÉIA SIMPLES ORDINÁRIA SOBRE UM BOND 261
3.9 DETERMINAÇÃO DA MEDIDA DE PROBABILIDADE NEUTRALIZADORA DO PREÇO DA INCERTEZA DE MÁXIMA ENTROPIA DO PREÇO DE AÇÃO 262
3.10 DETERMINAÇÃO DO PREÇO ÚNICO É LIVRE DE ARBITRAGEM DA OPÇÃO DE COMPRA EUROPÉIA SIMPLES ORDINÁRIA SOBRE UMA AÇÃO 266

4 ESTIMAÇÃO DA VOLATILIDADE IMPLÍCITA 267

4.1 INTRODUÇÃO 267
4.2 EXEMPLO PARADIGMÁTICO DA ESTIMAÇÃO DA VOLATILIDADE IMPLÍCITA DO PREÇO DA AÇÃO RCTB RPN 268
4.3 VALOR EM RISCO ENTRÓPICO 273
4.4 PREÇO MÍNIMO ATUAL 273
4.5 PREÇO MÍNIMO ATUAL ENTRÓPICO 274
4.6 PERDA MÁXIMA ATUAL ENTRÓPICA 274
4.7 PERDA MÁXIMA ATUAL ENTRÓPICA PADRONIZADA 274
4.8 COMPARAÇÃO DA MEDIDA DE PROBABILIDADE NEUTRALIZADORA DO PREÇO DA INCERTEZA DA TAXA DE RETORNO (MODELO ENTRÓPICO LOGNORMAL) COM A DO PREÇO DA AÇÃO (MODELO ENTRÓPICO GAMA) 275

5 CONCLUSÕES, LIMITAÇÕES, RECOMENDAÇÕES E TEMAS PARA PESQUISAS 279

5.1 CONCLUSÃO DA TESE 279
5.2 TESTE DO MODELO ENTRÓPICO 281
5.2.1 Os prós e contras da modelagem Black-Scholes-Merton versus a entrópica para ação 281
5.2.1.1 Prós da modelagem de Black-Scholes-Merton para o objeto subjacente ação: 281
5.2.1.2 Contras da modelagem de Black-Scholes-Merton para o objeto subjacente ação: 282
5.2.1.3 Prós da modelagem entrópico para o objeto subjacente ação: 283
5.2.1.4 Contras da modelagem entrópica para o objeto subjacente ação: 285
5.3 LIMITAÇÕES 286
5.4 PESQUISAS FUTURAS 286

6 APÊNDICE 289

6.1 PROGRAMAÇÃO DO MODELO DE BLACK-SCHOLES-MERTON E ENTRÓPICO EM VISUAL BASIC 289
6.2 PROGRAMAÇÃO DO MODELO DE BLACK-SCHOLES-MERTON E ENTRÓPICO DE DETERMINAÇÃO DO PREÇO RACIONAL DA OPÇÃO DE COMPRA EUROPÉIA SIMPLES ORDINÁRIA SOBRE UMA AÇÃO NO MATHEMATICA 3.0 293
6.2.1 Fórmula do preço racional da opção de compra européia simples ordinária sobre uma ação (modelo de Black-Scholes-Merton): 293

6.2.2 Fórmula do preço racional da opção de compra européia simples ordinária sobre uma ação (modelo de entrópico):.....294

7 BIBLIOGRAFIA.....295

1 Introdução

"In the usual MBA curriculum, students are presented with a number of alternative and apparently competing methods for valuing risk projects. In their decision analysis course, students learn about decision trees and utility theory and are taught to think of values in terms of expected utility and certainty equivalents. In their finance course, they learn about the dicounted cash flow model and are taught to think of values in terms of net present values computed using a discount rate refleting the risk of the project. In an advanced finance course, they might learn about option pricing methods and be taught to think of projects as being analogous to put and call options on a stock. The result of all this training is a graduate who may understand each method but fail to appreciate the relationships between them and their relative strengths and weaknesses. A similar gap between the decision theory and finance disciplines exists in the academic literature and professional practice."

Smith & Nau (1995, p. 795)

Esta tese apresenta uma abordagem alternativa para a determinação do preço único e livre de arbitragem ou racional da opção européia simples ordinária sobre ação e *bond* em condições mais gerais que as tradicionalmente utilizadas no modelo de Black-Scholes-Merton.

O modelo de Black-Scholes-Merton basicamente é utilizado para determinar o preço racional da opção européia simples ordinária sobre ação.

Este método alternativo consiste na aplicação do princípio de máxima entropia associado ao sistema de distribuições de Pearson para a estimação da medida de probabilidade neutralizadora do preço da incerteza. Esta medida de probabilidade é utilizada na determinação do preço racional da opção européia simples ordinária sobre valores mobiliários e *bonds*.

Esta nova abordagem é, basicamente, uma fusão das abordagens apresentadas em Yang (1997), Gulko (1998 e 1998a), Cover & Thomas (1991), Reza (1996) e Golan *et al.* (1996).

1.1 Justificativa do tema e motivação

1.1.1 A importância e o preço da incerteza na Economia

Conforme Sassatani (1999), um dos aspectos fundamentais no estudo de Finanças e que revolucionou a maneira como esta ciência é estudada e compreendida no mundo moderno é a incerteza que está relacionada com o risco. Conforme explica Stone (1970, p. 1), "o papel da incerteza na tomada de decisões tem sido a preocupação predominante na teoria de Finanças contemporânea". Não seria nenhum exagero dizer que a teoria de Finanças moderna nasceu em virtude da preocupação dos acadêmicos de buscar uma estrutura conceitual e lógica que auxiliasse a tomada de decisões em condição de incerteza e de risco. Neste sentido, a teoria de portfólio de Markowitz e os trabalhos de Modigliani e Miller na década de 50, o CAPM de Sharpe, Treynor, Lintner e Fama na década 60, o APT de Ross e o modelo de determinação do preço racional de opções de Black, Scholes e Merton na década de 70 e as metodologias de avaliação de risco como valor em risco são consideradas contribuições importantes e expressivas para o mundo das Finanças.

Conforme Sassatani (1999), a incerteza poderia ser imaginada como um elemento que influencia a atitude dos agentes econômicos. Dessa forma, ela expressaria comportamentos como aversão, neutralidade ou atração pela incerteza e, quando se trata de analisar o comportamento dos agentes econômicos com aversão à incerteza na formação de preços dos ativos, surge a noção do preço da incerteza. Nas publicações da área de Economia, Finanças e Teoria Estatística da Decisão, o preço da incerteza é conhecido por prêmio do risco.

Luenberger (1998), *apud* Sassatani (1999), afirma que uma função de utilidade côncava representa um indivíduo com aversão à incerteza. O equivalente monetário determinístico (*emd*) ou equivalente certo é o preço mínimo pelo qual um decisor estaria disposto a vender uma situação com incerteza por ele possuída. Para uma pessoa com aversão à incerteza, o *emd* é menor que o valor monetário esperado. A diferença entre o *emd* e o valor esperado de uma variável aleatória associada a uma situação com incerteza é chamada **preço da incerteza**. O preço

da incerteza, para uma pessoa a ela avessa, assume um valor real estritamente positivo. Ele seria então o valor máximo a que um indivíduo com aversão à incerteza estaria disposto a renunciar a partir do valor monetário esperado para se desfazer da loteria possuída.

Conforme Luenberger (1998, p. 222 e 429), incerteza é oportunidade. Oportunidade tem valor positivo para o investidor. Como será visto mais adiante, mesmo no mundo com o preço da incerteza neutralizado há incerteza e, portanto, volatilidade ou entropia. A análise de investimento multiperíodo evidencia este fato. Logo, o investidor racional deve *apreciar* a incerteza!

1.1.2 O conceito geral de opção

"*The real options approach is a way of thinking.*"

Amra & Kulatilaka (1999, p. 5)

Bernstein (1997) dedica o capítulo 18 à história da opção financeira (derivativo). Segundo ele (p. 315) "os derivativos não são transações envolvendo ações ou taxas de juros, vidas humanas, casas vulneráveis a incêndios ou hipotecas imobiliárias. O *produto das transações com derivativos é a própria incerteza*".

O conceito de opção é relativamente conhecido e aplicado em Finanças. O desenvolvimento dos mercados financeiros, a sofisticação dos produtos, a difusão dos negócios em bolsas de valores e o interesse acadêmico na questão da avaliação tornaram as opções o centro do grande debate. Contudo, é no campo das ciências da decisão que o conceito assume seu aspecto mais genérico e, ao constituir um significado mais abrangente, consegue incluir as definições utilizadas em Finanças e Economia.

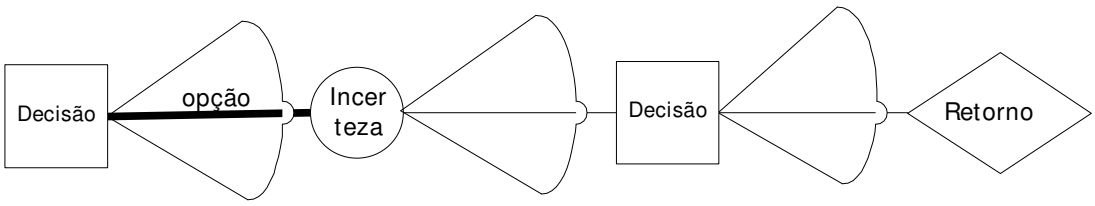
Além disso, conforme Nichols (1994, p. 91), para Judy Lewent, CFO da Merck & Co., todos os tipos de decisões de negócios são opções.

O termo **opção** está diretamente ligado à noção de alternativa, embora este não seja seu significado mais preciso.

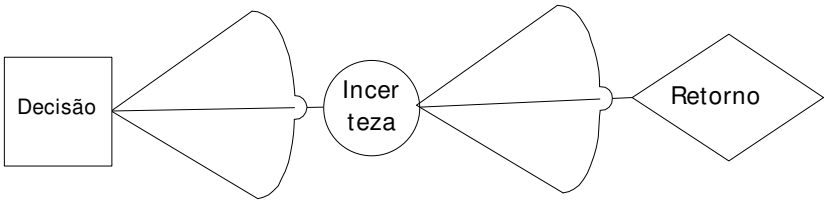
Howard (1996), *apud* Malinak (1996, p. 11) define opção como "uma alternativa que permite uma decisão futura provocada pela revelação de informação".

Note-se que a opção somente faz sentido em virtude da incerteza ou da falta de informação dos indivíduos. Se os fenômenos analisados pelas ciências da decisão ou Finanças fossem determinísticos e seus agentes, racionais, as preferências dos indivíduos convergiriam para a alternativa que trouxesse maior satisfação, lucro ou prazer. Nessa condição, qualquer discussão sobre opção seria redundante.

A árvore de decisão que auxilia no entendimento dessas relações é a seguinte:



Pode-se observar que esta estrutura decisória difere da que representa o problema de seleção ótima de carteira:

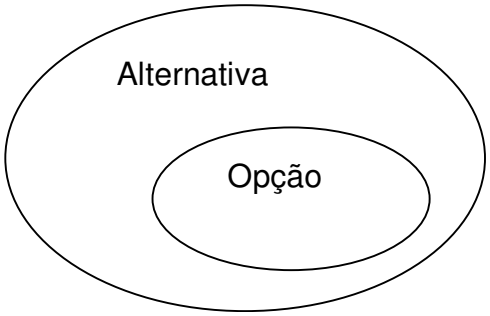


Torna-se, portanto, natural pensar em utilizar o conceito de valor da informação perfeita (valor de clarividência) e da imperfeita e o custo de implementação da decisão para que seja estabelecido o valor máximo da decisão de ter o direito de decidir depois da resolução da incerteza (Howard (1996)).

Howard (1988), *apud* Malinak (1996, p. 2), define análise da decisão como "um procedimento sistemático para transformar problemas decisórios opacos em problemas decisórios transparentes por meio de uma seqüência de passos transparentes. Análise da decisão é uma disciplina normativa criada para ajudar as pessoas a tomarem melhores decisões". Para Howard (1996), *apud* Malinak (1996), os elementos básicos da decisão são: alternativas, informação e preferência.

Assim, a própria decisão de fazer a análise da decisão é uma opção. O problema é de natureza recorrente (recursiva).

Temos ainda que toda opção é uma alternativa, mas nem toda alternativa é uma opção.



Outra idéia importante relacionada à opção é o livre arbítrio.

Em Finanças, a opção pode ser definida como um direito. Conforme explica Hull (1997, p. 5), a opção é algo que confere ao seu titular o direito de fazer algo e as pessoas exercerão esse direito se as circunstâncias forem vantajosas. Howard (1996, p.81) ainda ressalta que a opção está associada ao exercício da vontade humana em uma decisão futura.

O aspecto mais interessante sobre a opção é que constitui um elemento que agrega valor à decisão dos indivíduos. Em uma decisão caracterizada pela incerteza, qualquer informação que possa conduzir o sujeito a tomar uma decisão (ação) diferente possui valor [Clemen (1996, p. 438)].

Sobre esse aspecto, Howard (1996, p. 82) expõe o conceito de clarividência ou capacidade de antecipar todas as incertezas enfrentadas por um decisor. A existência de informações pode diminuir as incertezas, alterar o curso de uma ação e trazer mais satisfação ao indivíduo. Portanto, é interessante saber identificar situações em que estão presentes as opções. Contudo, a experiência de Howard (1996, p. 82) indica que as capacidades de reconhecer e de criar opções nos problemas de decisão são freqüentemente negligenciadas nas etapas de identificação (*framing*) e estruturação.

De acordo com Howard (1996, p. 83), a discussão sobre opções, na análise da decisão, se aplica a indivíduos que seguem algumas regras baseadas na idéia de von Neumann-Morgenstern. Bierman Jr. *et al.* (1991), Pratt *et al.* (1996) e Raiffa (1977) expõem detalhadamente essas regras.

Segundo Howard (1996, p. 84), o valor de uma opção depende da atitude, em relação à incerteza, de seu titular ou lançador.

A propensão das pessoas à incerteza pode ser observada pela curvatura da função de utilidade. Para indivíduos com aversão à incerteza (preço da incerteza positivo) pode-se atribuir uma função de utilidade exponencial ou logarítmica.

A partir da função de utilidade de cada indivíduo pode-se obter o equivalente monetário determinístico (*emd*) de uma dada alternativa de decisão que envolva incerteza.

A suposição de o decisor ser neutro à incerteza (decisor que considera apenas a expectativa de retorno, portanto, para ele o preço da incerteza é nulo) ou ter uma função utilidade exponencial (decisor avesso à incerteza - que considera um preço da incerteza positivo - e que não considera o valor de sua riqueza inicial) simplifica bastante os cálculos do equivalente monetário determinístico (*emd*), pois funciona a propriedade chamada **delta**:

$$emd(X + \delta) = emd(X) + \delta\,.$$

Isto é, o equivalente monetário determinístico de uma loteria X (situação com incerteza) mais uma constante delta é igual ao equivalente monetário determinístico da loteria mais o valor da constante delta. Em geral, a preocupação está em determinar um delta igual ao preço da opção. Além disso, decisores do tipo delta gozam da propriedade do preço de compra da loteria ser igual ao preço de venda da mesma loteria. Portanto, para eles o preço, num sentido pessoal, é único. Para os decisores com função utilidade exponencial, o preço da loteria varia em função do coeficiente de tolerância ao risco ρ, que, mantida a situação decisória, é constante para cada decisor, mas pode variar de decisor para decisor. Apenas os decisores neutros ao risco (função de utilidade linear) ou avessos ao risco com função de utilidade exponencial são do tipo delta. Para um decisor indiferente à incerteza o $emd(X)$ é o valor esperado da loteria. Por exemplo, para o decisor exponencial o $emd(X)$ é:

$$\mathrm{X} \sim \mathrm{normal}\left(\mu;\sigma^2\right) \Rightarrow emd(X) = \mu - \frac{\sigma^2}{2\rho}$$

Se, por exemplo, o decisor for avesso à incerteza com função utilidade logarítmica, ele não será do tipo delta. Além disso, a utilidade depende da riqueza inicial do decisor (α), o preço de compra da loteria é menor que o preço de venda da mesma loteria e o coeficiente de aversão à incerteza varia em função dos valores da loteria e da riqueza inicial, isto é, não é constante [Bekman & Costa Neto (1993, p. 79-80)]. Quando o valor da riqueza inicial é muito maior que os valores em jogo na loteria, a função utilidade do decisor é aproximadamente exponencial.

O lançamento de uma opção pode implicar o pagamento de um valor denominado prêmio ou preço da opção. Trata-se do custo de aquisição dessa opção, que confere ao seu titular um determinado direito. O preço dessa opção pode ser avaliado como o preço que torna o decisor indiferente entre sua aquisição ou não. Em geral, o prêmio costuma crescer de acordo com a incerteza e com o número de alternativas. Além disso, conforme Howard (1996, p. 84), a escolha de uma alternativa oferecida pela opção pode exigir o pagamento de um valor adicional denominado **preço de exercício**.

Embora os livros costumem associar as opções a uma modalidade de operação financeira ou a uma questão na análise da decisão, as pessoas costumam comprar ou vender opções em diversas situações do cotidiano.

Quadro 1: Exemplos de opções

Exemplo	Prêmio	Preço de exercício
Compra de um bilhete de loteria	preço do bilhete	imposto
Seguro de carro contra acidentes	prêmio do seguro	franquia
Plano de saúde	preço ou mensalidade	tabela da AMB

Siqueira & Sassatani (1999) mostram como formalizar matematicamente o conceito de opção, utilizando a Teoria Estatística Bayesiana da Decisão.

Kimura *et al.* (1999) mostram que há quatro produtos tradicionais no mercado financeiro - empréstimo com garantia, seguros, debêntures e cheque especial - que, em alguns aspectos, podem ser analisados como contratos de opções.

Herath (1998) mostra que projetos de investimento com incerteza podem também ser analisados como opções, utilizando a Teoria Estatística Bayesiana da Decisão.

1.1.3 Mundo com o preço da incerteza neutralizado e a medida de probabilidade neutralizadora do preço da incerteza

No entanto, seria útil neutralizar o preço da incerteza para que fosse necessária a consideração de funções de utilidade do decisor na determinação do preço da opção. Isto é possível pela introdução de um ente matemático chamado **medida de probabilidade neutralizadora do preço da incerteza** (*risk-neutral probability*), que permite a criação de um mundo virtual chamado **mundo com o preço da incerteza neutralizado** (*risk-neutral world*).

Essa medida de probabilidade elimina a necessidade de considerar o preço da incerteza de cada valor mobiliário. Dessa forma, todos os decisores podem tratar os ativos com preço da incerteza como se fossem ativos livres de preço da incerteza. Logo, a taxa de desconto desse mundo livre de preço da incerteza passa a ser a taxa de juros certa. Outras vantagens desse mundo virtual são (i) a ausência de arbitragem, isto é, não existe a possibilidade de ganho certo sem investimento, (ii) a linearidade dos preços, isto é, o preço de uma combinação linear de ativos é igual à combinação linear dos preços de cada ativo e (iii) unicidade dos preços, isto é, o preço de um ativo é único. Além disso, como não há preço da incerteza nesse mundo sintético, a esperança matemática é o operador linear adequado para a determinação de preço racional.

A medida de probabilidade neutralizadora do preço da incerteza é um elemento fundamental para a determinação do preço da opção financeira e é uma entidade não observável que não depende da distribuição do preço do valor mobiliário subjacente.

Conforme Hull (1998, p. 260) a determinação do preço de derivativos pode ser feita nesse mundo virtual, pois "qualquer valor mobiliário sobre outros valores mobiliários negociados pode ser avaliado sob a hipótese de que os decisores são indiferentes à incerteza" Ele observa ainda que a determinação de preço racional no mundo com preço da incerteza neutralizado não exige que os investidores

sejam neutros à incerteza. O que ele estabelece é que os derivativos, tais como opções, podem ter seus preços racionais determinados sob a hipótese de que os investidores são neutros à incerteza. Isso significa que as preferências dos investidores não têm efeito sobre o valor da opção quando ela é expressa como uma função do preço de um valor mobiliário subjacente, como a ação. A determinação de preço racional no mundo com preço da incerteza neutralizado é um poderoso instrumento porque, nesse mundo sintético, funcionam dois resultados particularmente simples: (i) a taxa de retorno esperada de todos os valores mobiliários é a taxa de juros certa do valor mobiliário com fluxo de caixa certo (ii) a taxa certa é a taxa de desconto apropriada para ser aplicada a qualquer fluxo de caixa futuro. Opções e derivativos que produzem um retorno num tempo específico podem ter preço determinado nesse mundo remido do preço da incerteza. Este procedimento segue as etapas para a opção européia: (i) assumir que a taxa de retorno esperada do valor mobiliário subjacente é a taxa certa, (ii) calcular o retorno esperado da opção na sua data de exercício e (iii) calcular o valor presente usando a taxa certa do valor esperado do retorno da opção na data de exercício.

Conforme Yang (1997, p. 7), a estimação e especificação da medida de probabilidade neutralizadora do preço da incerteza tornou-se um tema central em Finanças. Ainda, de acordo com essa autora, o princípio de não-arbitragem fornece uma condição suficiente para a existência de uma medida de probabilidade neutralizadora do preço da incerteza necessário para a formação de preço livre de arbitragem. Entretanto, esse princípio não fornece um meio de determinar essa medida de probabilidade.

Ross (1976), *apud* Yang (1997, p. 5), considerou o conceito de arbitragem como o conceito unificador em Finanças.

1.1.4 Teoria da Informação: uma abordagem alternativa para quantificar a incerteza e determinar preço

A tese expressa o problema da determinação do preço racional de opção européia simples ordinária por meio da Teoria da Informação de Shannon.

A Teoria da Informação de Shannon é também conhecida como Teoria da Informação. Ela quantifica o conceito de surpresa ou incerteza de uma maneira que altera o método de determinação do preço racional da opção.

A Teoria da Informação permite abordar os problemas de seleção de carteira log-ótima [Cover & Thomas (1991), Kelly Jr. (1956) e Vince (1999)] e a determinação de preço racional de opção [Gulko (1998 e 1998a) e Yang (1997)].

A Teoria da Informação, além disso, fornece uma maneira unificada de tratar problemas em condição de incerteza. Os conceitos de entropia, informação mútua, entropia diferencial, divergência, princípio da máxima entropia, princípio da mínima divergência e taxa de entropia fazem parte desta nova linguagem.

A Teoria da Informação, conforme E. L. Iñigo em *A filosofia, hoje*, *apud* Megale (1992, p. 72), descende do Pensamento Técnico e da Cibernética. A Cibernética, por sua vez, descende do Positivismo, Pragmantismo e Epistemologia. Finalmente, o Positivismo descende do Empirismo.

1.2 Formulação da questão de pesquisa

Como uma única medida de probabilidade neutralizadora do preço da incerteza, necessária para a determinação do preço racional da opção européia ordinária sobre valores mobiliários, pode ser extraída por meio do princípio da máxima entropia apenas sob as hipóteses de mercado perfeito e de mercado em equilíbrio competitivo?

1.3 Objetivo da tese

O objetivo da tese é desenvolver um método alternativo, robusto e parcimonioso para a determinação do preço racional da opção européia simples ordinária sobre ação e taxa de juros (*bond*) por meio da abordagem entrópica que aprimora a tecnologia de estimação da medida de probabilidade neutralizadora do preço da incerteza.

1.4 Metodologia

"*Statistical thinking is defined as a philosophy of learning and action, based on three fundamental principles: All works occurs in a system of interconnected processes, variation exists in all processes, and understanding and reducing variation are keys to success (Quality Press 1996). We must have a keen and deep understanding of the role of variation in all technical and managerial processes that drive the organization. Understanding variation and its effects is the core competency of statisticians.*"

Snee (1998, p. 208)

Conforme Efron & Tibshirani (1993, p. 392), o estatístico trabalha no campo do conhecimento no qual interagem a ciência, a matemática e a filosofia.

Metodologicamente a tese apóia-se em alguns princípios, métodos e abordagens.

Do ponto de vista filosófico, deseja-se obter um modelo parcimonioso em termos de hipóteses para a determinação do preço de opção européia. Dessa forma, será seguido o princípio da parcimônia:

- *Princípio da parcimônia (navalha de Occam)*: conforme William of Occam (1285-1349), *apud* Cover & Thomas (1991, p. 4), "as causas não devem ser aumentadas em número além do necessário", isto é, a mais simples explicação é a melhor. Ao relacionar Finanças com entropia, Samperi (1998, p. 4) utiliza a navalha de Occam. Merton (1973, p. 161) também sugere o uso deste princípio para a determinação do preço racional da opção. Segundo Poincaré (1984, p. 158), para um espírito acostumado a admirar os modelos da física, dificilmente uma teoria é satisfatória e ele não somente não tolerará a menor aparência de contradição, como exigirá que as diferentes partes da teoria estejam logicamente interligadas e que o número de hipóteses distintas seja o menor possível.

Como a tese é quantitativa e ligada a um problema econômico, segue uma abordagem econométrica:

- *Abordagem econométrica*[1]: conforme Frish[2] (1933), *apud* Greene (1997, p. 1), na primeira edição da *Econometrica* a Sociedade Econométrica estabeleceu

---

[1] Contra a matematização das Ciências Sociais temos as palavras de Plauto Faraco de Azevedo (1975, p. xii-xiii) na Introdução do livro - que ele próprio traduziu - intitulado *O caso dos exploradores de cavernas* de Lon L. Fuller publicado em 1993 em Porto Alegre pela Fabris: "Perelmam combate a opinião de tantos filósofos que consideram - e continuam considerando - que toda forma de raciocício que não se assemelhe ao matemático não pertence à lógica. Contra esta opinião injustificada e caduca sustenta Perelman que há mesmo *formas de raciocício mais elevadas*, que não constituem propriamente cálculos nem tampouco podem ser formuladas como '*demonstrações*', pertencendo, em contrapartida, à *argumentação*". E é este "*precisamente o tipo de raciocício empregado pelo jurista*... A tradição cartesiana, que busca acima de tudo a *evidência*, desdenha qualquer proposição que não possua o caráter de óbvio, do indiscutível, do exato, do preciso. Todavia, esta concepção logicista ou matematizante do pensamento é demasiada estreita, pois não abrange grande quantidade de raciocínios, que não têm e nem podem ter forma demonstrativa... Mas sucede que a própria índole da deliberação e da argumentação se opõem à evidência e à necessidade absoluta; porque não se delibera nos casos em que solução têm caráter de necessidade, como não se argumenta contra a evidência. A argumentação tem seu sentido no verossível, no plausível e no provável, escapando estes à certeza de um cálculo exato de que resulte uma única solução justuficável em termos absolutos... Já os cultores das Ciências Naturais apenas reconhecem a evidência da intuição sensível, da experiência e da indução... Tanto a concepção cartesiana quanto a dos cientistas empíricos mutilam o campo da razão, posto que lhe negam a capacidade para tratar dos domínios em que nem a dedução lógica nem a observação dos fatos podem fornecer-nos a solução dos problemas. A aceitar-se esta cinscunscrição da razão em tais domínios, não nos restaria outro recurso exceto o de neles entregar-nos às forças irracionais, a nossos instintos ou à violência". Perelman, Chaim - *De la justicia* (*De la justice*) Trad.

que seu objeto principal será promover estudos que contribuam para a unificação das abordagens teórico-quantitativa e empírico-quantitativa de problemas econômicos que estejam permeados por um raciocínio construtivo e rigoroso, semelhante àquele que predomina nas ciências naturais. Mas há várias facetas da abordagem quantitativa em Economia, e nenhuma delas isoladamente deve ser confundida com econometria. Então, a econometria não é o mesmo que a Estatística Econômica. Nem é idêntica à Teoria Geral Econômica, embora uma parte considerável desta teoria possua um caráter realmente quantitativo. Tampouco deve a Econometria ser tomada como sinônimo de aplicação de Matemática em Economia. A experiência tem mostrado que cada um destes três pontos de vista – Estatística, Teoria Econômica e Matemática – é uma condição necessária, mas não suficiente para o real entendimento das relações quantitativas na vida econômica moderna. É na unificação dessas três abordagens que reside o seu poder e é essa unificação que constitui a Econometria. Conforme Ebrahimi *et al.* (1999, p. 317), a principal atividade em Econometria é a estimação de funções de distribuição de probabilidade e de processos estocásticos baseada em informação incompleta (e.g. estimação, teste de hipótese, seleção de modelo, análise de portfólio etc.). O uso da teoria da probabilidade, conforme Theil (1967, p. vi-vii), é atualmente marcante. Esse autor destaca três estágios da Teoria da Probabilidade em Economia: (i) a Teoria da Probabilidade é básica para a Estatística; a Estatística é útil para a Economia (análise de regressão, amostragem etc.); desse modo, a Teoria da Probabilidade afeta *indiretamente* a Economia *Aplicada* por meio dos métodos estatísticos utilizados; (ii) mais recentemente (até 1966), tem-se visto o desenvolvimento da teoria da tomada da decisão em condição de incerteza. A maneira mais simples de tratar a incerteza é utilizar os conceitos da teoria da probabilidade; essa é a maneira

---

De Ricardo Guerra. México: Universidade Nacional Autónoma de México, 1964, p. ii-iii.

[2] Frish, R. 1933. "Editorial." *Econometrica* 1:1-4.

como a teoria da probabilidade entra *diretamente* na teoria econômica e (iii) Theil (1967, p. vi) pensa que há uma terceira e muito mais ampla área na qual a teoria da probabilidade pode ter um papel útil. Como ele parte da definição de Economia como sendo a disciplina que está preocupada com a alocação de recursos escassos para finalidades alternativas, considera que a Teoria da Probabilidade, no sentido usual da palavra, pode ser considerada como um problema de alocação especial, isto é, dados n eventos, há n probabilidades parciais que necessitam ser alocadas a eles, baseando-se na confiança de duas realizações. Segundo Theil (1967, p. vi), a Teoria da Informação é a abordagem adequada para atingir esse fim[3].

Essa tese pretende também aprimorar a tecnologia de estimação da medida de probabilidade neutralizadora do preço da incerteza. Dessa forma, adota um método de pesquisa denominado **método de desenvolvimento**:

- *Método de pesquisa de desenvolvimento*: esse método é definido por Contandriopoulos *et al.* (1994, p. 41) como uma "estratégia de pesquisa que visa, utilizando-se de maneira sistemática os conhecimentos existentes, elaborar uma nova intervenção ou melhorar consideravelmente uma intervenção existente ou, ainda, elaborar ou melhorar um instrumento, um dispositivo ou um método de medição".

Conforme Lauand (1987, p. 78), "...acadêmico significa filosófico e filosófico significa essencialmente (entre outras coisas) teorético e, portanto, algo voltado unicamente para a captação da realidade e alheio a fins práticos...". Portanto, a tese pretende ser acadêmica, i.e., possuir uma abordagem acadêmica, portanto, teórica:

---

[3] É interessante notar que Myron Scholes participou da elaboração desse livro, merecendo os seguintes agradecimentos: "[...]*Myron Scholes, who programmed the computations; Myron Scholes again, who read the manuscript and advised me on its intelligibility;*[...]".

- *Abordagem teórica*: parafraseando Vince (1999, p. XII), esta tese é sobre teoria musical. Não é uma tese que ensina, passo a passo, como se toca um instrumento. Da mesma forma, não é uma tese sobre como vencer o mercado. Mais precisamente, é uma tese sobre conceitos matemáticos, que dá aquele importante passo entre teoria e aplicação. Ela não irá conferir habilidade para lidar com a dor emocional que a atuação no mercado de capitais reserva, quer se ganhe se perca. Conforme Sharpe & Alexander (1990, p. 195-6), a determinação do preço racional de um ativo necessita da construção de um modelo, isto é, da construção de uma teoria. Isso exige simplificação das situações complexas por um processo de abstração dos seus elementos essenciais. Essa simplificação é atingida por meio de hipóteses sobre o ambiente. Essas hipóteses necessitam ser simples de modo que forneçam um grau de abstração que permita o sucesso na construção do modelo. Por outro lado, o teste de um modelo está em sua capacidade de auxiliar no entendimento e na predição do processo que está sendo modelado.

O exemplo canônico ou paradigmático permite a introdução integrada de conceitos com o mínimo de formalização matemática:

- *Exemplo canônico ou paradigmático:* segundo Howard (1996), o exemplo paradigmático possui as seguintes características:
  - ❖ É um problema hipotético.
  - ❖ É o problema decisório não trivial mais simples possível.
  - ❖ Apresenta questões desafiadoras.
  - ❖ Possui a estrutura de problemas reais.

1.5 Contribuições originais e resultados principais da tese

Conforme Salomon (1972, p. 206-13), **originalidade** significa, pela própria etimologia da palavra, **volta às fontes** (**origem**, que quer dizer princípio, **arqué**). Não se identifica, portanto, com novidade ou singularidade, mas retorno à origem, à essência, à verdade, ainda que essa verdade se tenha perdido, obscurecido ou olvidado. Hoje se verifica o esquecimento do sentido prístino,

devido talvez às idéias de progresso e moda trazidas no bojo do desenvolvimento científico e tecnológico. Mas exigir originalidade como total novidade num trabalho, para que ele seja tese ou monografia, é uma colocação ingênua, para não dizer inatingível: (i) porque a ciência é um processo cumulativo em que verdades provisórias se assentam com maior freqüência, e a **revisão** é uma constante; (ii) porque o que está em jogo, desde que pela primeira vez se formulou a exigência de originalidade para teses e monografias, é a questão da atualização de que se falou quando se tratou da relevância contemporânea na escolha de problema que mereça ser investigado cientificamente; (iii) porque, desde tempos remotos, considerou-se original também o trabalho que apresenta modo novo de abordar um assunto já tratado ou que consegue estabelecer relações novas, ou, finalmente, que se propõe uma nova interpretação de questões controversas. Ademais, nem todas as ciências são como as naturais, que, através de experimentos, conseguem realmente descobertas e inventos.

De acordo com Castro[4] (1978, p. 55), *apud* Sassatani (1999), uma tese deve ser original, importante e viável. Um tema é importante quando polariza ou afeta um segmento substancial da sociedade. O fato de não haver sido feito não confere necessariamente originalidade. Um tema original é aquele que tem o potencial de nos surpreender.

Essa tese apresenta algumas características inovadoras.

Ela promove uma integração entre Finanças e Teoria da Informação, capaz de criar uma tecnologia alternativa para a determinação do preço racional da opção européia simples ordinária sobre ação e *bond*. Uma das características desta nova tecnologia de determinação de preço racional é poder continuar utilizando o cálculo newtoniano em vez do estocástico. Cria uma notação matemática precisa e completa para a Teoria da Informação e a integra com a teoria de Finanças em condições de incerteza. Integra as abordagens entrópicas da determinação de preço racional da opção européia simples ordinária de Gulko (1998 e 1998a) e de Yang (1997).

---

[4] Castro, C. de M. e. 1977. *A prática da pesquisa.* 1ª ed. São Paulo: McGraw-Hill.

Define precisamente o mundo com preço da incerteza neutralizado (*risk-neutral world*), o mundo martingale, o mundo informacionalmente eficiente e o mundo entrópico e suas implicações para a Ciência do Investimento e para a determinação de preço racional de ativos básicos e derivativos.

Demonstra detalhadamente a fórmula do preço racional da opção européia simples ordinária de Black-Scholes-Merton, melhorando a notação matemática, simplificando (eliminando a abordagem de martingale) e complementando a demonstração feita por Baxter & Rennie (1998).

Interrompe uma sucessão de trabalhos que estabelecem uma forma equivocada de calcular o preço de opção européia simples ordinária. Esse erro teve sua origem, muito provavelmente, numa edição de Brealey & Myers, que equivocadamente utilizou um resultado de Cox & Rubinstein (1985); esse resultado facilitava o cálculo do preço racional da opção européia simples ordinária por meio de uma tabela que evita o uso direto da fórmula de Black-Scholes-Merton. Brealey & Myers (desde a quarta edição de 1991), Luehrman (nos seus dois artigos da HBR de 1998 e um caso de 1995 pela HBS) e Edleson (caso publicado em 1994 pela HBS) ensinam que o valor percentual encontrado nessa tabela deve ser multiplicado pelo preço do valor mobiliário, quando deveria ser multiplicado pelo valor presente do preço de exercício.

Os resultados mais importantes desta tese para Finanças são: (i) desenvolvimento de um método alternativo, robusto e parcimonioso, baseado no princípio da máxima entropia da Teoria da Informação e do Sistema de Distribuições de Pearson para obtenção de uma única medida de probabilidade neutralizadora do preço da incerteza (*risk-neutral probability*), (ii) obtenção de fórmula prática para a determinação do preço racional da opção européia simples ordinária para ação, (iii) validação da fórmula de Black-Scholes-Merton para ação, (iv) obtenção de uma fórmula adequada para a determinação do preço racional da opção européia simples ordinária sobre um *bond*, (v) estimação da volatilidade implícita entrópica do preço do valor mobiliário e (iv) definição e estimação do valor em risco entrópico.

Há ainda dois resultados importantes para a Teoria da Informação e Economia: (i) distinção mais precisa entre incerteza e risco e (ii) desenvolvimento da medida de ganho informacional da previsão aprimorando o resultado de Theil (1967) e Benish (1999) pela utilização do conceito de divergência de Kullback-Leibler.

1.6 Escopo da tese

A tese analisa um problema das Ciências Humanas - Economia Aplicada - com a metodologia e linguagem das Ciências Naturais - Física - e Exatas - Matemática, Estatística e Computação.

A seguir são sintetizadas as áreas de conhecimento da qual derivam a de determinação de preço racional de opção:

- **Ciências Humanas**: universo de conhecimento cujo objeto comum de investigação é o próprio ser humano enquanto animal social [Megale (1992)];
- **Ciências Sociais**: investigam a relação social entre os homens, ou a ação social [Megale (1992)];
- **Economia**: disciplina que está preocupada com a alocação de recursos escassos para finalidades alternativas [Theil (1967)];
- **Econometria**: a principal é a estimação de funções de distribuição de probabilidade e de processos estocásticos baseada em informação incompleta [Ebrahimi *et al.* (1999)]; análise e resolução de problemas econômicos por meio da metodologia das Ciências Naturais com a linguagem das Ciências Exatas [Frish (1933), *apud* Greene (1997)];
- **Administração**: é um *processo* de atingir certos objetivos utilizando recursos tais como pessoas, dinheiro, energia, materiais, espaço e tempo; Administração é sinônimo de Tomada de Decisão [Turban & Meredith (1994)];
- **Ciência da Administração ou Tecnologia da Decisão**: é a aplicação de do método científico para a análise e resolução de problemas decisórios em Administração [Turban & Meredith (1994)];

- **Ciência da Administração Financeira ou Finanças**: está estreitamente ligada à Economia e Contabilidade e pode ser vista como uma forma de Economia Aplicada [Gitman (1987)]; é a ciência e a profissão do manejo do dinheiro, particularmente do dinheiro do Estado [Ferreira (1986)];
- **Ciência do Investimento**: é a aplicação de instrumentos científicos, sendo estes eminentemente matemáticos, para o emprego de recursos disponíveis com o intuito de obter benefícios futuros investimentos, i.e., investimentos; a linguagem dessa ciência é em grande parte matemática e alguns aspectos do assunto podem ser expresso apenas em termos matemáticos. [Luenberger (1998)];
- **Análise de investimento**: é o processo de examinar alternativas e decidir qual delas é a preferível dentro da estrutura de um mercado financeiro, sendo que este mercado provê alternativas não encontradas em outras situações. Essa estrutura é o que faz a análise de investimento ser única e excepcionalmente poderosa [Luenberger (1998)];
- **Análise de investimento em condição de incerteza**: a incerteza surge quando pode haver mais que um possível fluxo de caixa que pode ser produzido por um dado ativo num dado tempo futuro [Bierman Jr. & Smidt (1993)];
- **Valor mobiliário em condição de incerteza**: ação, taxa de juros (*bond)*, taxa de câmbio, derivativo etc.
- **Derivativo**: ou derivado ou instrumento financeiro derivado é um contrato financeiro ou direito contingencial se seu valor, na data de expiração, é determinado exatamente pelo preço de mercado do valor mobiliário subjacente nessa data. [Ingersoll Jr. (1987), *apud* Neftci (1996)]; há três grupos de derivativos: (i) contrato futuro e a termo, (ii) opção e (iii) *swap*. Teoricamente o contrato a termo e a opção podem gerar qualquer outro valor mobiliário, inclusive o futuro e o *swap*. [Neftci (1996)];
- **Opção**: é um contrato que dá o direito, mas não a obrigação de fazer algo numa data futura: direito de comprar ou vender um bem, concedido mediante pagamento de uma quantia acordada entre as partes. Se o direito não for

exercido depois de um período especificado, a opção termina pelo vencimento e o comprador da opção perde a quantia paga para obtenção da opção [Downes & Goodman (1993)];

- **Opção européia ordinária simples de compra**: é uma opção cujo valor intrínseco é o máximo entre o valor zero e a diferença entre o preço do valor mobiliário subjacente $S(t)$ e o preço de exercício $k(t)$ na data de exercício t, i. e., $C(t)=\max[0;S(t)-k(t)]$.

1.7 Estrutura da tese

No capítulo *Fundamentação teórica* são definidos os conceitos básicos da Ciência do Investimento, os princípios a que o decisor racional deve obedecer, as hipóteses necessárias para o desenvolvimento do modelo de Black-Scholes-Merton (B-S-M) e do modelo entrópico. Em seguida o modelo de B-S-M é explicado detalhadamente. Como introdução ao modelo entrópico, há a apresentação dos elementos da Teoria da Informação.

No capítulo seguinte há a aplicação da Teoria da Informação por meio do princípio de máxima entropia e do Sistema de Distribuições de Pearson para a estimação da medida de probabilidade neutralizadora do preço da incerteza entrópica. Em seguida, são determinadas as fórmulas entrópicas de determinação do preço da opção européia simples ordinária sobre ação e obrigação (*bond*).

No capítulo final são tratados os problemas de implementação computacional das fórmulas, os modelos de B-S-M e entrópico são comparados, as limitações do modelo entrópico são apresentadas e há recomendações para pesquisas futuras.

1.8 Notação matemática utilizada na tese e termos técnicos

Bons glossários sobre os termos de cálculo estocástico encontram-se em Shreve (1997, p. 349-53) e Baxter & Rennie (1998).

Um bom glossário sobre os termos da teoria da decisão encontra-se em Spradlin (1997).

Um bom glossário da Teoria da Informação pode ser encontrado em Mackay (1995).

O glossário e o uso dos termos técnicos em Finanças tem por base os glossários de Brealey & Myers (1996) e (1998), Pindyck & Rubinfeld (1994), Baxter & Rennie (1998), Bernstein (1997), Downes & Goodman (1993) e Gélédan & Brémond (1988).

Na maioria das vezes procurar-se-á seguir a notação matemática destacada abaixo:

X: (letra maiúscula) variável estocástica ou aleatória (matricial ou escalar)

x: (letra minúscula) constante ou parâmetro (variável determinística) podendo ser um escalar ou matriz

t: tempo (período da taxa de juros)

$x(t) = x_t$

$\{X_i : i = 1,2,...\}$: processo estocástico discreto

$\{X_i : i \geq 0\}$: processo estocástico contínuo

$\mathbb{P}$: medida de probabilidade subjetiva, real ou prévia

$\mathbb{Q}$: medida de probabilidade neutralizadora do preço da incerteza (sintética ou martingale)

$E_{\mathbb{Z}}(X) = \langle X \rangle_{\mathbb{Z}}$: esperança matemática da variável aleatória X na medida de probabilidade $\mathbb{Z}$. Conforme Jaynes (1996, p. 3-14), os estatísticos e matemáticos costumam utilizar a primeira notação, enquanto os físicos utilizam a segunda.

$Var_{\mathbb{Z}}(X) = E^2_{\mathbb{Z}}(X) - E_{\mathbb{Z}}(X^2)$: variância da variável aleatória X na medida de probabilidade $\mathbb{Z}$

$\exp(x) = \sum_{i=0}^{\infty} \frac{1}{i!} = e \cong 2,7182$

$\ln(x) = \log_e(x)$

r_f : taxa de juros certa (livre de risco) nominal ou efetiva ao período

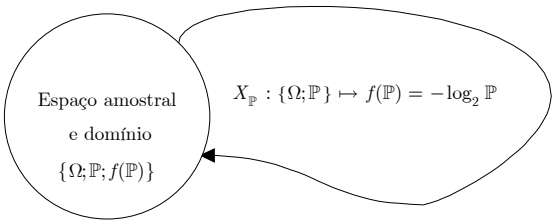
♦: fim de exemplo canônico ou demonstração

x^T : matriz transposta

x^{-1} : matriz inversa

$s(p)$: informacional ou quantidade de surpresa, informação ou incerteza do evento com probabilidade p (bit)

$X_{\mathbb{P}} = -\log_2 \mathbb{P}$: auto-variável aleatória ou variável aleatória informacional



$h_{\mathbb{P}}(X_{\mathbb{T}})$: entropia ou quantidade esperada de variável aleatória informacional $X_{\mathbb{T}} = -\log_2 \mathbb{T}$ na medida de probabilidade $\mathbb{P}$ (bit/evento)

$d(X_{\mathbb{T}} \,||\, X_{\mathbb{P}})$: divergência de Kullback-Leibler, *cross-entropy* ou entropia relativa (bit/evento)

$\gamma = \sum_{i=1}^{n} d(X_{\mathbb{T}^i} \,||\, X_{\mathbb{P}})$: divergência total (bit/evento)

$\gamma^i = \dfrac{d(X_{\mathbb{T}^i} \,||\, X_{\mathbb{P}})}{\sum_{i=1}^{n} d(X_{\mathbb{T}^i} \,||\, X_{\mathbb{P}})}$ divergência relativa

para $i = 1, 2, \ldots, n$

2 Fundamentação téorica

Nihil lucri cepit qui nulla pericla subivit.
Nothing venture, nothing have.
Quem não arrisca, não petisca[5].
Mota (1987, p. 190-1)

Qui nimis inquirit, multa pericla subit.
Quem as coisa muito apura, não vive vida segura.
Mota (1987, p. 177-8)

2.1 A Ciência do Investimento

"Portanto, devemos afirmar que uma empresa de negócios já constituída tem três objetivos essenciais:
1) prestar um serviço real à sociedade civil,
2) aperfeiçoar seu próprio corpo social e conseguir os recursos instrumentais - técnicas e bens materiais - a fim de melhor realizar suas operações, e
3) obter um lucro econômico."

Figuerola, J. "A estratégia da empresa de negócios: importância, definição e conteúdo" *Instituto Superior da Empresa* - Barcelona.

Conforme Luenberger (1998, p. 1), investimento é o emprego de recursos disponíveis com o intuito de obter benefícios futuros. A Ciência do Investimento é a aplicação de instrumentos científicos para investimentos. Esses instrumentos são fundamentalmente matemáticos. Na página xiii o autor afirma que a linguagem dessa ciência é em grande parte matemática e alguns aspectos do assunto podem ser expresso apenas em termos matemáticos. O surgimento e desenvolvimento desta ciência deve-se basicamente à expansão global da atividade de investimento, pelo explosivo crescimento da tecnologia da

[5] **Petisco** é, conforme Ferreira (1986, p. 1322), uma iguaria saborosa, preparada com esmero. **Iguaria**, por sua vez, conforme Ferreira (1986, p. 916), é uma comida fina, delicada e/ou apetitosa. Logo, **petiscar** é saborear uma iguaria.

computação e da informação e por alguns extraordinários desenvolvimentos teóricos em Finanças.

Conforme Moore (1986), investimento é o uso de capital para gerar mais recursos, seja por meio de negócios geradores de renda ou de empreendimentos mais arriscados destinados a ganhos de capital. Pode se referir a um investimento financeiro (quando um investidor põe dinheiro em determinado negócio) ou a um investimento de esforço e tempo por parte de um indivíduo que deseja colher lucros pelo sucesso de seu trabalho. Investimento sugere a idéia de que a segurança do valor principal investido é importante. A especulação, por outro lado, é bem mais arriscada.

Conforme Luenberger (1998, p. 3), a análise de investimento é o processo de examinar alternativas e decidir qual delas é a preferível. Nesse sentido esta análise é semelhante à análise de outras decisões e esta ciência baseia-se em alguns dos instrumentos utilizados para a análise de outros tipos de decisões. No entanto, os problemas de investimento diferem dos outros problemas decisórios num aspecto importante. Esse importante aspecto consiste no fato de que a maioria dos investimentos são levados a cabo dentro da estrutura de um mercado financeiro e esse mercado provê alternativas não encontradas em outras situações. Essa estrutura é o que faz a análise de investimento ser única e excepcionalmente poderosa.

2.1.1 Entes do mercado financeiro

2.1.1.1 Pessoas

Conforme o manual do DPL 4.0 (1998, p. 4-5) o decisor ou tomador de decisão é uma pessoa ou o grupo que tem a palavra final numa decisão. O decisor é a pessoa que tem autoridade e responsabilidade de comprometer recursos e iniciar (ou interromper) as atividades de uma organização. Geralmente, mas nem sempre, o decisor não é o que realiza a análise da decisão, porém irá analisar os resultados e recomendações.

O decisor racional, além das características operacionais anteriormente definidas, deve seguir alguns princípios fundamentais para ser chamado de **investidor racional** ou simplesmente **investidor**. Esses princípios serão detalhados adiante.

Será suposta apenas a existência de muitos investidores igualmente importantes interagindo.

2.1.1.2 Ativos e instrumentos

Ativo ou bem (*asset*) é algo que seja capaz de produzir fluxo monetário para o proprietário. Qualquer bem com valor comercial ou valor de troca pertencente a uma sociedade, instituição ou pessoa física.

Valor mobiliário (*security*) é um instrumento que indica participação em uma companhia (ações), relacionamento de um credor com uma empresa ou entidade governamental (obrigações), ou direitos de propriedade representados por instrumentos como opção, direito de subscrição e bônus de subscrição.

Valor mobiliário negociável é um valor mobiliário que pode ser facilmente vendido.

Valor mobiliário-base, valor mobiliário objeto ou valor mobiliário subjacente é um valor mobiliário básico do mercado, tal como ações, obrigações e taxas de câmbio. É um valor mobiliário que deverá ser entregue se um contrato de opção de venda ou de compra for exercido. Entretanto, dada a impossibilidade de entregar índices de ações, todas as opções sobre índices de ações e futuros de índices de ações são liquidados em moeda.

Obrigação ou título de dívida (*note and bond*) são valores mobiliários representativos de um empréstimo a médio ou a longo prazos. Nos E.U.A. as obrigações de médio prazo (até 10 anos) são conhecidas por *notes* e as de longo prazo (acima de 10 anos), por *bonds*. É qualquer título de dívida pública ou privada subscrito e negociado com desconto, que rende juros e obriga a emitente a pagar, geralmente a intervalos específicos, determinado montante ao detentor da obrigação, e a reembolsar o principal do empréstimo na data do vencimento. Seus detentores recebem uma promessa de pagamento da emitente, mas não têm

direito a quaisquer dos privilégios assegurados aos acionistas. Um titular de obrigações ao portador apresenta os cupons dos títulos e recebe os juros correspondentes, ao passo que o titular de obrigações nominativas ou registradas aparece nos livros de registros da emitente da obrigação. Uma obrigação garantida (*secured bond*) tem o benefício de certos bens dados como garantia real, que podem ser vendidos pelo detentor da obrigação para satisfação de seu crédito, caso a emitente deixe de pagar os juros do principal quando vencidos. Uma obrigação sem garantia ou debênture tem como garantia apenas a boa-fé e o crédito da emitente, não havendo qualquer garantia real específica. Uma obrigação conversível concede a seu titular o privilégio de permutá-la por outros valores mobiliários da companhia emitente, em data futura e de acordo com determinadas condições. Na área financeira, o termo é também usado para definir a obrigação, de uma pessoa, de pagar uma dívida contraída por um terceiro no caso de inadimplemento deste. A denominação pode também corresponder a recursos em espécie ou títulos dados em caução judicial.

Ação ordinária (*common stock*) é um valor mobiliário representativo da posse de uma parcela da empresa. É a fração do capital social de uma companhia. Os titulares têm o direito de votar para a escolha dos conselheiros e para a solução de questões importantes, tendo também direito a receber dividendos sobre os valores mobiliários de sua propriedade. Caso a companhia seja liquidada, as reivindicações dos credores garantidos e não garantidos e dos proprietários de obrigações e ações preferenciais têm prioridade sobre as reivindicações dos detentores de ações ordinárias. De modo geral, no entanto, estas detêm maior potencial de valorização.

Commodities são mercadorias como cereais, metais e alimentos negociados em uma bolsa de mercadorias ou no mercado à vista. Termo inglês que significa **mercadorias**, designa nas relações comerciais internacionais bens em estado bruto, os chamados produtos primários de importância mercantil, como café, cacau, algodão, lã, chá, estanho, cobre, soja, minério de ferro, etc., cujos preços, em face de sua importância comercial, acabam sendo determinados pelas cotações dos mercados – denominados mercados de *commodities* – de Londres, Nova Iorque e

Chicago, nos quais são realizados a termo (para pagamento e entrega em data futura) [Firmo (1993, p. 74)].

Derivativo, derivado ou instrumento financeiro derivado, conforme Ingersoll Jr. (1987), *apud* Neftci (1996, p. 2), é um contrato financeiro ou direito contingencial se seu valor, na data de expiração, é determinado exatamente pelo preço de mercado do valor mobiliário subjacente nessa data.

Conforme Neftci (1996, p. 2), há três grupos de derivativos: (i) contrato futuro e a termo, (ii) opção e (iii) *swap*. Teoricamente o contrato a termo e a opção podem gerar qualquer outro valor mobiliário, inclusive o futuro e o *swap*.

Contrato a termo é um contrato para entrega futura ou a termo (*forward contract*) para a compra ou venda de uma quantidade especificada de uma *commodity*, títulos do governo, moeda estrangeira ou outro instrumento financeiro ao preço atual ou preço do mercado à vista (*spot price*), com a entrega e liquidação determinadas para data futura. Como se trata de um contrato para entrega física – em oposição a um contrato de opções, em que o proprietário pode escolher entre liquidá-la por entrega ou apenas por diferença -, um contrato para entrega futura pode ser uma cobertura para a venda de contratos de futuros.

Contrato a futuro (*future contract*) é um contrato para a compra de uma mercadoria ou de um ativo financeiro numa data futura, por um preço fixado no presente. Ao contrário dos contratos *forward*, os futuros são transacionados em bolsas e sujeitos a reavaliação diária ao preço de referência. São contratos para compra ou venda de uma quantidade específica de uma *commodity* ou instrumento financeiro a um preço predeterminado, em data futura estipulada. O preço é estabelecido entre comprador e vendedor no pregão de uma bolsa de *commodities*, por meio de um sistema de pregão a viva voz. Um contrato futuro obriga o comprador a adquirir e o vendedor a vender a *commodity*-objeto, salvo se o contrato for vendido a outro antes da data de liquidação, o que pode ocorrer se o investidor deseja realizar lucro ou interromper perdas. Isso contrasta com a negociação de opções, em que o comprador pode escolher entre exercer ou não a opção na data de exercício.

Opção é um contrato que dá o direito, mas não a obrigação de fazer algo numa data futura: direito de comprar ou vender um bem, concedido mediante pagamento de uma quantia acordada entre as partes. Se o direito não for exercido depois de um período especificado, a opção termina pelo vencimento e o comprador da opção perde a quantia paga para obtenção da opção.

Uma carteira de títulos, carteira de investimentos ou portfólio é um conjunto de ativos financeiros possuídos. Carteira de títulos que contém ações, obrigações, *commodities*, investimentos em imóveis, investimentos de títulos de liquidez imediata ou outros ativos de um investidor pessoa física ou investidor institucional. Uma das finalidades do portfólio é reduzir o risco por meio da diversificação.

2.1.1.3 Opções

Opção é um contrato que dá o direito, mas não a obrigação de fazer algo numa data futura. É o direito de comprar ou vender um bem, concedido mediante pagamento de uma quantia acordada entre as partes. Se o direito não for exercido depois de um período especificado, a opção termina pelo vencimento e o comprador da opção perde a quantia paga para obtenção da opção. Uma opção de compra (*call option*) assegura a seu titular o direito de comprar ações do valor mobiliário subjacente ao preço fixado, até uma data especificada no futuro, geralmente 3, 6 ou 9 meses. Por esse direito, o adquirente da opção de compra paga ao vendedor da opção, chamado **lançador**, uma comissão denominada **prêmio** (*premium*), que será perdida se o comprador não exercer a opção até a data acordada. Portanto, o adquirente de uma opção de compra especula, esperando que o preço das ações-objeto suba dentro do período especificado. Por exemplo, uma opção de compra do tipo americana de 100 ações XYZ pode assegurar ao comprador o direito de adquiri-las a R$100,00 cada, em qualquer data nos próximos 3 meses. Para adquiri-la, o comprador pode ter que pagar um prêmio de R$2,00 por ação, ou R$200,00. Se, na data de fechamento do contrato de opção, as ações XYZ estão sendo vendidas a R$95,00 cada, o comprador da opção poderá ter lucro se o preço da ação aumentar. Se, em dois meses, o preço saltar para R$120,00 por ação, por exemplo, o comprador da opção de compra

pode exercer sua opção de comprar 100 ações XYZ a R$100,00 e em seguida revendê-las por R$120,00 cada, retendo a diferença como lucro (excluídos os R$2,00 de prêmio pago por ação). Por outro lado, se as ações XYZ caírem abaixo de R$95,00 e permanecerem nesse nível por 3 meses, tendo a opção de compra terminada ao final do período, o comprador da opção não receberá qualquer retorno sobre o seu investimento de R$200,00 correspondente ao prêmio de R$2,00 por ação. O oposto da opção de compra é a opção de venda (*put option*), que assegura ao comprador o direito de vender um número determinado de ações a um preço, até a data de vencimento. Os adquirentes de opções de venda apostam na queda do preço da ação-objeto. Alguém que previsse uma queda no preço da ação XYZ poderia comprar uma opção de venda de 100 ações para 3 meses, ao preço de exercício de R$100,00 cada ação, e pagar um prêmio de R$2,00. Se o preço da ação XYZ recuar para R$80 cada, o comprador da opção de venda poderá exercer seu direito de vender 100 ações XYZ a R$100,00. Primeiramente, o comprador irá adquirir 100 ações a R$80,00 cada e, em seguida, vendê-las ao vendedor da opção de venda (lançador) a R$100,00 cada, realizando, portanto, um lucro de R$18,00 por ação (lucro de R$20,00 por ação menos os R$2,00 correspondentes ao custo do prêmio por ação). Na prática, as opções de compra e venda são raramente exercidas. Os investidores preferem, ao contrário, comprar e vender opções antes de seus vencimentos, negociando no aumento e queda dos preços dos prêmios. Como o comprador de uma opção precisa de apenas uma pequena quantia (prêmio) para controlar grandes quantidades de ações, as operações no mercado de opções oferecem um alto nível de alavancagem e podem ser altamente rentáveis. Os negociantes de opções podem lançar opções cobertas, em que detêm a propriedade do valor mobiliário subjacente, ou opções descobertas, muito mais arriscadas, pois não têm a propriedade do valor mobiliário subjacente. Geralmente os negociantes de opções perdem muitos prêmios em operações malsucedidas antes de realizar uma operação muito rentável. Os negociantes mais sofisticados combinam várias opções de compra e de venda em posições de *spread* e operações casadas. Seus lucros ou prejuízos resultam da diminuição ou aumento dos *spreads* entre preços das opções.

Uma opção de ações para incentivo é oferecida aos executivos de uma empresa na hipótese de esta atingir determinadas metas financeiras, como determinado nível de vendas ou lucros. Ao executivo é assegurada a opção de comprar ações da empresa a um preço inferior ao do mercado e vendê-las no mercado para realizar lucros.

Opção americana é uma opção que pode ser exercida em qualquer momento até a data final de exercício.

Opção européia é uma opção que pode ser exercida só na data de vencimento.

Opção composta é uma opção sobre uma opção.

Opção de compra (*call*) é uma opção de compra de determinado ativo por um preço de exercício especificado, na data de vencimento ou até essa data. É o direito de comprar, por exemplo, 100 ações de determinada ação ou índice de ação a um preço pré-determinado, antes de um prazo final pré-estabelecido, em contrapartida a um prêmio (pagamento).

Opção de venda (*put*) é uma opção de venda de um ativo por um preço de exercício especificado até uma data pré-definida ou nessa mesma data.

Opções simples são opções cujos resultados dependem apenas do preço do valor mobiliário subjacente.

Opção européia simples ordinária (*vanilla option*) é uma opção cujo valor intrínseco é $\max\{S_\tau - k_\tau; 0\}$ para a opção de compra e $\max\{k_\tau - S_\tau; 0\}$ para a de venda, sendo S_τ o preço de mercado do valor mobiliário e k_τ o preço de exercício na data de vencimento τ.

Opção fora do preço (*out of the money*) é uma opção que não seria vantajoso exercer se vencesse imediatamente.

Opção dentro do preço (*in the money*) é uma opção que seria vantajoso exercer se vencesse imediatamente.

Opção no preço (*at the money*) é uma opção cujo preço de exercício é igual ao preço corrente da opção ou dos contratos futuros subjacentes.

Valor intrínseco da opção (*intrinsic value*) é a diferença entre o preço de exercício de uma opção dentro do preço e o valor de mercado dos ativos-objeto. Por exemplo, se o preço de exercício é de R$53,00 numa opção de compra de ação cujo preço de mercado é R$55,00, a opção tem o valor intrínseco de R$2,00. Ou, no caso de uma opção de venda, se o preço de exercício é R$55,00 e o preço de mercado é R$53,00, o valor instrínseco é também R$2,00. Opções no preço e fora do preço não têm valor intrínseco ou este pode ser considerado igual a zero.

Valor do tempo (*time value*) é a parte do preço de uma opção que reflete o tempo remanescente de um contrato de opção até o seu término.

O preço da opção é composto pelo valor do tempo e o valor intrínseco da opção.

2.1.1.3.1 Um pouco de história

Duas formas de acordo comercial, baseados na expectativa de um cumprimento futuro, tiveram origem direta na aposta.

Para Huizinga (1980), o final da Idade Média assiste, tanto em Gênova como em Antuérpia, ao surgimento do seguro de vida sob a forma de apostas sobre futuras eventualidades de caráter não econômico. Apostava-se, por exemplo, "sobre a vida e a morte de pessoas, o nascimento de um menino ou uma menina, o resultado de viagens e peregrinações, a conquista de várias terras, praças e fortes ou cidades". Este tipo de contrato, embora houvesse já assumido um caráter puramente comercial, foi diversas vezes proibido sob a alegação de tratar-se de jogo ilegal, entre outros, por Carlos V. Apostava-se sobre a escolha de um novo Papa tal como hoje se aposta em corridas de cavalo. Ainda no século XVII os contratos de seguro de vida eram conhecidos pelo nome de "apostas". ♦

2.1.1.4 Preços

Conforme Gélédan & Brémond (1988), preço é o número de unidades monetárias necessárias para poder obter uma mercadoria ou um serviço, em dado momento, dado local e relativamente a uma quantidade específica precisa.

Preço *spot* ou à vista (*spot price*) é o preço de um ativo para entrega imediata (em contraste com o preço do termo (*forward*) ou o preço de um futuro).

Preço de compra é o preço a que um investidor está disposto a comprar.

Preço de venda é o preço a que um investidor está disposto a vender.

Preço de exercício (*strike or exercise price*) é o preço a que uma opção de compra ou de venda pode ser exercida. Preço pelo qual uma ação ou *commodity*-objeto de uma opção de compra ou de venda pode ser comprada (*call*) ou vendida (*put*) ao longo de determinado período. Por exemplo, um contrato de compra pode permitir ao comprador a aquisição de 100 ações de XYZ a qualquer momento nos próximos três meses, a um preço de exercício de R$ 63,00.

Preço ou valor de mercado (*market value or price*) é o último preço pelo qual um valor mobiliário foi vendido em uma bolsa. Para ações ou obrigações vendidas no mercado de balcão, é o preço combinado das ofertas de compra e de venda. O preço de mercado de um valor mobiliário de negociação infreqüente pode ser determinado por avaliadores ou outros analistas. No mundo dos negócios, em geral, o preço de mercado é aquele acordado entre compradores e vendedores por um produto ou serviço, na forma determinada pela oferta e procura.

Preço por arbitragem é o preço único para um ativo financeiro que não permite oportunidade de arbitragem.

Preço da incerteza ou "prêmio do risco" é o valor esperado do excesso de rentabilidade de um investimento com incerteza em relação a outro certo. Soma em dinheiro que um indivíduo avesso à incerteza pagará para evitar assumir uma situação de investimento com incerteza.

Preço de uma opção é o prêmio de uma opção. Cada quantia paga por ação, pelo comprador ao vendedor de uma opção, pelo direito de comprar ou vender o valor mobiliário subjacente a preço determinado em data pré-estabelecida. Nos E.U.A. os preços dos prêmios das opções são cotados em aumentos de 1/8 e 1/16 de 1% e publicados nas tabelas dos jornais diários. Um prêmio de R$5,00 por ação significa que o comprador da opção pagará R$500,00 por uma opção em 100 ações.

Variação mínima de preço (*tick*) é o montante mínimo que o preço de um valor mobiliário pode variar.

Preço racional é o preço único e livre de arbitragem de um ativo.

Uma forma mais primitiva de definir preço correto é o preço justo.

Conforme Downes & Goodman (1991), o valor justo de mercado (*fair market value*) é o preço pelo qual um bem ou serviço é transferido ou fornecido por alguém que deseja vender a outrem que deseja comprar. Presume-se que tanto o comprador quanto o vendedor sejam sensatos e dotados de conhecimento razoável dos fatos pertinentes.

Conforme Neftci (1996, p. 13), hodiernamente usa-se o conceito de arbitragem para obter uma definição operacional de preço justo de um valor mobiliário. Diz-se que o preço de um valor mobiliário está num nível justo ou que o valor mobiliário está com seu preço corretamente determinado, se não há possibilidade de qualquer tipo de arbitragem naquele preço. Tal preço livre de arbitragem irá ser utilizado como *benchmark*, i.e., um padrão ou referência pelo qual outros podem ser medidos ou avaliados.

2.1.1.4.1 Um pouco de história

Huizinga (1980) faz a análise etimológica das palavras **aposta** (*vadium*=penhor no sentido de um objeto puramente simbólico que é atirado dentro do campo de um jogo a título de desafio) e **prêmio**, **preço** e **apreço**, todas originadas de *pretium*.

Pretium surgiu originariamente na esfera da troca e da avaliação, e pressupõe valor. O *pretium justum* medieval corresponde aproximadamente à idéia atual de "valor de mercado".

A palavra **preço** permaneceu limitada à esfera econômica, **prêmio** passou para a do jogo e da competição e **apreço** passou a significar exclusivamente *laus* (estima, louvor)[6].

Já a palavra *wage* (salário) passou de um sentido de aposta (esfera lúdica) para o de preço (esfera econômica), tendo o significado de **ordenado**, aquilo que se

[6] Nota da Revisora.

ganha. A palavra **ganho**, por sua vez, não tem conexão etimológica com as palavras anteriores e pertence tanto à esfera lúdica quanto à econômica.♦

Portanto, o objetivo desta tese é desenvolver um método alternativo, robusto e parcimonioso para a determinação do preço racional da opção européia simples ordinária sobre ação e obrigação (*bond*).

2.1.1.5 Mercados

Conforme Baxter & Rennie (1998), mercado é um lugar geralmente situado num espaço digital para troca de informações sobre preços.

Mercado financeiro é o mercado para negociação de capital e crédito na Economia. O mercado financeiro é composto por valores mobiliários (*securities*) que representam direitos sobre ativos reais. Instrumento que indica participação em uma companhia (ações), relacionamento de um credor com uma empresa ou entidade governamental (obrigações) ou direitos de propriedade representados por instrumentos como opção, direito de subscrição e bônus de subscrição.

Em sentido estrito, o mercado monetário (*money market*) concentra-se em instrumentos de dívida de curto prazo.

O mercado de capitais (*capital markets*) negocia dívidas de longo prazo e valores mobiliários. É um mercado em que são negociados fundos de capital: mercado de títulos representativos de dívida e mercado acionário. Estão aí incluídas as fontes de levantamento de recursos por ações ou obrigações, por meio de colocações privadas, bem como de mercados e bolsas organizados. Exemplos de mercados financeiros: mercado de ações, mercado de *commodities* e mercado de moeda estrangeira.

Mercado acionário (*stock market*) é o termo geral que se refere à negociação de valores mobiliários nas várias bolsas e no mercado de balcão. Os valores mobiliários negociados incluem: ações ordinárias, ações preferenciais, obrigações, valores mobiliários conversíveis, opções, direitos e bônus de subscrição. Quando usado em sentido mais geral, o termo também pode referir-se a *commodities*, mas

geralmente os mercados acionário e o mercado de *commodities* (ou mercado futuro) são diferenciados.

Mercado à vista (*cash market*) é um mercado no qual as transações são realizadas no mercado físico ou à vista; nele a propriedade da *commodity* é transferida imediatamente do vendedor ao comprador e o pagamento é efetuado mediante entrega. O mercado à vista contrasta com o mercado futuro, em que os contratos são para entrega em determinada data futura.

Mercado a futuro (*future market*) é uma bolsa de *commodities* na qual se negociam contratos a futuro.

2.1.2 Os princípios fundamentais da Ciência do Investimento

"*Então será semelhante o Reino dos Céus a [...] um homem que ao empreender uma viagem chamou os servos e lhes entregou os próprios bens. A um deu cinco talentos, a outro dois e a outro um, a cada qual segundo a sua capacidade, e partiu. O que recebera cinco talentos foi logo negociar com eles e lucrou outros cinco. Do mesmo modo, o que recebera dois lucrou, também ele, outros dois. Mas o que recebera um só, foi abrir uma cova na terra e escondeu o dinheiro do seu senhor. Depois de muito tempo, voltou o senhor daqueles servos e chamou-os a contas. O que recebera cinco talentos chegou, apresentou outros cinco talentos e disse: "Senhor, cinco talentos me entregaste, aqui estão outros cinco que eu ganhei". Respondeu-lhe o senhor: "Muito bem! Servo bom e fiel! Foste fiel em coisas de pouca monta, dar-te-ei a superintendência de coisas grandes. Entra no gozo do teu Senhor." Chegou também o que recebera dois talentos e disse: " Senhor dois talentos me entregaste, aqui estão outros dois que eu ganhei." Disse-lhe o senhor: "Muito bem! Servo bom e fiel! Foste fiel em coisas de pouca monta, dar-te-ei a superintendência de coisas grandes. Entra no gozo do teu Senhor." Chegou também o que tinha recebido um só talento e disse: "Senhor, eu sabia que és homem duro, que ceifas onde não semeaste, e recolhes donde não espalhastes. Com medo fui esconder o teu talento na terra. Aqui tens o que é teu". Respondeu-lhe o senhor: "Servo mau e preguiçoso! Sabias que eu ceifo onde não semeei e recolho donde não espalhei? Por isso mesmo devias dar o meu dinheiro aos banqueiros e eu, quando voltasse, receberia o que era meu com os juros. Tirai-lhe, pois, o talento e dai-o ao que tem dez talentos. Porque a todo aquele que tem , dar-se-lhe-á e terá em abundância; ao que não tem, ainda o que tem lhe será tirado. Quanto a esse servo inútil, lançai-o nas trevas exteriores. Aí haverá choro e ranger de dentes*".

Evangelho segundo Mateus[7] 25, 1 e 14-30. Parábola dos talentos. *Bíblia Sagrada: Santos Evangelhos*. 1994. Reimpressão corrigida. Braga, Portugal: Edições Theologica.

Há, baseado em Luenberger (1998, cap. 1), quatro princípios básicos que definem o decisor racional para a construção de modelos para problemas de investimento:

2.1.2.1 Princípio da insaciabilidade

"*B) Dos cosas tambiém, según hemos indicado, acrecientam la caridade una vez obtenida.*

La primera consiste en mantener separado el corazón dos bienes terrenos. En efecto, nuestro corazón no es capaz de darse íntegramente a amores diversos. Nadie pude amar a Dios y al mundo. Por conseguiente, cuanto más se aleja nuestro espíritu del amor de lo terreno, tanto más sólidamente se afianza en el amor de Dios. Y así, Agustín en su libro De diversis quaestionibus octoginta tribus escribe: 'El veneno de la caridade es el afán de obtener y conservar bienes temporales; su auge reside en la mengua del ansia; la perfección de aquélla, en la ausencia de ésta; porque la raíz de todos los males es el ansia'. Por tanto, trabaje en cercenar su ansia quien desee fomentar su caridade. Es el ansia un afán de alcanzar y poseer bienes temporales."

Santo Tomás de Aquino. *Escritos de Catequesis*. 1974. Madri: RIALP. P. 211

O decisor deve preferir mais a menos riqueza.

Quando o valor de um objeto é não observável, como a utilidade, por exemplo, é necessário exigir que o decisor siga o princípio da transitividade para evitar que ele se torne uma bomba de dinheiro.

No entanto, como no mercado financeiro os valores são expressos na forma de preços e os retornos também em valor monetário, torna-se direta a comparação. Isto é, se há duas alternativas de investimento cujos retornos são distintos e certos no mesmo prazo, o decisor deve escolher a de maior retorno ou valor.

---

[7] "Levi-Mateus, pela sua condição de publicano (Mt 9, 9-13), estava a cobrar os impostos no "telônio" - posto público para o pagamento de tributos -. Os publicanos eram cobradores ao serviço dos romanos. Por isso era um ofício odiado e desprezado pelo povo, ainda que, ao mesmo tempo, apetecido pela facilidade de enriquecimento." [Nota explicativa da *Bíblia Sagrada: Santos Evangelhos* (1994, p. 485)]

Merton (1973, p. 141 e 143) emprega apenas este princípio quando define o investidor racional para demonstrar um conjunto de restrições sobre as fórmulas da determinação do preço da opção.

2.1.2.1.1 Exemplo paradigmático de aplicação do princípio da insaciabilidade

Imagine-se que alguém tenha R$1.000,00 e não saiba o que fazer com essa quantia e que seu tio (uma pessoa honesta) - sabendo disso - oferece-lhe a possibilidade de dar-lhe R$1.100,00 daqui a um ano se o sobrinho lhe der R$1.000,00 agora. Considere-se, ainda, que o sobrinho não tem nenhuma obrigação de fazer esse investimento! Imagine-se, também, que só há um banco no mercado financeiro e esse banco, por sua vez, tem o pagamento do depósito garantido pelo Tesouro Nacional, oferecendo aos clientes a possibilidade de depositar R$1.000,00 e, daqui a um ano, retirar R$1.090,00. Qual deve ser a decisão: investir no banco ou no tio? A resposta é investir no tio, pois o retorno é maior.♦

2.1.2.2 Princípio da comparação

O decisor deve avaliar o investimento comparando-o com os outros investimentos disponíveis no mercado financeiro, pois o mercado financeiro provê uma base para a comparação.

2.1.2.2.1 Exemplo paradigmático de aplicação do princípio da comparação

Imagine-se a mesma situação anterior com a diferença de que o decisor pode retirar do banco daqui a um ano R$1.000,00*(1+$r$). Qual deve ser decisão: investir no banco ou no tio?

A resposta depende da taxa de retorno r oferecida pelo banco. Se essa taxa for menor que 0,1, é preferível investir no tio. Caso seja maior, deve-se investir no banco. Caso ela seja igual a 0,1, este princípio é insuficiente para que o decisor possa escolher uma alternativa preferível. No entanto, como será exposto a seguir, essa igualdade é importante para o estabelecimento de um outro importante princípio.♦

2.1.2.3 Princípio de não-arbitragem

"*O melhor negócio do mundo*
é comprar um homem por quanto ele pensa que vale
e vendê-lo por quanto ele realmente vale."

Adágio popular

O mercado financeiro deve ser livre de arbitragem.

Quando há duas alternativas de investimento semelhantes disponíveis no mercado financeiro, surgem conclusões mais fortes que o princípio da comparação.

Uma forma básica de arbitragem é a possibilidade de obter retorno positivo (ganhar dinheiro) sem investir nada.

Uma condição necessária para haver a possibilidade de arbitragem é a existência de um decisor que use o princípio de comparação e da insaciabilidade.

Arbitragem (*arbitrage*) é a compra de um valor mobiliário e a sua venda simultânea para a obtenção de um lucro sem risco ou a realização de lucro garantido sem incerteza com uma ou uma série de transações no mercado.

Arbitragem é a obtenção de lucros com as diferenças de preço quando o mesmo título, moeda ou *commodity* é negociado em dois ou mais mercados. Por exemplo, um arbitrador compra simultaneamente um contrato de ouro no mercado de New York e vende um contrato de ouro no mercado de Chicago, obtendo um lucro porque, naquele momento, há um diferencial de preço entre os dois mercados. (O preço de venda do arbitrador é superior ao de compra.) A arbitragem de índice (*index arbitrage*) explora diferenças de preço entre os futuros de índices de ações (*stock index futures*) e as ações-objeto. Aproveitando-se dos breves diferenciais de preço entre os mercados, os arbitradores desempenham a função econômica de fazer com que tais mercados operem com maior eficiência.

Um mercado livre de arbitragem não possui oportunidades de lucros certos. Uma oportunidade de arbitragem poderia ser uma estratégia de negociação autofinanciável que iniciasse com zero e terminasse numa data futura, com um valor positivo. Um mercado é livre de arbitragem se não houver, de modo algum, tais oportunidades de arbitragem [Baxter & Rennie (1998, p. 197)].

2.1.2.3.1 Exemplo paradigmático de aplicação do princípio de não-arbitragem

Imagine-se que há dois bancos ideais que emprestam e aceitam depósitos à mesma taxa de juros, independentemente do prazo e do montante, e que também não cobram taxas. Imagine-se que o banco A estabeleça a taxa de juros ao ano no valor de 10% e o B, em 12%. Pelo princípio da comparação e da insaciabilidade, o decisor deve tomar o máximo que puder de empréstimo do banco A e depositar todo esse valor em B, uma vez que o ganho de 2% é certo e o decisor não investiu nada.♦

É claro que esta possibilidade é economicamente indesejável. No entanto, se ela existir, rapidamente deve desaparecer no mercado. No caso do exemplo anterior, os dois bancos estabelecerão a mesma taxa de juros.

Quando os bancos, neste exemplo, definem que a taxa de juros anual é de 11%, o decisor não consegue mais arbitrar. Isto é, ele terá que ganhar dinheiro de outra forma que, sem dúvida, será mais árdua que esta.

Portanto, admitir-se-á que no mercado financeiro não há a possibilidade de qualquer tipo de arbitragem. Esta é a hipótese de não-arbitragem.

Assim, dois valores mobiliários distintos com idênticas propriedades devem ter o mesmo preço para que não haja possibilidade de arbitragem no mercado.

2.1.2.3.2 Exemplo paradigmático da aposta entre o circunspeto professor de Economia e o matuto

Um circunspeto professor de Economia e um matuto estavam esperando um ônibus numa importante cidade. Para passar o tempo, o matuto sugeriu que eles disputassem um jogo que consistia em cada um fazer uma pergunta ao outro, começando pelo matuto. Caso a pergunta não fosse respondida, um real deveria ser pago ao inquiridor. O prudente e circunspeto professor de Economia lembrou ao humilde matuto quem ele era: um douto professor de Economia! O matuto pensou e resolveu alterar a aposta. Caso ele – o matuto – não fosse capaz de responder à pergunta do circunspeto professor, ele pagaria apenas cinqüenta centavos. E, caso o circunspeto professor não fosse capaz de responder à pergunta

do matuto, um real seria ainda pago ao humilde matuto. O circunspeto professor ainda comentou que agora o jogo parecia ser justo. O matuto logo disparou a primeira pergunta: "O que sobe o morro com sete pernas e desce o morro com três pernas?". O circunspeto professor de Economia meditou profundamente por uns momentos e finalmente exclamou com um ar de surpresa: "Eu não sei...". O matuto calmamente retrucou: "Bem, eu também não sei. Mas se o senhor me der um real, eu lhe darei meus cinqüenta centavos." Essa estória ilustra o conceito de arbitragem: conseguir uma transação não envolvendo dispêndio de dinheiro que resulta em lucro certo. O texto foi adaptado de Varian (1987, p. 55).♦

Evidentemente estas oportunidades devem ser raras em mercados bem constituídos. Conforme Varian (p. 56) e Neftci (1996, p. 12), é geralmente aceito que parte das condições gerais de equilíbrio é que não haja possibilidade de arbitragem de qualquer tipo.

Ross (1976) *apud* Yang (1997, p. 5), Luenberger (1998, p. 4-5) e Varian (1987, p. 56) consideram o princípio de arbitragem como o mais básico princípio unificador no estudo da moderna teoria da Ciência do Investimento.

Conforme Neftci (1996, p. 12-13), no seu sentido mais básico, **arbitragem** significa tomar posições simultâneas em diferentes ativos de tal forma que se garanta um lucro sem risco mais alto que o retorno sem risco dado pelas letras do Tesouro dos E.U.A. Se tais lucros existem, diz-se que há uma oportunidade de arbitragem. Os preços de ativos livres de arbitragem são utilizados como referência.

Luenberger (1998, cap. 9) distingue dois tipos de arbitragem: (A) obtenção de algo do nada agora e (B) obtenção de uma possibilidade de algo no futuro sem custo agora. A impossibilidade da arbitragem do tipo A permite a existência do princípio da formação de preço linear. A impossibilidade dos dois tipos de arbitragem permite determinar, de maneira bem definida a carteira, que maximiza a utilidade esperada.

A formalização do princípio de não-arbitragem deveu-se a S. A. Ross em 1976.

Conforme Yang (1997, p. 12-3), a hipótese de não-arbitragem afirma que, se há uma carteira de ativos que replica os resultados de uma opção, então a opção e a carteira devem ter o mesmo preço.

Conforme Neftci (1996, p. xix-xx), o chamado teorema da arbitragem ou teorema fundamental de Finanças é o maior resultado da teoria de Finanças. Este autor afirma que, mais importante do que fornecer uma estrutura matemática, este teorema justifica a existência e a utilização da medida de probabilidade neutralizadora do preço da incerteza , pois esta medida de probabilidade torna possível evitar considerações relativas aos prêmios pelo risco do valor mobiliário.

Este princípio, baseado na existência de mercados bem desenvolvidos, permeia uma boa parte da moderna Ciência do Investimento.

Conforme Luenberger (1998, cap. 1) o princípio de não-arbitragem implica a relação linear (aditiva) dos preços dos valores mobiliários e que os preços dos derivativos como futuros e opções podem ser determinados analiticamente.

Ross (1997, p. 530-6) exemplifica o apreçamento da opção européia simples ordinária sobre ação. Ele emprega o exemplo para ilustrar a aplicação do teorema fundamental de Finanças, isto é, o teorema da arbitragem. A medida de probabilidade neutralizadora do preço da incerteza é manifesta de uma maneira muito elegante.

2.1.2.3.3 Exemplo paradigmático de determinação de preço de opção e da aplicação do teorema da arbitragem

"Muito menos, ninguém de senso se daria ao trabalho de definir a tecelagem só por amor da própria arte de tecer. Mas o que a maioria não percebe, segundo penso, é que algumas coisas apresentam semelhanças naturais, muito fáceis de apontar aos que pedem explicações de certas realidades, quando não queremos ter trabalho de grandes explanações; ao passo que para as mais importantes verdades e da mais alta relevância, não existe imagem capaz de fornecer aos homens uma idéia clara que, para completa satisfação do espírito, bastaria apresentar a quem nos interroga, com a devida acomodação a qualquer dos sentidos. Por isso mesmo, precisamos esforçar-nos para ficar em condições de dar ou compreender a razão de cada coisa. As realidades imateriais, as maiores e as mais belas, só nos podem ser claramente reveladas por meio da razão, nada mais, sendo a elas que se refere tudo o que dissemos até agora. Mas, é muito mais fácil tratar das coisas pequenas do que das grandes."

Platão. 1980. *Diálogos*: *Político* ou *Da Realeza*, 285de a 286ab.
Tradução de Carlos Alberto Nunes. Universidade Federal do Pará.

Ross (1997) supõe que há uma taxa de desconto positiva. Aqui vai-se supor que esta taxa de desconto é nula. Esta alteração não muda o problema essencialmente.

Imagine-se que uma ação valha R$100,00 hoje. Daqui a um ano, ela pode valer R$200,00 ou R$50 (estes valores podem ser entendidos como valores presentes, pois a taxa de desconto é nula). O preço de exercício de uma opção sobre esta ação, daqui um ano, é R$150,00. A pergunta é: "Qual o preço racional da opção de compra da ação?".

Suponha-se que seja possível montar uma carteira de ações e opções sobre estas ações. De forma a eliminar completamente o risco (*hedging* perfeito), deseja-se montar uma carteira que não varie de preço no futuro. Portanto:

$$200x + (200 - 150)y = 50x + 0y$$

sendo:

x quantidade de ações adquiridas hoje

y quantidade de opções de compra adquiridas hoje

\+ comprado

\- vendido.

Observe-se que o valor 200-150=50 é ganho por exercer a opção e 0 (zero) é o ganho por não tê-la exercido.

Portanto:

$$y = -3x\,.$$

Isto é, necessita-se comprar x unidades de ações e vender $3x$ unidades de opções ou o contrário.

Suponha-se que o preço racional da opção seja c. Então o custo da carteira hoje é:

$$100x + cy = 100x - 3cx\,.$$

Logo, o ganho hoje pode ser computado:

$$\begin{aligned}\text{ganho}_{\text{hoje}} &= \text{valor da carteira}_{hoje}^{R\$50,00} - \text{custo da carteira}_{\text{hoje}} \\ &= 50x - (100x - 3cx) \\ &= x(3c - 50)\end{aligned}$$

Portanto, se $c^* = \text{R\$}\frac{50}{3} \cong \text{R\$}16,67$, o ganho é nulo. Logo, não há possibilidade de arbitragem. Caso contrário, há a possibilidade de ganho certo (sem incerteza) e sem ter dispêndio de dinheiro, isto é, há a possibilidade de arbitragem.

A arbitragem ocorre se, por exemplo, o preço em real da opção for $c = 20 > c^*$. Então, comprando-se uma ação $(x = 1)$ e simultaneamente vendendo-se três opções $(y = -3)$ custa $100 - 60 = 40$ reais. No entanto, no ano seguinte, o valor do portfólio será R$50,00 se a ação subir para R$200,00 ou se ela cair para R$50,00. Então, obtém-se um lucro garantido de R$10,00.

Similarmente, se o preço em real da opção for $c = 15 < c^*$, então, vendendo-se uma ação $(x = -1)$ e comprando três opções $(y = 3)$, tem-se um ganho inicial de $100 - 45 = 55$ reais. Por outro lado, o valor desta carteira, no ano seguinte, será de -50 reais. Obtém-se, então, um lucro certo de R$5,00.

Um plano de aposta com ganho certo é chamado de uma **arbitragem**.

O preço da opção que impossibilita a arbitragem é chamado de **preço racional** e utilizado no mercado como *benchmark.*

Este mesmo resultado pode ser obtido aplicando-se o teorema da arbitragem [Ross (1997, p. 533-6)].

Suponha-se que haja dois conjuntos:

$$\begin{aligned} s &= \{200; 50\} \\ \mathbb{Q} &= \{q; 1 - q\} \end{aligned}$$

O primeiro é chamado de **domínio** (conjunto de apoio) e o segundo, de **medida de probabilidade**.

Serão criadas duas apostas: (i) comprar/vender uma ação e (ii) comprar/vender uma opção.

(i) Aposta 1: comprar/vender uma ação

$$R_1 = \begin{cases} 200 - 100 = 100, \text{ se açao valer } 200 \\ 50 - 100 = -50, \text{ se açao valer } 50 \end{cases}$$

$$\mathrm{E}_{\mathbb{Q}}(R_1) = 100q - 50(1-q) = 0$$

$$q = \frac{1}{3}$$

$$\therefore \mathbb{Q} = \left\{ q = \frac{1}{3}; 1 - q = \frac{2}{3} \right\}$$

R é a variável aleatória retorno.

Portanto, existe apenas uma medida de probabilidade neutralizadora do preço da incerteza e, por conseqüência, não há possibilidade de arbitragem. Além disso, existe e é único o preço da opção que será determinado a seguir.

(ii) Aposta 2: comprar/vender uma opção

$$R_2 = \begin{cases} 200 - 150 - c, \text{ se açao valer } 200 \\ -c, \text{ se açao valer } 50 \end{cases}$$

$$\mathrm{E}_{\mathbb{Q}}(R_2) = (50 - c)\frac{1}{3} - c\frac{2}{3} = 0$$

$$\therefore c = \text{R\$}\frac{50}{3} \cong \text{R\$}16{,}67$$

Observe-se que a medida de probabilidade neutralizadora do preço da incerteza independe da probabilidade real ou subjetiva.

O teorema da arbitragem, segundo Ross (1997, p. 533-4), é uma conseqüência do teorema da separação de hiperplano, que é freqüentemente utilizado como um mecanismo para provar o teorema da dualidade da programação linear.♦

Portanto, o preço racional de uma opção existe se e somente se existir uma única medida de probabilidade neutralizadora do preço da incerteza $\mathbb{Q}$.

Conforme Yang (1997, p. 7), o princípio de arbitragem fornece uma condição suficiente para a existência de uma medida de probabilidade neutralizadora do preço da incerteza. Entretanto, ele não fornece um meio de determiná-la. Estimação e especificação da medida de probabilidade neutralizadora do preço da incerteza tornou-se um tema central em Finanças. Tornou-se comum caracterizar a medida de probabilidade neutralizadora do preço da incerteza supondo um processo estocástico para o preço do valor mobiliário subjacente.

Admitindo-se que não exista arbitragem no mercado, o preço de um valor mobiliário cujos resultados (*e.g.* retornos) são uma combinação linear de outros ativos é exatamente igual à combinação linear dos preços dos valores mobiliários. Este resultado é conhecido como Teorema da Aditividade de Valor [Varian (1987, p. 61)]. Quando a hipótese do mundo com preço da incerteza neutralizado for definida, este teorema será mais detalhadamente tratado.

2.1.2.4 Princípio da aversão à incerteza

O decisor deve preferir menos a mais incerteza.

O investidor racional deve ser avesso à incerteza.

Conforme Bierman Jr. & Smidt (1993, p 365-7), a incerteza surge quando pode haver mais que um possível fluxo de caixa que pode ser produzido por um dado ativo num dado tempo futuro.

O estado de conhecimento certo é o mais confortável para o decisor.

Em geral o decisor deve preferir mais incerteza a menos se desejar um nível de retorno maior do que o retorno associado ao investimento com menos incerteza.

Moore (1986, cap. 9) imagina que há dois investimentos com o mesmo custo e que ambos tenham o mesmo retorno esperado maior que este custo inicial. O termo esperado é definido probabilisticamente. Considera ainda que o retorno é certo para um dos investimentos e incerto para o outro. Logo, o decisor que procura investir, ao invés de especular, deve preferir o investimento sem incerteza.

Há evidências empíricas de que o investidor típico é avesso à incerteza.

Pindyck & Rubinfeld (1994, p. 191-202) perguntam (i) se os executivos de empresas seriam mais apreciadores de incerteza do que a maioria das pessoas e (ii) se, quando em face de diferentes alternativas de estratégia, algumas envolvendo incerteza, e outras garantidas, quais seriam aquelas que elas escolheriam.

Foi realizado um estudo com 464 executivos que responderam a um questionário com descrições de situações de risco com as quais eles poderiam se defrontar imaginando serem vice-presidentes de uma empresa hipotética[8]. O estudo revelou que aproximadamente 20% foram neutros à incerteza, 40% propensos à incerteza, 20% avessos à incerteza e 20% não responderam ao questionário. O estudo mostra que os executivos de maneira geral fizeram significativos esforços para reduzir ou eliminar as incertezas, geralmente mediante postergação de decisões ou coleta de informações adicionais.

Em seguida, Pindyck & Rubinfeld (1994, p. 191-202) afirmam que, diante de uma ampla variedade de situações de risco, os consumidores geralmente mostram-se avessos à incerteza. Há normalmente três maneiras de os consumidores reduzirem a incerteza: diversificação, seguros e obtenção de informações adicionas a respeito de opções e desfechos. Afirmam ainda que a maioria das pessoas tem aversão à incerteza. Outras perguntas importantes de Pindyck & Rubinfeld (1994, p. 191-202) são: "Por que razão indivíduos que têm aversão à incerteza são capazes de fazer investimentos no mercado acionário, portanto, arriscando-se a perder parte ou a totalidade de seus investimentos? Como será que as pessoas tomam suas decisões quanto ao grau de risco que assumirão ao fazer seus investimentos e ao realizar planos para o futuro?"

Fishburn & Kochenberg (1979)[9], *apud* Poulton (1994, p. 199-208), analisaram as funções utilidade estimadas empiricamente para mudanças na riqueza ou retorno sobre investimento de trinta decisores em várias áreas de negócios. Os autores analisaram quão adequadas são as funções potência de Kahneman & Tversky (as pessoas subavaliam grandes perdas relativamente a menores ganhos e perdas), logarítmica de Bernoulli e exponencial para explicar o comportamento de cada

---

[8] MacCrimmon, K. R. & Wehrung D. A. 1984. "The risk in-basket" *Journal of Business* 57: 367-87.

[9] Fischburn, P.C. & Kochenberger, G. A. 1979. "Two-piece Von Neumann-Morgenstern utility functions." *Decision Sciences* 10: 503-18.

decisor. Concluem que oito dos trinta decisores têm a função potência de Kahneman & Tversky e dois terços têm funções logarítmicas ou exponenciais.

No entanto, Metrick[10] (1995), *apud* Grosjean (1998), utilizou dados obtidos na final de um jogo chamado *Jeopardy!* (*quiz show*) para analisar questões econômicas e concluiu que os jogadores são aproximadamente neutros à incerteza.

2.1.2.4.1 Quantificação da incerteza

"In the Physics, Aristotle´s inquiry into randomness and the nature of chance events is conceptual or ontological and not mathematical. This does not mean, however, that this contributions to chance studies are worthless. Aristotle presented a careful analysis of the random and intentional properties of chance events. As a result, we may wonder whether it is necessary or possible to make a strict dichotomy between statistical and epistemic probability. If so, then epistemic questions must be given priority, because before calculating, one must determine whether what is to be measured is a suitable subject for calculation and wich aspects are to be quantified."

Styan (1998, p.15)

A probabilidade normalmente é tida como um quantificador de incerteza.

No entanto, conforme Jaynes (1996, p. 2-18), qualquer distribuição de probabilidade é necessariamente subjetiva no sentido de descrever apenas um estado de conhecimento do decisor, e não poder de modo algum ser mensurada por um experimento físico. Ela pode ser objetiva se independer da personalidade do usuário (sentimentos como esperança, medos, valores de julgamento etc.). Portanto, a probabilidade denota um estado informacional do investidor sobre um evento futuro. Logo, a probabilidade não é a medida da incerteza ou da surpresa de um evento de um espaço amostral. A Teoria da Informação estabelece que a medida da incerteza é o informacional cuja unidade de medida é o bit. O informacional é uma função logarítmica do inverso da probabilidade de ocorrência de um evento.

---

[10] Metrick, A. 1995. "A natural experiment in Jeopardy!" *American Economic Review* 85: 240-253.

A incerteza tem como sinônimos as palavras, conforme Ebrahimi *et al.* (1999), **volatilidade**, **variabilidade**, **dispersabilidade** e **variedade**.

Uma das mais populares medidas de incerteza é a variância.

Uma medida que concorre com a variância é a entropia oriunda da Teoria da Informação.

A medida do retorno esperado é a esperança matemática.

A análise de média-variância é uma das maneiras de formalizar matematicamente este princípio. Essa análise estabelece que, se várias oportunidades de investimento têm a mesma média, mas diferentes variâncias, o decisor deve preferir aquele com a menor variância.

Outra forma mais geral de selecionar uma entre várias oportunidades de investimento com incerteza é utilizar funções de utilidade individual, conforme Luenberger (1998, cap. 1, 9, 15 e 16).

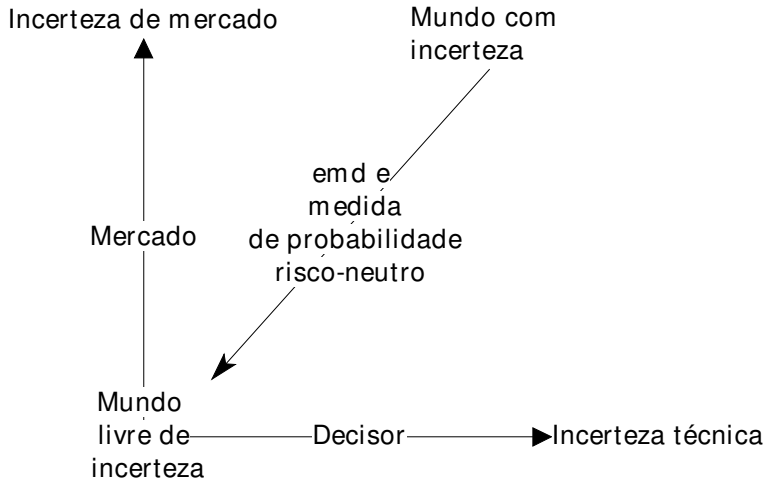
No entanto, no espírito de Kelly Jr. (1956), quanto menos informações subjetivas (e.g. função de utilidade) forem necessárias para a obtenção do preço único e livre de arbitragem, melhor (princípio da parcimônia). Dessa forma, a Teoria da Informação é uma terceira maneira alternativa de modelar problemas de investimento em condições de incerteza.

2.1.2.4.2 Incerteza técnica e de mercado

Conforme Luenberger (1998, p. 458-63), há dois tipos de incerteza: a técnica e a de mercado. Por exemplo, o fluxo de caixa de uma mina de ouro depende tanto da incerteza de mercado do preço do ouro quanto da incerteza técnica da quantidade de ouro que ainda há para ser extraído. Incerteza técnica inclui eficiência de produção desconhecida (devido a novos processos de produção), incerteza sobre recursos, tal como a quantidade de petróleo num campo petrolífero), incerteza sobre resultado (como numa pesquisa ou projeto de desenvolvimento) e um componente de incerteza do preço de bens com pouca liquidez (preço futuro de um terreno isolado). Incertezas de mercado são aquelas

associadas aos preços de *commodities* e valores mobiliários negociados e outros ativos reais.

Em situações com os dois tipos de incerteza, isto é, com a presença de incerteza técnica e de mercado, a solução deverá conter as medidas de equivalente monetário determinístico (*certainty equivalent*) para contemplar a atitude do investidor diante de risco em condição de incerteza técnica (*private uncertainty*) e da probabilidade neutralizadora do preço da incerteza para contemplar a condição de incerteza de mercado (*market uncertainty*). O que se observa é o movimento da Ciência do Investimento de um mundo com prêmio pela incerteza para um mundo virtual sem preço da incerteza.



A medida de probabilidade neutralizadora do preço da incerteza manifesta-se tanto em condição de incerteza técnica quanto em condição de incerteza de mercado. A natureza da incerteza acarreta a forma de *hedging*.

2.1.2.4.3 Um pouco de história

Para Bernstein (1997, p. 53), o *Livro dos Jogos de Azar*, iniciado em 1525 e reescrito em 1565, de Girolamo Cordano (1500-1571), médico e jogador inveterado, foi o primeiro esforço de pôr a medição a serviço do risco. No entanto, para Stigler (1986, p. 63) o pai da quantificação da incerteza é Jacob Bernoulli (1654-1705), matemático, filósofo e professor da Universidade de Basel a partir de 1687, contemporâneo e rival ocasional de Isaac Newton. Conforme Bernstein (1997, cap. 7), por meio de sua obra póstuma *A Arte da Conjectura*, publicada

em 1713, ocorre o início da formalização do conceito de probabilidade, evento, probabilidade prévia (*a priori*) e posterior (*a posteriori*) e a lei dos grandes números. Esse livro procura demonstrar onde termina a arte de pensar objetivamente e começa a arte da conjectura, de modo a entender a conjectura como o processo de estimar o todo a partir das partes.♦

2.1.2.4.4 Relação entre risco financeiro e incerteza

[...] *A mulher é um jogo*
Difícil de acertar
E o homem como um bobo
Não se cansa de jogar
O que eu posso fazer
É se você jurar,
Arriscar a perder
ou desta vez então ganhar [..]

Ismael Silva, Nilton Bastos e Francisco Alves (1931).
1982. *História da música popular brasileira.* São Paulo: Abril.

Risco é uma conseqüência da decisão livre e consciente de expor-se a uma situação na qual se luta pela realização do bem havendo a possibilidade de ferimento.

Perda ou dano é um ferimento humano físico ou psicológico. Conforme Pieper (1960, p. 173), "entende-se por ferimento toda a violação da incolumidade natural contrária à nossa vontade, toda a ofensa ao ser que em si próprio descansa, tido o que contraria a nossa vontade, tudo o que é de qualquer modo negativo, tudo o que magoa e prejudica, atemoriza e oprime. O ferimento mais profundo e extremo é, porém, a morte. E mesmo os ferimentos não mortais são imagem da morte". Portanto, conforme Pieper (1960, p. 179-80), o ser humano que possui a virtude cardeal da fortaleza expõe-se ao perigo da morte por causa do bem. E só quem é prudente (vê a realidade como ela é, i.e., objetivamente) e justo (dá a cada um o que lhe é devido) pode ser forte.

Luenberger (1998, cap. 1), por exemplo, utiliza a palavra **risco** ao invés de **incerteza**. Preferiu-se aqui utilizar o termo **incerteza** devido à seguinte definição para **risco financeiro**:

Risco financeiro é uma conseqüência da decisão livre e consciente de expor-se a uma situação na qual há a expectativa de ganho sabendo-se que há a possibilidade de perda ou dano.

As definições de risco e de risco financeiro estão baseadas em Pieper (1960), Styan (1998), Steiner Neto (1998, p. 50-6), Spradlin (1997), Luenberger (1998), Moore (1986), Dargahi-Noubary & Growney (1998), *The new Merriam-Webster dictionary* (1989), *Merriam-Webster´s pocket guide to synonyms* (1995) e Ferreira (1986).

Desta forma, **risco financeiro**, ao longo da tese, será chamado apenas de **risco**.

Logo, a incerteza é um componente necessário mas não suficiente para o conceito de risco.

Em geral, a palavra **risco** aparece com o significado de perigo ou chance. No entanto, nenhum destes constitui seu sentido estrito.

Para March & Shapira (1987, p. 1404), *apud* Steiner Neto (1998, p. 50), **risco** é normalmente definido como reflexo das eventuais variações nas distribuição dos retornos possíveis, com as suas probabilidades e com os seus valores subjetivos.

Para Maccrimmon & Wehrung (1986, p. 9), *apud* Steiner Neto (1998, p. 51), existem três condições para a definição de risco, que são denominados componentes do risco:

- Deve existir a possibilidade de haver perda ou dano (magnitude da perda);
- Deve haver uma possibilidade associada a essa perda (possibilidade de perda);
- Deve haver a possibilidade de o decisor agir de forma tal que aumente ou diminua a magnitude ou a probabilidade dessa perda ou dano (exposição à perda);

Os mesmos autores também determinam as fontes causadoras do risco, que seriam:

- Ausência de controle;
- Ausência de informações;
- Ausência de tempo.

Conclui Steiner Neto (1998, p. 51) que o conceito de risco implica a possibilidade de escolha.

Steiner Neto (1998, p. 70) afirma que Maccrimmon & Wehrung (1986, p. 99) mostram que retornos positivos são tratados como oportunidades e de maneira significativa, enquanto o risco é associado apenas aos retornos negativos.

Conforme March & Shapira (1987, p. 1407), *apud* Steiner Neto (1998, p. 70), a questão da utilização da probabilidade de ocorrência de um retorno como algo importante é questionada por eles, que citam as palavras de um executivo entrevistado: "Eu aceito grandes riscos, independentemente da probabilidade, mas não grandes valores".

Conforme Steiner Neto (1998, p. 50), o risco faz parte do dia-a-dia dos gerentes e executivos de tal forma que, muitas vezes, passa despercebido e a convivência com o risco pode ser definida como algo natural na vida.

A típica situação de risco é o jogo de azar. O jogador decide jogar determinado jogo de azar porque espera ganhar, mesmo sabendo que pode perder.

Uma distinção entre **jogatina** (*gambling*), **especulação** e **investimento** encontra-se em Moore (1986, cap. 9).

O que há em comum entre os três conceitos é a presença da incerteza.

Além disso, para ser bem-sucedido em qualquer um deles, o conhecimento dos componentes dos riscos (perdas/ganhos e incertezas) envolvidos é essencial.

Para a jogatina, o ganho esperado é negativo, e a única vantagem ocorre se a utilidade de um alto prêmio com baixa probabilidade supera a perda do dinheiro apostado. Uma empresa normalmente não irá jogar, já que no jogo ela não pode

controlar os riscos, mas pode especular ou investir em um novo produto, argumentando que, nessas situações, ela pode controlar os resultados no sentido de que ele é o melhor preditor dos resultados dos outros participantes do mercado.

Conforme Bernstein (1997, p. 21) o executivo de sucesso é antes de tudo um previsor; comprar, produzir, vender, fixar preços e organizar vêm depois.

Especulação e investimento diferem principalmente nos níveis de incertezas envolvidas, mas, na especulação, a presença da incerteza é maior.

Portanto, a especulação (*speculation*) é a aceitação de risco que prevê ganho mas reconhece possibilidade de prejuízos maiores que a média. A especulação é uma atividade necessária e produtiva. Pode ser lucrativa a longo prazo, quando realizada por profissionais que geralmente limitam suas perdas utilizando várias técnicas de *hedging*, inclusive operações com opções, venda a descoberto, ordens para prevenção de prejuízos e negociações de contratos futuros. O termo **especulação** indica que um negócio ou investimento de risco pode ser analisado e avaliado e se diferencia do termo **investimento** pela graduação do risco.

Logo, investimento (*investment*) é o uso de capital para gerar mais recursos, seja por meio de negócios geradores de renda ou de empreendimentos mais arriscados destinados a ganhos de capital. Pode se referir a um investimento financeiro (quando um investidor põe dinheiro em determinado negócio) ou a um investimento de esforço e tempo por parte de um indivíduo que deseja colher lucros pelo sucesso de seu trabalho. Investimento sugere a idéia de que a segurança do valor principal investido é importante.

A especulação, por outro lado, é bem mais arriscada.

Segundo Sassatani (1999), na Economia, não é incomum encontrar o termo **risco** como sinônimo de **incerteza**. De acordo com Machina e Rotstchild (1990, p. 227), *apud* Sassatani (1999), risco pode ser utilizado como sinônimo de **incerteza** ou **informação incompleta**. Um dos tópicos estudados na Microeconomia trata da escolha na incerteza [Kreps (1990) e Pyndick (1995)]. Contudo, coube ao economista Frank Knight (1972, p. 249) fazer a distinção entre risco e incerteza:

"A diferença prática entre as duas categorias, risco e incerteza, é que na primeira a distribuição do resultado num grupo de casos é conhecida (quer através do cálculo a priori, quer das estatísticas da experiência passada), enquanto no caso da incerteza isso não ocorre, em geral devido ao fato de que é impossível formar um grupo de casos, porque a situação que se enfrenta é, em alto grau, singular". Pela definição de Knight, a diferença básica entre risco e incerteza consiste na presença ou não de uma distribuição de probabilidades sobre um certo evento. Dessa forma, **incerteza** se refere a situações em que não se conhece a distribuição de probabilidade dos resultados. **Risco** é a situação em que se podem estabelecer os possíveis resultados e suas respectivas probabilidades de ocorrência.

2.1.2.4.4.1 Gestão do risco

A gestão do risco consiste em obter informações adequadas para conhecer melhor a situação de risco e/ou intervir nela, tendo como resultado a melhoria da qualidade das decisões nesta situação, com possibilidade de perda ou dano.

Os componentes da perda potencial (componente negativa do risco) são: (i) magnitude, (ii) chance de ocorrência e (iii) grau de exposição.

Os elementos determinantes da perda numa situação com risco, adaptado de Steiner Neto (1998, p. 52), são: (i) impossibilidade de dominar as forças da natureza (condições climáticas, leis da natureza etc.), o comportamento humano (livre arbítrio, ações e atitudes individuais e coletivas etc.) e os recursos limitados (tempo, capital etc.) e (ii) informação incompleta, podendo ela ser inadequada, inconfiável, não familiar, imprevisível ou inacessível.

O exemplo paradigmático de gestão de risco a seguir foi inspirado no exemplo de Steiner Neto (1998, p. 50).

2.1.2.4.4.1.1 Exemplo paradigmático de gestão de risco

Situação de risco:	dirigir um automóvel na cidade de São Paulo
Ganhos esperados:	flexibilidade de locomoção na cidade e visibilidade da condição financeira e social

Perdas e danos potenciais:	ferimentos pessoais, parentes, outros motoristas e transeuntes, ofensas morais, defeitos do carro, doenças provocadas pela poluição e estresse, multas, aumentos de impostos, desvalorização, roubo/furto, atropelamento, batida etc.

O quadro a seguir foi adaptado de Steiner Neto (1998, p. 52).

	Componentes da perda potencial		
Determinantes do perda potencial	*Magnitude*	*Chance*	*Grau de exposição*
Falta de controle **Como reduzir...?**	Fazer seguro do carro, possuir um carro barato, locar um carro etc.	Estacionar sempre em estacionamento pago, evitar caminhos perigosos, possuir um carro discreto e não muito visado por ladrões, dirigir com cuidado e usar cinto de segurança, possuir um carro seguro (ar condicionado, *air bag* e blindado com vidros à prova de bala; não dirigir e utilizar transporte coletivo; morar perto do trabalho e escola etc.	Não dirigir e utilizar transporte coletivo ou motorista particular, morar perto do trabalho e escola.
Falta de informação	**Qual a sua magnitude?** Acompanhar os preços de mercado de novos e usados e os preços dos serviços, manutenção e peças de reposição.	**Qual é a probabilidade de sua ocorrência?** Acompanhar as Estatísticas de ocorrências por tipo de sinistro.	**Quais são as alternativas para compartilhá-la?** Há várias empresas seguradoras, corretores, tipos de seguro (prazo, preço, condições etc.).

Conforme Steiner Neto (1998, p. 53) a atitude do decisor diante do risco é uma prerrogativa individual.

A atitude diante do risco supõe a consideração da atitude diante da incerteza e da perda potencial e da expectativa de ganho da situação.

2.1.2.4.4.2 Um pouco de história

"[...] Em verdade, os negócios humanos não merecem a menor consideração da nossa parte. Todavia, somos como que obrigados a considerá-los de perto, o que não deixa de ser aborrecido. [...] O que eu digo, é que devemos aplicar-nos seriamente ao que é sério, não ao que não for. Por sua própria natureza, Deus é digno de todo o nosso zelo religioso, ao passo que o homem, conforme já o observamos, foi feito para servir de joguete nas mãos da divindade, no que, aliás, consiste todo o seu merecimento. Importa, pois, ao homem conformar-se com sua sorte e entreter-se a vida inteira com belos jogos: eis como os homens e as mulheres precisam viver, em contrário, justamente, à sua atual maneira de pensar."

Platão. Leis (1980, 803bcd)

Duas formas de acordo comercial, baseados na expectativa de um cumprimento futuro, tiveram origem direta na aposta.

Para Huizinga (1980), o final da Idade Média assiste, tanto em Gênova como em Antuérpia, ao surgimento do seguro de vida sob a forma de apostas sobre futuras eventualidades de caráter não econômico. Apostava-se, por exemplo, "sobre a vida e a morte de pessoas, o nascimento de um menino ou uma menina, o resultado de viagens e peregrinações, a conquista de várias terras, praças e fortes ou cidades". Este tipo de contrato, embora houvesse já assumido caráter puramente comercial, foi diversas vezes proibido sob a alegação de tratar-se de jogo ilegal, entre outros por, Carlos V. Apostava-se sobre a escolha de um novo Papa tal como hoje se aposta em corridas de cavalo. E, ainda no século XVII, os contratos de seguro de vida eram conhecidos pelo nome de "apostas".

Huizinga (1980) afirma que a Revolução Francesa deflagrou uma onda de seriedade no Ocidente, que perdura até os nossos dias. O esporte, as artes, os negócios, a moda, a ciência vivem uma crise lúdica. Entretanto, no homem, o desejo de jogar é um traço indelével da sua natureza. Ainda segundo o autor, a avareza em seu estado puro não comercia e muito menos joga, i.e., jamais arrisca. Faz notar que a essência do espírito lúdico é ser audacioso, correr riscos e suportar a incerteza e a tensão. A tensão faz com que o jogo adquira mais importância, levando o jogador a esquecer que está jogando. Um jogador de roleta admite facilmente que está jogando. No entanto, um corretor de valores sustentará que a sua atividade de comprar e vender em função do comportamento

da Bolsa fazem parte das coisas sérias da vida ou, pelo menos, da vida dos negócios, e constitui uma função econômica da sociedade. Nos dois casos o fator operante é a expectativa de lucro. Mas, enquanto no primeiro caso o caráter puramente fortuito da coisa é geralmente reconhecido, no segundo, o jogador ilude-se a si mesmo com a idéia de que é capaz de prever a tendência futura do mercado. Seja como for, é ínfima a diferença de mentalidade entre ambos os casos.

Conforme Epstein (1977, p. 295-305) o mercado acionário é um fenômeno de tomada de decisão gigantesco cujos jogadores são denominados eufemisticamente de **investidores**. O autor compara o mercado acionário à aposta em cavalos. Uma característica peculiar do mercado acionário está na duração extremamente variável de um jogo completo e na exata determinação do retorno durante um longo período de jogo, devido à variação do poder de compra do dinheiro, taxas de impostos, disponibilidade de mercadorias, invenções, guerras, mudanças de governos, reviravoltas políticas, deslocamentos em massa, que influenciam psicologicamente o comportamento dos investidores. Refletindo um amálgama de fatores econômicos, monetários e psicológicos, o mercado acionário representa o mais sutilmente intrincado jogo inventado pelo homem. Nas páginas 287 a 294, Epstein relata que a mais antiga descrição de uma corrida de cavalos aparece no livro XXIII da *Ilíada* de Homero, e que ela parece remontar do povo hitita de 1500 a.C. O autor aborda matematicamente o problema da aposta em corridas de cavalo e culmina com a aplicação do princípio da máxima entropia da Teoria da Informação.

Baxter & Rennie (1998) iniciam o livro discutindo o problema da aposta em corrida de cavalos, para discutir os elementos essenciais da formação de preços racionais de derivativos. Este exemplo foi adaptado nesta tese, substituindo a aposta em corrida de cavalos por uma aposta em partida de futebol (final de um campeonato entre o Corinthians e o São Paulo).

Bekman & Costa Neto (1980) também utilizam a partida de futebol (final de uma copa do mundo entre Brasil e Argentina) como situação decisória enfrentada

por um torcedor brasileiro chamado João, para exposição dos principais conceitos de análise da decisão.

Do ponto de vista psicológico, Zuckerman *et al.* (1980) relatam que parece haver basicamente dois tipos de traços psicológicos nos seres humanos: buscadores e os não-buscadores de sensação. Esta distinção explicaria uma série de comportamentos dos indivíduos. O indivíduo buscador de sensação, do tipo que persegue prazeres hedonisticamente por meio de atividades de extroversão, incluindo beber socialmente, festas e atividades sexuais, tende a estar mais interessado em jogar e é mais sensível às probabilidades e aos valores envolvidos nos jogos.

Cover & Thomas (1991, p. 131) afirmam que, quando o mercado acionário funciona como o mercado de apostas em corridas de cavalos, então há um aumento da variação da taxa de duplicação que é igual à entropia mútua entre a informação privilegiada e a aposta na corrida de cavalos.

Finalmente, Cox & Rubinstein (1985, cap. 8) utilizam quase um capítulo, mostrando que o mercado de opções difere de um simples mercado de apostas em corridas de cavalos.♦

A seguir será exposto um exemplo paradigmático para que o conceito de incerteza e risco possa ser mais bem distinguido.

2.1.3 Hipóteses normalmente utilizadas para determinação do preço racional de opção européia simples ordinária

Serão definidas as hipóteses normalmente utilizadas para a determinação do preço racional de opção européia simples ordinária.

2.1.3.1 Hipótese de mercado perfeito

"*Veja - A ocasião faz o ladrão?*

James Wygand, Presidente da Control Risks *- As pessoas têm a noção errada de que um sujeito que entra para o crime faz isso porque não teve oportunidades na vida ou não achou emprego honesto. Pode ser que existam alguns criminosos por necessidade, mas esses são ladrões de pequenos crimes, que roubam para sobreviver no dia-a-dia. A verdade é que a maioria dos bandidos perigosos opta pela atividade criminosa porque sabe que ela é muito rentável, não paga*

*impostos, tem uma série de vantagens. Ou seja, é uma questão econômica em que os riscos (*sic*) do negócio são menores que os lucros, em que o criminoso pode ganhar mais do que sendo honesto."*

Entrevista à Revista *Veja*: Predadores à solta. 17 de novembro de 1999, p. 15.

"*Num livro de 1988, o professor Robert Klitgaard, de Harvard, definiu um grande problema em uma simples equação: corrupção = monopólio + arbitrariedade - transparência. [...]. Para* Transparency International*, uma ONG dedicada ao assunto, estamos na 45ª posição entre os 99 países, com uma pontuação de 4,1 (numa escala de 0 a 10), empatados com Zimbábue, Marrocos e Malaui. [...] Segundo estimativas de Ades e Di Tella, valores entre 16% e 44% dos investimentos viabilizados por políticas industriais ativas (internacionais) em sua amostra se transformam em propina: 30% em média!"*

Gustavo Franco. Corrupção e lubrificação. *Veja*: Em Foco. 17 de novembro de 1999, p. 224.

Mercado perfeito é um mercado sem custos de transação e leilões; nele todos os acordos são cumpridos; há possibilidade de comprar/vender qualquer montante de cada valor mobiliário [Luenberger (1998, cap. 9)]; as transações ocorrem continuamente e há a possibilidade da venda a descoberto (*short-selling*) ilimitada [Yang (1997, p. 12)]; há ausência de impostos; a liquidação é instantânea, a transação ocorre à vista (sem parcelamento) e existe um banco ideal constante. No caso das ações, não se considera o dividendo e dos *bonds*, o cupom.

Esta hipótese é análoga à da ausência de atrito da Mecânica Clássica. Desta forma, alguns autores denominam esta hipótese como **hipótese da ausência de atrito**.

Custos de transação (*transaction costs*) são custos da compra ou venda de um valor mobiliário, que consistem principalmente na comissão de corretagem, margem do investidor ou de uma taxa (como seria a cobrada por um banco ou por uma corretora para negociar títulos do governo), mas inclui também tributos diretos tais como a comissão da SEC nos E.U.A., quaisquer impostos de transferência determinados pelo governo e outros impostos diretos.

2.1.3.1.1 Banco ideal constante

Num banco ideal, conforme Luenberger (1998, p. 19) e Merton (1973, p. 162), as taxas juros de depósito e empréstimo são iguais e não há taxas de serviços e de

transações. As taxas de juros também independem do montante do principal. Esta definição não implica que as taxas de juros para todas as transações devam ser idênticas.

Se um banco ideal tem uma taxa de juros que independe do prazo para o qual ela se aplica e a taxa de juros é composta de acordo com as regras usuais, diz-se que ele é um banco ideal constante.

Ao longo do texto, admitir-se-á que as taxas de juros são constantes.

Portanto, pode-se supor a existência de ativo sem incerteza (obrigação) *b(t)* cujos preços seguem um processo

$$\begin{aligned} db(t) &= b(t) r_f dt \\ \frac{db(t)/dt}{b(t)} &= r_f \\ b(t) &= b(0) \exp\left(r_f t\right) \end{aligned}$$

sendo r_f a taxa de juros certa nominal (composta continuamente) constante. Sua unidade é o inverso do período (ano, mês, dia útil etc.).

r_f será chamada de **preço do tempo**.

2.1.3.2 Hipótese de mercado competitivo, livre e aberto como vários investidores igualmente importantes

A partir deste momento, o decisor racional, isto é, o que segue os princípios anteriores, será chamado de **investidor**.

No mercado financeiro será suposta a existência de vários investidores igualmente importantes.

Para haver a possibilidade de arbitragem no mercado financeiro, é necessária a existência de pelo menos um decisor que use os princípios de comparação e da insaciabilidade e pelo menos dois valores mobiliários de propriedades idênticas com preços diferentes. Logo, os modelos de determinação dos preços de valores mobiliários e derivativos não devem possibilitar a arbitragem.

Mercado competitivo é um mercado no qual compradores e vendedores individualmente têm pouca ou nenhuma condição de afetar os preços. A moderna

teoria de Finanças está baseada no paradigma de mercado competitivo (Duffie (1992) e Jarrow & Turnbull (1966)). Este paradigma possui duas hipóteses. A primeira é que os mercados de valores mobiliários são perfeitamente elásticos, isto é, os investidores atuam como tomadores de preços. A segunda é que a compra/venda tem liquidação imediata. Ambas as hipóteses são aproximações que não são satisfeitas mesmo nos mercados com maior liquidez. A ausência destas condições é algumas vezes chamada de **risco de liquidez**.

Mercado livre e aberto (*free and open market*) é um mercado em que o preço é determinado pela oferta e procura de forma livre e sem regulamentações. O oposto é um mercado controlado, no qual oferta, procura e preço são artificialmente estabelecidos, resultando em um mercado ineficiente.

Pindyck & Rubinfeld (1994, cap. 16) mostram que, num mercado competitivo, livre, aberto e perfeito, com vários investidores, os preços dos ativos são únicos, pois o mercado encontra-se em equilíbrio[11] e possui a propriedade de ser Pareto-eficiente[12]. No mercado financeiro há ainda a facilidade de estes preços serem de fácil acesso. Logo, os preços à vista dos valores mobiliários são teoricamente únicos e de equilíbrio, e o valor observado no mercado é aproximadamente igual a este valor em mercados líquidos.

---

[11] Um equilíbrio é um conjunto de preços pelos quais a quantidade demandada se torna igual à quantidade ofertada em cada um dos mercados. Isto é também um equilíbrio competitivo, pois todos os vendedores e compradores serão estabelecedores de preço. A alocação de valores mobiliários em equilíbrio competitivo é economicamente eficiente. O equilíbrio competitivo é freqüentemente preferido, pois consegue alocar os recursos com um mínimo de informação, isto é, com os preços disponíveis no mercado.

[12] Uma alocação tem eficiência de Pareto quando não pode ser realocada para tornar maior o bem-estar de uma pessoa sem que haja diminuição do bem-estar de outra.

Negociabilidade (*marketability*) é a rapidez e facilidade com que determinado valor mobiliário pode ser comprado e vendido. Uma grande quantidade de ações em circulação, que sejam negociadas ativamente, é um fato que caracteriza a liquidez de um título. Em linguagem corriqueira, o termo **negociabilidade** pode ser usado no lugar de **liquidez** (*liquidity*), mas liquidez implica a preservação do valor quando um valor mobiliário é comprado ou vendido.

2.1.3.3 Hipótese do equilíbrio competitivo do mercado de capitais

Pindyck & Rubinfeld (1994, cap. 16) afirmam que, num mercado competitivo, livre, aberto e perfeito com vários investidores, os preços dos ativos são únicos, pois o mercado encontra-se em equilíbrio e possui a propriedade de ser Pareto-eficiente.

No mercado financeiro há ainda a facilidade de esses preços serem de fácil acesso.

Logo, os preços à vista dos valores mobiliários são teoricamente únicos e de equilíbrio e o valor observado no mercado é aproximadamente igual a esse valor em mercados líquidos.

Lisboa (1999) estuda as relações entre equilíbrio, arbitragem e mercado completo e incompleto.

2.1.3.4 Hipótese de mercado completo

"Significação das teorias físicas – Existem muitas pessoas que se espantam ao ver quão efêmeras são as teorias científicas. Essas pessoas vêem teorias serem sucessivamente abandonadas depois de alguns anos de prosperidade; vêem ruínas se acumulando sobre ruínas; prevêem que teorias hoje na moda sucumbirão, por sua vez, dentro em pouco, e concluem que elas são interiamente vãs. É o que chamam a falência da ciência. [...] O objetivo das teorias matemáticas não é nos revelar a verdadeira natureza das coisas – o que não seria uma pretensão razoável – e sim coordenar as leis físicas que a experiência nos revela mas que, sem o auxílio da matemática, não poderíamos sequer enunciar. Pouco nos importa que o éter exista realmente: é um problema para os metafísicos (filósofos). O importante para nós é que tudo se passa como se ele existisse, e essa é uma hipótese cômoda para a explicação de fenômenos. Afinal, temos outras razões para crer na existência de objetos materiais? Essa também é uma hipótese cômoda e que nunca deixará de o ser, ao passo que um dia, certamente, o éter será rejeitado, por inútil."

Poincaré (1984, p. 127 e 157)

Se os valores mobiliários (ativos financeiros) forem definidos como um conjunto finito de estados futuros de preços e se houver pelo menos a mesma quantidade de ativos financeiros linearmente independentes, então o mercado é denominado completo uniperíodo (adaptado de Luenberger (1998, cap. 9)).

Conforme Cox & Rubinstein (1985, cap. 19), sempre que o número de ativos financeiros linearmente independentes for igual ao número de estados da natureza, qualquer padrão de retornos pode ser reproduzido e o mercado, nesta situação, chama-se **completo**. Ele é **completo** porque nenhum outro ativo financeiro com padrão de retorno diferente, i.e., que não seja mera combinação linear dos retornos dos ativos anteriores, pode ser gerado.

Um mercado é dito **dinamicamente completo** ou **completo multiperíodo** se todo possível resultado de um derivativo pode ser replicado para cada instante de tempo seguinte por uma carteira autofinanciável de valores mobiliários negociáveis [Baxter & Rennie (1998, p. 197)].

Um dos objetivos da tese é dispensar o uso desta hipótese para a determinação do preço racional da opção européia simples ordinária sobre ação e *bond.*

Gulko (1998), Samperi (1998) e Lisboa (1999) discutem as relações entre eficiência, equilíbrio e mercado completo e incompleto.

2.1.3.5 Hipótese de mercado eficiente

Mercado eficiente é um mercado no qual toda a informação se reflete instantaneamente nos preços dos valores mobiliários. É uma hipótese segundo a qual os preços de mercado refletem o conhecimento e as expectativas de todos os investidores. Seus adeptos consideram inútil procurar ações subavaliadas ou fazer previsões de movimentos de mercado. Qualquer novo desenvolvimento se reflete no preço das ações de uma companhia, afirmam, tornando impossível vencer o mercado. Essa mesma teoria, objeto de ruidosas disputas, afirma que um investidor que arremessasse dardos na lista de ações de um jornal teria as mesmas chances de desempenho no mercado que qualquer investidor profissional.

Uma questão importante em Finanças é se o mercado de ações é eficiente. Conforme Brealey & Myers (1998, p. 324-6), "como muitas vezes acontece com idéias importantes, o conceito de mercado eficiente foi conseqüência de uma descoberta fortuita. Em 1953, a Royal Statistical Society reuniu-se em Londres para discutir um documento bastante original[13]. Seu autor, Maurice Kendall, era um estatístico distinto, e o tema era o comportamento dos preços das acções e das mercadorias. Kendall tinha procurado identificar ciclos regulares de preços, mas, para sua surpresa, não os encontrou. Cada série parecia ser "errática, quase como se, uma vez por semana, o Démonio do Azar (*Demon of Chance*) tirasse um número aleatório...e o acrescentasse ao preço corrente para determinar o preço da semana seguinte". Por outras palavras, os preços pareciam seguir um *passeio aleatório*. [...] Quando Maurice Kendall sugeriu que os preços das acções seguem um passeio aleatório, estava a admitir que as variações dos preços são tão independentes umas das outras como os ganhos e as perdas do nosso jogo[14]. Para a maioria dos economistas, esta idéia era bizarra e alarmante. Na realidade, a idéia *não* era completamente nova. Já tinha sido proposta numa tese de doutoramento quase esquecida, escrita 53 anos antes por um francês, Louis Bachelier[15]. A sugestão de Bachelier era suficientemente original, mas a teoria matemática dos processos aleatórios que a acompanhava antecipou-se em cerca de 5 anos ao famoso trabalho de Einstein sobre o movimento browniano aleatório das moléculas de gás em colisão. O trabalho de Kendall não caiu no esquecimento como o de Bachelier. Com a generalização do uso dos computadores e o aumento

---

[13] Kendall, M.G. 1953. "The analysis of economic time series, part I. Prices." *Journal of Royal Statistical Society,* 96:11-25.

[14] Jogo definido anteriormente cujos resultados estavam baseados em lançamentos independentes de uma moeda.

[15] Bachelier, L. 1990. *Théorie de la speculation.* Gautheir-Villars, Paris. Reeditada em inglês (A. J. Bones, trad.) *in* P.H. Cootner (ed.), *Random character of stock market prices,* M.I.T. Press. Cambridge, Mass., 1964, p. 17-78.

dos bancos de dados, os economistas e os estatísticos conseguiram acumular rapidamente uma grande quantidade de informações comprovativas".

Nos anos 30, o economista especializado no setor alimentício Holbrook Working também reparou no comportamento aleatório dos preços das mercadorias (*commodities*)[16].

O mercado é eficiente se o conhecimento das mudanças passadas do preço da ação não informa nada sobre as futuras mudanças no preço desta ação [Winston (1998, p. 73-78)]. Se os preços sempre refletem a informação total relevante para os investidores, então eles irão mudar apenas quando surge informação nova. Mas informação nova, por definição, não pode ser predita, caso contrário ela não seria nova. De outra maneira, se o preço da ação reflete tudo o que é previsível, então a variação de preço da ação deve refletir apenas o que é imprevisível. Portanto, a série de variações de preços deve ser aleatória [Brealey & Myers (1998, p. 323-347)].

A versão contínua do passeio aleatório (*random walk*) é o movimento browniano. Adota-se, normalmente, que o preço do valor mobiliário subjacente segue um movimento browniano geométrico:

$$dS(t) = S(t)[\mu dt + \sigma dZ(t)]$$

sendo $dZ(t)$ possui distribuição normal com média zero e variância dt, μ e σ são constantes positivas.

Note-se que a hipótese de mercado eficiente conduz à idéia de passeio aleatório. Esta, por sua vez, leva ao movimento browniano. Isto é, há uma formalização deste princípio por meio de uma equação diferencial estocástica. Evidentemente, há outras equações que modelam situações mais específicas.

Há pelo menos duas teorias reacionárias. Uma delas é totalmente determinística e é conhecida pelo nome de *Elliott Wave Principle* de Prechter & Frost (1995).

---

[16] Veja H. Working. 1934. "A random difference series for use in the analysis of time series" *Journal of American Statistical Association*, 29:11-24.

Esta teoria baseia-se fundamentalmente na seção áurea [Ricieri (1994)]. A outra é uma combinação ou meio termo entre a abordagem estocástica (passeio aleatório) e a determinística. Ela é chamada de **hipótese de mercado fractal** e está relacionada com a teoria do caos (Peters (1994), Brown (1995), Gleiser (1994), Ricieri (1990) e Trippi (1995)). Em Peters (1994) há o relato de uma fórmula de formação do preço da opção européia simples ordinária utilizando a hipótese de mercado fractal.

Essa hipótese é normalmente usada para que se possa utilizar o modelo do movimento browniano geométrico para descrever o comportamento do preço da ação. Como a abordagem entrópica não necessita da informação sobre a distribuição real ou subjetiva do preço da ação para a determinação do preço racional da opção européia simples ordinária, então esta hipótese não será necessária.

2.1.3.6 Hipótese da atitude do decisor diante da incerteza

Rubinstein (1976), *apud* Trigeorgis (1998, cap. 1), e Luenberger (1998) derivam as fórmulas tradicionais de formação de preço de opções na condição de aversão à incerteza.

Luenberger (1998, cap. 9) afirma que utilizou a condição de aversão à incerteza por esta abordagem "parecer ser a mais útil" mas o preço racional independe da escolha da função de utilidade do investidor racional. O fato de Luenberger escolher uma abordagem que é útil mas desnecessária contraria o princípio da parcimônia.

Para Jaynes (1996, p. 19-22) a utilidade do dinheiro permanece mal definida. Smidts (1997) discute as diferenças e semelhanças entre os modelos de utilidade de Bernoulli (1738) e de von Neumann & Morgenstern (1947).

No entanto, com ativos em condição de incerteza técnica, a consideração da atitude do decisor diante da incerteza pode ser relevante (Herath (1998)). Logo, a Teoria Estatística Bayesiana da Decisão torna-se adequada para determinar o preço racional da opção. Contudo, para a aplicação desta abordagem, o decisor necessita ter coerência ou racionalidade no sentido de que os axiomas sejam

respeitados [Bekman & Costa Neto (1980, p. 56-7), Lee (1996, p. 13-4), Howard (1996) e Bierman *et al.* (1991)].

Pela construção do mundo com preço da incerteza neutralizado a seguir elimina-se matematicamente o preço da incerteza dos valores mobiliários do mercado financeiro. Neste mundo virtual, os investidores são considerados contingencialmente como neutros ao preço da incerteza. Porém, no mundo real eles têm alguma atitude diante da incerteza. Desta forma, com a consideração deste mundo hipotético, a atitude diante da incerteza de cada indivíduo é irrelevante para a determinação do preço racional dos valores mobiliários do mercado financeiro.

2.1.3.7 Hipótese do mundo com preço da incerteza neutralizado

"You really know where you are.
For the first time in history.
He quoted the planetary motto.
'Community, Identity, Stability.'
Grand words."

Aldous Huxley. 1989. *Brave New World*. New York: HarperPerennial.

Inicialmente temos que o mercado de capitais é um mundo no qual os valores mobiliários possuem preço da incerteza positivo para os investidores racionais que são considerados avessos à incerteza. Isto é, no mercado financeiro real, os investidores são avessos à incerteza e, portanto, há preço da incerteza positivo para os valores mobiliários.

A cada um dos ativos está associado um preço da incerteza positivo. Quanto mais incerto for um valor mobiliário, maior deverá ser esse preço.

Para um valor mobiliário sem incerteza o preço deverá ser nulo. Este ativo será denominado **valor mobiliário livre de incerteza**.

Em termos teóricos, este valor mobiliário é contrato de empréstimo ou depósito num banco ideal constante.

Conforme Luenberger (1998, p. 40), a taxa de juros livre de incerteza ou taxa de juros certa é o preço ou valor da locação da *commodity* mais popular: o dinheiro.

Um exemplo concreto é o retorno das letras do tesouro norte-americano (*U.S. Treasury bills*). É o rendimento (*yield*) sobre um investimento sem risco. As letras do Tesouro americano com vencimento em 3 meses são consideradas investimento sem incerteza, uma vez que são obrigações diretas do governo norte-americano e o prazo é suficientemente curto para minimizar os riscos de inflação e alterações nas taxas de juros.

O modelo de determinação do preço dos ativos (*capital asset pricing model-CAPM*) apresenta a premissa de que o retorno sobre um valor mobiliário é igual ao retorno sobre um investimento certo (preço do tempo), acrescido de um preço da incerteza.

Taxa de capitalização ou de desconto (*capitalization or discount rate*) é a taxa utilizada para comparar o valor, numa moeda, que será recebido no futuro, com o valor na mesma moeda recebido hoje. É a taxa de juros usada para converter uma série de fluxos de pagamentos futuros em um só valor presente. É também a taxa de juros que a Reserva Federal (E.U.A.) exige dos bancos membros para empréstimos, utilizando-se como garantia, em tais empréstimos, títulos do governo ou títulos aceitáveis para redesconto (*eligible paper*). Configura um referencial mínimo, uma vez que os bancos estabelecem suas taxas de empréstimos em percentuais sempre acima da taxa de desconto.

Taxa de desconto ajustada à incerteza (*risk-adjusted discount rate*) é a taxa necessária para determinar o valor presente de um fluxo de renda incerto; é a taxa certa (geralmente o retorno sobre títulos do Tesouro dos E.U.A.), acrescida de um preço da incerteza, que se baseia em uma análise das características de risco de um determinado investimento ou projeto.

Pode-se pensar que $E_{\mathbb{P}_i}[S_i(\tau)] > \left(1 + r_f\right)s_i(0)$, pois, caso contrário, não haveria motivo para ter o valor mobiliário com fluxos de caixa incertos [Neftci (1996, p. 280)].

Então, uma forma de expressar o preço de valor mobiliário em função do preço da incerteza é a seguinte:

$$
\begin{aligned}
s_i(0) &= vp_{r_i}\left\{E_{\mathbb{P}_i}\left[S_i(\tau)\right]\right\} \\
&= \frac{1}{1+r_i}E_{\mathbb{P}_i}\left[S_i(\tau)\right] \\
&= \frac{1}{1+r_f+r_{p_i}}E_{\mathbb{P}_i}\left[S_i(\tau)\right]
\end{aligned}
$$

Isto é, o preço $s_i(0)$ de um valor mobiliário i hoje (data 0) é o valor presente vp_{r_i} da esperança matemática do preço numa data futura $E_{\mathbb{P}_i}\left[S_i(\tau)\right]$, na medida de probabilidade real ou subjetiva $\mathbb{P}_i$. Note-se que a taxa de desconto ajustada à incerteza r_i do valor mobiliário i pode ser decomposta na taxa de desconto certa r_f (pela hipótese de mercado perfeito e mais especificamente pela suposição da existência do banco ideal constante) (preço do tempo) e no preço da incerteza específico do ativo i.

Outra forma de expressar o mundo com preço da incerteza é em termos do retorno que um valor mobiliário produz.

Retorno é o fluxo monetário total que um ativo produz, como fração de seu preço.

$$
\begin{aligned}
E_{\mathbb{P}_i}\left[\frac{S_i(\tau)}{s_i(0)}\right] &= 1+r_i \\
&= 1+r_f+r_{p_i} \\
E_{\mathbb{P}_i}\left[1+R_i(\tau)\right] &= 1+r_f+r_{p_i} \\
E_{\mathbb{P}_i}\left[R_i(\tau)\right] &= r_f+r_{p_i}
\end{aligned}
$$

Portanto, a esperança do retorno $\frac{S_i(\tau)}{s_i(0)}$ na medida de probabilidade real ou subjetiva $\mathbb{P}_i$ é um mais a taxa de desconto ajustada à incerteza do valor mobiliário i.

Observa-se também que a esperança, nesta mesma medida de probabilidade da taxa de retorno, é igual à soma do preço do tempo com o preço da incerteza específico.

Sob a hipótese de mercado perfeito, a taxa certa r_f é única e observável.

Sob a hipótese de mercado competitivo, livre e aberto com vários investidores (mercado em equilíbrio competitivo), o preço de um valor mobiliário $s_i(0)$ é único e observável.

No entanto, o preço da incerteza e a medida de probabilidade do preço do valor mobiliário são não observáveis e não necessariamente únicos.

Além disso, se for exigido que o mercado seja livre de arbitragem, a linearidade dos preços é obrigatória e vice-versa [Luenberger (1998, p. 188-90)]. Em outras palavras, conforme Luenberger (1998, p. 189), o preço da soma de dois ativos é a soma de seus preços, e o preço de um múltiplo do ativo é o múltiplo do preço. Desta forma, sendo a esperança matemática um operador linear, ela torna-se séria candidata a determinador de preço. Portanto, o valor presente da esperança de uma combinação linear dos preços futuros dos valores mobiliários deve ser igual à combinação linear dos preços atuais dos respectivos valores mobiliários.

A linearidade dos preços também implica necessariamente, conforme Luenberger (1998, p. 190), a lei do preço único. Isto é, no mercado livre de arbitragem o preço de um valor mobiliário deve ser único.

Imagine-se que haja dois ativos não idênticos com risco no mercado de capitais perfeito e em equilíbrio competitivo.

Portanto, o preço de cada um deles é:

$$s_1(0) = \frac{1}{1 + r_f + r_{p_1}} E_{\mathbb{P}_1}\left[S_1(\tau)\right]$$
$$s_2(0) = \frac{1}{1 + r_f + r_{p_2}} E_{\mathbb{P}_2}\left[S_2(\tau)\right]$$

O preço da soma dos dois ativos é:

$$\begin{aligned} s_{1+2}(0) &= \frac{1}{1 + r_f + r_{p_{1+2}}} E_{\mathbb{P}_{1+2}}\left[S_1(\tau) + S_2(\tau)\right] \\ &= \frac{1}{1 + r_f + r_{p_{1+2}}} \left\{ E_{\mathbb{P}_{1+2}}\left[S_1(\tau)\right] + E_{\mathbb{P}_{1+2}}\left[S_2(\tau)\right] \right\} \\ &= \frac{1}{1 + r_f + r_{p_{1+2}}} E_{\mathbb{P}_{1+2}}\left[S_1(\tau)\right] + \frac{1}{1 + r_f + r_{p_{1+2}}} E_{\mathbb{P}_{1+2}}\left[S_2(\tau)\right] \end{aligned}$$

Observe-se que, mesmo que as variáveis aleatórias sejam dependentes, as equações anteriores valem.

Note-se que seria interessante obter o seguinte resultado para que o mercado também fosse livre de arbitragem:

$$
\begin{aligned}
s_{1+2}(0) &= s_1(0) + s_2(0) \\
&= \frac{1}{1 + r_f + r_{p_1}} E_{\mathbb{P}_1}\left[S_1(\tau)\right] + \frac{1}{1 + r_f + r_{p_2}} E_{\mathbb{P}_2}\left[S_2(\tau)\right]
\end{aligned}
$$

Portanto, temos que:

$$
\frac{1}{1 + r_f + r_{p_{1+2}}} E_{\mathbb{P}_{1+2}}\left[S_1(\tau)\right] + \frac{1}{1 + r_f + r_{p_{1+2}}} E_{\mathbb{P}_{1+2}}\left[S_2(\tau)\right] = \frac{1}{1 + r_f + r_{p_1}} E_{\mathbb{P}_1}\left[S_1(\tau)\right] + \frac{1}{1 + r_f + r_{p_2}} E_{\mathbb{P}_2}\left[S_2(\tau)\right]
$$

Percebe-se que o problema fica cada vez mais complicado, pois novas grandezas não observáveis e não únicas (subjetivas) são introduzidas.

Uma forma de resolver este problema é supor a validade do modelo de determinação do preço de valores mobiliários, isto é, do CAPM. Este modelo foi desenvolvido entre 1964 e 1968 [Sharpe & Alexander (1990, p. 194)].

O CAPM (*capital asset pricing model*) é um modelo de determinação de preço de valores mobiliários que relaciona o risco esperado com o retorno esperado. Baseia-se na teoria de que os investidores exigem maiores retornos para maiores riscos e estabelece que o retorno de um ativo ou de um valor mobiliário é igual ao dos investimentos livres de risco - como, por exemplo, o retorno de uma obrigação de Tesouro de curto prazo -, acrescido de um preço da incerteza.

Conforme Sharpe & Alexander (1990, p. 195-6), o CAPM está apoiado em algumas hipóteses, como mercado perfeito, decisor insaciável e avesso à incerteza, todos os investidores possuem o mesmo horizonte de investimento de um período, a taxa certa é a mesma para todos os investidores, informação é disponibilizada gratuitamente e instantaneamente para todos os investidores, investidores têm expectativas homogêneas (eles têm as mesmas percepções dos retornos esperados, das variâncias e das covariâncias entre os retornos dos valores mobiliários).

Supondo a validade do CAPM, conforme Luenberger (1998, p. 187), o preço dos dois valores mobiliários pode ser escrito da seguinte maneira:

$$
\begin{aligned}
s_1(0) &= \frac{1}{1 + r_f + \beta_1(r_m - r_f)} E_{\mathbb{P}_1}\left[S_1(\tau)\right] \\
s_2(0) &= \frac{1}{1 + r_f + \beta_2(r_m - r_f)} E_{\mathbb{P}_2}\left[S_2(\tau)\right]
\end{aligned}
$$

sendo $\beta_i = \dfrac{Cov\left[\dfrac{S_i(\tau) - s_i(0)}{s_i(0)}; R_m\right]}{Var[R_m]}$, R_m a taxa de retorno da carteira eficiente de mercado e $r_m = E_{\mathbb{M}}[R_m]$.

No entanto, há uma forma alternativa mais interessante, denominada **forma do equivalente certo**:

$$s_1(0) = \frac{1}{1+r_f}\left\{E_{\mathbb{P}_1}[S_1(\tau)] - \frac{Cov[S_1(\tau); R_m](r_m - r_f)}{Var[R_m]}\right\}$$
$$s_2(0) = \frac{1}{1+r_f}\left\{E_{\mathbb{P}_2}[S_2(\tau)] - \frac{Cov[S_2(\tau); R_m](r_m - r_f)}{Var[R_m]}\right\}$$
$$s_{1+2}(0) = \frac{1}{1+r_f}\left\{E_{\mathbb{P}_{1+2}}[S_1(\tau) + S_2(\tau)] - \frac{Cov[S_1(\tau) + S_2(\tau); R_m](r_m - r_f)}{Var[R_m]}\right\}$$

Como $Cov[X+Y;Z] = Cov[X;Z] + Cov[Y;Z]$ para quaisquer variáveis aleatórias, então:

$$\begin{aligned} s_{1+2}(0) &= \frac{1}{1+r_f}\left\{E_{\mathbb{P}_{1+2}}[S_1(\tau) + S_2(\tau)] - \frac{Cov[S_1(\tau) + S_2(\tau); R_m](r_m - r_f)}{Var[R_m]}\right\} \\ &= \frac{1}{1+r_f}\left\{E_{\mathbb{P}_1}[S_1(\tau)] - \frac{Cov[S_1(\tau); R_m](r_m - r_f)}{Var[R_m]}\right\} + \frac{1}{1+r_f}\left\{E_{\mathbb{P}_2}[S_2(\tau)] - \frac{Cov[S_2(\tau); R_m](r_m - r_f)}{Var[R_m]}\right\} \\ &= s_1(0) + s_2(0) \end{aligned}$$

Logo, o CAPM respeita a condição de linearidade dos preços no mercado de capitais. No entanto, ele depende ainda de grandezas não observáveis (mas estimáveis) e não únicas, como as medidas de probabilidade dos preços e dos betas dos valores mobiliários.

Esta forma do equivalente certo nos fornece uma expressão importante do retorno dos ativos:

$$\begin{aligned} \frac{E_{\mathbb{P}_1}[S_1(\tau)] - \dfrac{Cov[S_1(\tau); R_m](r_m - r_f)}{Var[R_m]}}{s_1(0)} &= 1 + r_f \\ \frac{E_{\mathbb{P}_1}[S_1(\tau)] - \beta_1 s_1(0)(r_m - r_f)}{s_1(0)} &= 1 + r_f \\ E_{\mathbb{P}_1}\left[\frac{S_1(\tau)}{s_1(0)}\right] - \beta_1(r_m - r_f) &= 1 + r_f \\ E_{\mathbb{P}_1}\left[\frac{S_1(\tau)}{s_1(0)}\right] &= 1 + r_f + \beta_1(r_m - r_f) \\ E_{\mathbb{P}_1}\left[\frac{S_1(\tau)}{s_1(0)}\right] &= 1 + r_f + r_{p_1} \end{aligned}$$

Seria, então, muito bom, conforme Neftci (1996, p. 22-6 e cap. 14), se fosse possível expressar os retornos dos ativos com incerteza da seguinte forma:

$$E_{\mathbb{Q}}\left[\frac{S_i(\tau)}{s_i(0)}\right] = 1 + r_f$$
$$i = 1,2,\ldots,a$$

Se isso pudesse ser feito, haveria linearidade dos preços, portanto, um mercado livre de arbitragem e preço único para os valores mobiliários e suas combinações lineares.

Dessa forma, a existência e unicidade de uma medida de probabilidade $\mathbb{Q}$ origina o que se chamará de **mundo com preço da incerteza neutralizado**, pois todos os valores mobiliários possuem, como retorno, o retorno certo. A medida de probabilidade $\mathbb{Q}$ é chamada de **medida de probabilidade neutralizadora do preço da incerteza**.

Desta forma, o preço de um valor mobiliário possui a seguinte expressão no mundo com preço da incerteza neutralizado:

$$s_i(0) = \frac{1}{1+r_f} E_{\mathbb{Q}}\left[S_i(\tau)\right]$$
$$i = 1,2,\ldots,a$$

Note-se que, se ela existir e for única, então uma combinação linear de valores mobiliários tem seu preço determinado da seguinte maneira:

$$\begin{aligned}\frac{1}{1+r_f} E_{\mathbb{Q}}\left[\sum_{i=1}^{a} \alpha_i S_i(\tau)\right] &= \frac{1}{1+r_f}\sum_{i=1}^{a} \alpha_i E_{\mathbb{Q}}\left[S_i(\tau)\right] \\ &= \sum_{i=1}^{a} \alpha_i s_i(0)\end{aligned}$$

Logo, o problema de determinação de preço de valores mobiliários reduz-se à determinação de uma única medida de probabilidade neutralizadora do preço da incerteza.

Desta maneira, conforme Neftci (1996, cap. 14), a medida de probabilidade neutralizadora do preço da incerteza é chamada de **medida de probabilidade equivalente** e não depende da medida de probabilidade real ou subjetiva $\mathbb{P}_i$. Esta transformação permite, conforme Neftci (1996, p. 276-7), tratar um valor

mobiliário que carrega um preço da incerteza positivo como se ele fosse um valor mobiliário sem incerteza.

2.1.3.7.1 Mudança de medida de probabilidade

"- Adeus, disse a raposa. Eis o meu segredo.
É muito simples: só se vê bem com o coração.
O essencial é invisível para os olhos.

-O essencial é invisível para os olhos,
repetiu o principezinho, a fim de se lembrar."

Antoine de Saint-Exupéry. *O Pequeno Príncipe*. 1999. 47ª ed. 3ª reimpressão, p. 72
Tradução de Dom Marcos Barbosa. R.J.: Agir.

Conforme Neftci (1996, p. 283), duas medidas de probabilidade são equivalentes se elas são atribuídas ao mesmo domínio (conjunto de apoio). Em Finanças, a transformação para uma medida equivalente permite tratar um ativo que carrega um preço da incerteza positivo como se fosse um ativo sem incerteza [Neftci (1996, p. 276-7)].

2.1.3.7.1.1 Exemplo de mudança de medida de probabilidade de variável aleatória discreta

O exemplo a seguir encontra-se em Neftci (1996, p. 278 e 282-3).

Seja Z uma variável aleatória discreta uniforme, definida sobre o experimento de lançamento de um dado de seis faces.

Evento	Probabilidade ($\mathbb{P}$)	*Valor*
w_1 =face 1 ou 2	$\Pr(w_1) = \frac{1}{3} \approx 0,33$	$z_1 = 10$
w_2 =face 3 ou 4	$\Pr(w_2) = \frac{1}{3} \approx 0,33$	$z_2 = -3$
w_3 =face 5 ou 6	$\Pr(w_3) = \frac{1}{3} \approx 0,33$	$z_3 = -1$

Tem-se que a esperança e a variância valem:

$$
\begin{aligned}
E_{\mathbb{P}}[Z] &= \frac{1}{3}[10] + \frac{1}{3}[-3] + \frac{1}{3}[-1] \\
&= 2 \\
Var_{\mathbb{P}}[Z] &= E_{\mathbb{P}}\left\{[Z - E_{\mathbb{P}}[Z]]^2\right\} \\
&= \frac{1}{3}[10-2]^2 + \frac{1}{3}[-3-2]^2 + \frac{1}{3}[-1-2]^2 \\
&= \frac{98}{3}
\end{aligned}
$$

Imagine-se que se deseja transformar esta variável aleatória de tal modo que:

- o domínio seja mantido;
- a variância seja mantida;
- e a esperança passe a valer 1, isto é, $E_{\mathbb{S}}[Z] = 1$.

Qual medida de probabilidade equivalente $\mathbb{S}$ satisfaz as restrições anteriores?

Evento	Probabilidade ($\mathbb{S}$)	*Valor*
w_1 =face 1 ou 2	$\Pr(w_1) = \frac{122}{429} \approx 0{,}28$	$z_1 = 10$
w_2 =face 3 ou 4	$\Pr(w_2) = \frac{22}{39} \approx 0{,}56$	$z_2 = -3$
w_3 =face 5 ou 6	$\Pr(w_3) = \frac{5}{33} \approx 0{,}15$	$z_3 = -1$

Portanto:

$$
\begin{aligned}
E_{\mathbb{S}}[Z] &= \frac{122}{429}[10] + \frac{22}{39}[-3] + \frac{5}{33}[-1] \\
&= 1 \\
Var_{\mathbb{S}}[Z] &= E_{\mathbb{S}}\left\{[Z - E_{\mathbb{S}}[Z]]^2\right\} \\
&= \frac{122}{429}[10-1]^2 + \frac{22}{39}[-3-1]^2 + \frac{5}{33}[-1-1]^2 \\
&= \frac{98}{3}
\end{aligned}
$$

Note-se que a nova medida de probabilidade independe da original.♦

Outro resultado interessante derivado do mundo com preço da incerteza neutralizado é o seguinte:

$$
\begin{aligned}
&E_{\mathbb{Q}}\left[S_i(\tau)\right] = \left(1 + r_f\right)s_i(0) \\
&i = 1, 2, \ldots, a
\end{aligned}
$$

Isto é, a expectativa na medida de probabilidade neutralizadora do preço da incerteza do preço futuro de um valor mobiliário é o valor futuro do preço à vista do valor mobiliário, utilizando-se a taxa certa do seu preço atual.

Se o conjunto de valores possíveis do preço futuro do valor mobiliário (domínio) for enumerável (quantidade contável de estados da natureza), então:

$$\begin{aligned} E_{\mathbb{Q}}\left[\frac{S_i(\tau)}{s_i(0)}\right] &= \sum_{j=1}^{\infty} q_j \frac{s_{ij}(\tau)}{s_i(0)} \\ &= \sum_{j=1}^{\infty} q_j \left(1 + r_{ij}\right) \\ &= 1 + r_f \\ i &= 1,2,\ldots,a \end{aligned}$$

Como neste mercado existe um valor mobiliário certo, então a soma da probabilidades é igual a um:

$$\begin{aligned} E_{\mathbb{Q}}\left[\frac{S_l(\tau)}{s_l(0)}\right] &= \sum_{j=1}^{\infty} q_j \frac{s_{lj}(\tau)}{s_l(0)} \\ &= \sum_{j=1}^{\infty} q_j \left(1 + r_{lj}\right) \\ &= \sum_{j=1}^{\infty} q_j \left(1 + r_f\right) \\ &= 1 + r_f \\ \sum_{j=1}^{\infty} q_j \left(1 + r_f\right) &= 1 + r_f \\ \sum_{j=1}^{\infty} q_j &= 1 \end{aligned}$$

Portanto, a existência de um valor mobiliário sem incerteza equivale à condição de a soma das probabilidades ser igual a um.

Outra forma muito utilizada é a que evidencia os preços de ativos Arrow-Debreu:

$$
\begin{aligned}
E_{\mathbb{Q}}\left[S_i(\tau)\right] &= \sum_{j=1}^{\infty} q_j s_{ij}(\tau) \\
&= \left(1 + r_f\right) s_i(0) \\
\sum_{j=1}^{\infty} q_j s_{ij}(\tau) &= \left(1 + r_f\right) s_i(0) \\
\sum_{j=1}^{\infty} \pi_j \left(1 + r_f\right) s_{ij}(\tau) &= \left(1 + r_f\right) s_i(0) \\
\sum_{j=1}^{\infty} \pi_j s_{ij}(\tau) &= s_i(0) \\
i &= 1,2,\ldots,a
\end{aligned}
$$

Portanto, os ativos Arrow-Debreu são os valores mobiliários elementares associados a cada estado da natureza. Isto é, $q_j = (1 + r_f)\pi_j$ é o preço de ativo Arrow-Debreu que paga uma unidade monetária se o estado da natureza j ocorrer, e zero, caso contrário. Pelas equações anteriores verifica-se que este valor é maior que zero, menor que um e sua soma é um mais a taxa de juros certa, isto é:

$$
\begin{aligned}
&0 < \pi_j < 1 \\
&\sum_{j=1}^{\infty} \pi_j = \sum_{j=1}^{\infty} \frac{q_j}{(1 + r_f)} = \frac{1}{1 + r_f}
\end{aligned}
$$

Então o mercado é livre de arbitragem se e somente se existirem os preços dos ativos Arrow-Debreu [Luenberger (1998, p. 249)].

Se o domínio $\xi \subset \mathbb{R}_*^+$ do preço futuro for não enumerável, então:

$$
\begin{aligned}
E_{\mathbb{Q}}\left[\frac{S_i(\tau)}{s_i(0)}\right] &= \int_{\xi} \frac{s_i(\tau)}{s_i(0)} q\left[s_i(\tau)\right] ds_i(\tau) \\
&= \int_{\xi} \left[1 + r_i(\tau)\right] q\left[s_i(\tau)\right] ds_i(\tau) \\
&= 1 + r_f \\
i &= 1,2,\ldots,a
\end{aligned}
$$

O preço racional de um valor mobiliário é:

$$
\begin{aligned}
s_i(0) &= \frac{1}{1 + r_f} E_{\mathbb{Q}}\left[S_i(\tau)\right] \\
&= \frac{1}{1 + r_f} \int_{\xi} s_i(\tau) q\left[s_i(\tau)\right] ds_i(\tau) \\
i &= 1,2,\ldots,a
\end{aligned}
$$

O preço racional de um ativo Arrow-Debreu é:

$$\pi(\tau) = \frac{q\left[s_i(\tau)\right]}{1 + r_f}$$
$$i = 1,2,\ldots,a$$

2.1.3.7.2 Processo estocástico martingale

Suponha-se que hoje seja t_0 e a data futura seja t_τ. Então o prazo será $t_\tau - t_0 = \tau$.

Desta forma, tem-se que:

$$\begin{aligned}
s_i(t_0) &= \exp\left(-r_f\tau\right)E_{\mathbb{Q}}^{t_0}\left[S_i(t_\tau)\right] \\
&= \exp\left[-r_f\left(t_\tau - t_0\right)\right]E_{\mathbb{Q}}^{t_0}\left[S_i(t_\tau)\right] \\
&= \exp\left(r_f t_0\right)\exp\left(-r_f t_\tau\right)E_{\mathbb{Q}}^{t_0}\left[S_i(t_\tau)\right] \\
\exp\left(-r_f t_0\right)s_i(t_0) &= \exp\left(-r_f t_\tau\right)E_{\mathbb{Q}}^{t_0}\left[S_i(t_\tau)\right] \\
\exp\left(-r_f t_0\right)s_i(t_0) &= E_{\mathbb{Q}}^{t_0}\left[\exp\left(-r_f t_\tau\right)S_i(t_\tau)\right] \\
\overline{s_i}(t_0) &= E_{\mathbb{Q}}^{t_0}\left[\overline{S_i}(t_\tau)\right]
\end{aligned}$$

Isto é, o processo estocástico $\overline{S_i}(t) = \exp\left(-r_f t\right)S_i(t)$ é chamado **martingale** [Luenberger (1998, p. 373)]. Este processo estocástico não apresenta tendência ou periodicidade.

Conforme Neftci (1996, p. 101), martingale é um instrumento central na moderna teoria de Finanças.

Portanto, $\overline{S_i}(t)$ é um martingale no mundo com preço da incerteza neutralizada.

Neste mundo virtual, conforme Neftci (1996, p. 103), a melhor estimativa do preço futuro é a última observação de $\overline{S_i}(t)$.

Neftci (1996, p. 103-4) expõe que martingales são variáveis aleatórias cujas variações futuras são completamente imprevisíveis mesmo que a série história do preço seja dada. Isto é, é impossível prever a direção do movimento futuro do preço num processo martingale.

No entanto, no mundo real com preço da incerteza, conforme Neftci (1996, p. 104-6), sabe-se que o preço da ação e do *bond* não são completamente imprevisíveis. Espera-se que o preço do *bond* descontado cresça com o passar do tempo. Em geral, o mesmo ocorre com o preço da ação. Espera-se que seus preços

cresçam em média com o passar do tempo. Portanto, estes processos estocásticos são chamados de **submartigales**.

Algo similar ocorre com os derivativos. Por exemplo, a opção tem o preço do tempo (valor do tempo) e, com o passar do tempo, o preço de uma opção européia declina, considerando-se todos os outros parâmetros fixos. Portanto, estes processos são chamados de **supermartingales**.

No mercado financeiro, conforme Neftci (1996, p. 115), não se pode esperar que o preço observado no mercado de um valor mobiliário com incerteza seja igual ao seu preço esperado, trazido a valor presente pela taxa de juros certa. Há um preço da incerteza. Portanto, qualquer preço de um ativo com incerteza, quando descontado pela taxa de juros certa, não será um martingale.

Então os preços dos valores mobiliários são mais provavelmente sub ou supermartingales.

Como foi visto anteriormente, a consideração da hipótese do mundo com preço da incerteza neutralizada transforma processos não-martingale em martingale por meio de uma substituição matemática da medida de probabilidade. Isto é:

$$E_{\mathbb{P}}^{t_0}\left[\exp\left(-r_f\tau\right)S_i(t_\tau)\right] > s_i(t_0)$$
$$E_{\mathbb{Q}}^{t_0}\left[\exp\left(-r_f\tau\right)S_i(t_\tau)\right] = s_i(t_0)$$

Conforme Neftci (1996, p. 122), é mais conveniente e significativamente mais fácil converter o preço do ativo em martingale, não pela subtração de seu *drift*, mas pela mudança da medida de probabilidade.

Neftci (1996, p. 107-8) expõe que um processo estocástico com variância finita é chamado de **martingale integrável quadrático contínuo**. Isto significa que se pode representar todos esses martingales pelo movimento browniano [Karatzas & Shreve (1991)]. Isto sugere que mudanças imprevisíveis e ausência de saltos são duas propriedades do movimento browniano no tempo contínuo.

Além disso, no mundo com preço da incerteza neutralizado, o processo do preço de um derivativo linear, $F_i(t) = aS_i(t) - b$, ou parcialmente linear, $F_i(t) = \max\left[0; aS_i(t) - b\right]$ ou $F_i(t) = \max\left[0; b - aS_i(t)\right]$, e que não gera fluxos de caixa

intermediários, sendo $a > 0$ e $b > 0$ constantes, deve também ser um martingale sob a mesma estrutura de probabilidade. Isto é:

$$\overline{f}_i(t_0) = E_{\mathbb{Q}}^{t_0}\left[\overline{F}_i(t_\tau)\right]$$
$$f_i(t_0) = \exp\left[-r_f\left(t_\tau - t_0\right)\right]E_{\mathbb{Q}}^{t_0}\left[F_i(t_\tau)\right]$$

2.1.3.7.3 Integral de Lebesgue: esperança e variância

Da integral de Lebesgue $\int g(x)dP(x)$ deriva-se o valor esperado, variância etc. tanto de variáveis aleatórias discretas como de contínuas.

Para esperança de uma variável aleatória discreta, $g(x) = x$, $dP(x) = \Pr(X = x)$ e a integral transforma-se num somatório.

Para uma contínua, $g(x) = x$ e $dP(x) = f(x)dx$.

Para variância $g(x) = [x - E(X)]^2$.

2.1.3.7.4 Esperança parcial

O domínio da integração é particionado no ponto k.

Portanto, para $g(x) = x$:

$$E^{(e)}(k) = \int_{-\infty}^{k} xdP(x)$$
$$E^{(d)}(k) = \int_{k}^{\infty} xdP(x)$$

Tem-se ainda que: $E^{(e)}(k) + E^{(d)}(k) = \begin{cases} E(X), \text{ se X for contínua} \\ E(X) + kf(k), \text{ se X for discreta} \end{cases}$

2.1.3.7.5 Função de perda linear

A função de perda é medida do custo de oportunidade de uma decisão, com informação adicional perfeita ou imperfeita, em relação à melhor decisão sem a informação adicional. Ela avalia o quanto deixamos de perder quando informação adicional torna-se disponível. Conforme Pratt *et al.* (1996), normalmente, em problemas econômicos, a função de perda é linear. Desta forma, o seu valor está em valor monetário (real, dólar etc.).

Considere-se $g_1(x) = \max\{x - k; 0\}$ e $g_2(x) = \max\{k - x; 0\}$.

Assim $g_1(x) - g_2(x) = k - x$.

Portanto:

$$\begin{aligned} L^{(e)} &= E\{\max(k - X; 0)\} \\ &= \int_{-\infty}^{k} (k - x) dP(x) \\ L^{(d)} &= E\{\max(X - k; 0)\} \\ &= \int_{k}^{\infty} (x - k) dP(x) \end{aligned}$$

Integrando os dois lados:

$$L^{(e)}(k) = L^{(d)}(k) + k - E(X)$$

Finalmente, tem-se que:

$$\begin{aligned} L^{(e)}(k) &= kF(k) - E^{(e)}(k) \\ L^{(d)}(k) &= E^{(e)}(k) - kG(k) \end{aligned}$$

Esta duas últimas equações são muito úteis na prática, e as fórmulas para $E^{(e)}(k)$ para as principais distribuições estão desenvolvidas no capítulo 9 de Pratt *et al.* (1996).

2.1.4 Problemas de investimento típicos

2.1.4.1 Determinação de preço

Conforme Luenberger (1998, p. 7), o problema da determinação de preço é um dos problemas básicos da moderna Ciência do Investimento.

Conforme Sharpe & Alexander (1990, p. 195-6), a determinação do preço racional de um ativo necessita da construção de um modelo, isto é, da construção de uma teoria. Isso exige simplificação das situações complexas por um processo de abstração dos seus elementos essenciais. Essa simplificação é atingida por meio de hipóteses sobre o ambiente. Essas hipóteses necessitam ser simples, de modo que forneçam um grau de abstração que permita o sucesso na construção do modelo.

Por outro lado, o teste de um modelo está em sua capacidade de auxiliar no entendimento e na predição do processo que está sendo modelado.

2.1.4.1.1 Exemplo paradigmático de determinação de preço em condição de certeza

Suponha-se uma situação determinística de determinação de preço. Imagine-se que há uma oportunidade de investimento que paga R$110,00 ao final de um ano. Qual o preço desse investimento hoje, dado o ambiente financeiro completo? Se há uma taxa de juros para um investimento de um ano de 10%, então (pelo princípio da comparação) este investimento tem o preço exato de R$100,00. Se a taxa atual de juros para este prazo de investimento é menor que 10%, então (pelo princípio da comparação) o preço do investimento deve ser maior que R$100,00. Em geral, se a taxa de juros corrente para este tipo de investimento for r, então o preço do investimento para x após um ano deve ser, pelo princípio da comparação, $\frac{x}{1+r}$. Este investimento pode ser diretamente comparado com um investimento de capital de um ano em certificado de depósito (ou dívida do Tesouro de um ano). Ele deve gerar a mesma taxa efetiva de juros.♦

Este exemplo paradigmático de determinação de preço em condição determinística é um exemplo simples do problema de determinação de preço geral. Dado um investimento com características conhecidas de retorno (que podem ser aleatórias), qual é o preço razoável? Ou equivalentemente, qual é o preço consistente com outros valores mobiliários que estão disponíveis?

No entanto, nem sempre a aplicação do princípio da comparação é tão simples como a do exemplo anterior.

2.1.4.1.2 Exemplos paradigmáticos de determinação de preço em condição de incerteza de mercado

*2.1.4.1.2.1 Contrato a termo (*forward*)*

Imagine-se um contrato chamado **a termo** (*forward*) F cujos investidores envolvidos (um comprador e um vendedor) combinam hoje $t_0 = 0$ um preço certo

$k(t_0;t_\tau) = k(t_\tau)$ para um valor mobiliário com risco S numa data futura t_τ. Note-se que os dois investidores têm respectivamente a obrigação de comprar e vender o ativo na data futura, ao preço futuro que foi pré-estabelecido hoje.

Portanto, o preço deste contrato com obrigação bilateral tem hoje o preço nulo, isto é $f(t_0) = 0$.

O retorno deste contrato é:

$$F(t_\tau) = S(t_\tau) - k(t_\tau).$$

Portanto, o preço racional do contrato a termo é:

$$\begin{aligned}
f(t_0) &= \exp\left[-r_f(t_\tau - t_0)\right]E_{\mathbb{Q}}[F(t_\tau)] \\
&= \exp\left(-r_f t_\tau\right)E_{\mathbb{Q}}[F(t_\tau)] \\
&= \exp\left(-r_f t_\tau\right)E_{\mathbb{Q}}[S(t_\tau) - k(t_\tau)] \\
&= \exp\left(-r_f t_\tau\right)E_{\mathbb{Q}}[S(t_\tau)] - \exp\left(-r_f t_\tau\right)k(t_\tau) \\
0 &= s(t_0) - \exp\left(-r_f t_\tau\right)k(t_\tau) \\
k(t_\tau) &= \exp\left(r_f t_\tau\right)s(t_0)
\end{aligned}$$

Logo, o preço futuro do contrato a termo racional é o preço atual do valor mobiliário levado a valor futuro pela taxa certa nominal.

Imagine-se que em instantes de tempo seguintes $t_\tau \geq t > t_0$ novos contratos a termo tenham preços $k(t;t_\tau) = \exp\left[r_f(t_\tau - t)\right]s(t)$. Qual o preço racional $f(t)$ do contrato a termo original?

$$\begin{aligned}
\overline{f}(t) &= E^t_{\mathbb{Q}}\left[\overline{F}(t_\tau)\right] \\
\exp\left(-r_f t\right)f(t) &= \exp\left(-r_f t_\tau\right)E^t_{\mathbb{Q}}[F(t_\tau)] \\
f(t) &= \exp\left[-r_f(t_\tau - t)\right]E^t_{\mathbb{Q}}[F(t_\tau)] \\
&= \exp\left[-r_f(t_\tau - t)\right]E^t_{\mathbb{Q}}[S(t_\tau)] - \exp\left[-r_f(t_\tau - t)\right]k(t_\tau)
\end{aligned}$$

Observe-se que um outro contrato a termo qualquer na data t possui preço nulo também, isto é, $f^*(t) = 0$.

Portanto:

$$\begin{aligned}
0 &= \exp\left[-r_f(t_\tau - t)\right]E^t_{\mathbb{Q}}[S(t_\tau) - k(t;t_\tau)] \\
E^t_{\mathbb{Q}}[S(t_\tau)] &= k(t;t_\tau)
\end{aligned}$$

Desta forma, o preço do contrato a termo original na data t é:

$$
\begin{aligned}
f(t) &= \exp\left[-r_f(t_\tau - t)\right]k(t;t_\tau) - \exp\left[-r_f(t_\tau - t)\right]k(t_\tau) \\
&= \exp\left[-r_f(t_\tau - t)\right][k(t;t_\tau) - k(t_\tau)] \\
&= \exp\left[-r_f(t_\tau - t)\right]\left\{\exp\left[r_f(t_\tau - t)\right]s(t) - \exp\left[r_f(t_\tau - t_0)\right]s(t_0)\right\} \\
&= \exp\left[-r_f(t_\tau - t)\right]\left\{\exp\left[r_f(t_\tau - t)\right]s(t) - \exp\left(r_f t_\tau\right)s(0)\right\} \\
&= s(t) - \exp\left(r_f t\right)s(0)
\end{aligned}
$$

♦

2.1.4.1.2.2 Contrato de opção européia simples ordinária sobre ação

Imagine-se que um investidor tenha o direito de comprar uma ação S numa data futura de exercício t_τ por um preço certo pré-definido no mercado $k(t_\tau)$. Suponha-se que este direito tenha hoje $(t_0 = 0)$ o preço $c(t_0) > 0$. O investidor que vendeu esta opção de compra tem a obrigação de vender a ação caso o primeiro investidor exerça a opção de compra dela na data de exercício. A ação hoje vale $s(t_0)$.

O retorno deste contrato é:

$$C(t_\tau) = \max[0; S(t_\tau) - k(t_\tau)].$$

O preço racional da opção por martingale em $t_0 \le t \le t_\tau$ é:

$$
\begin{aligned}
\bar{c}(t) &= E_{\mathbb{Q}}^t\left[\overline{C}(t_\tau)\right] \\
\exp\left(-r_f t\right)c(t) &= E_{\mathbb{Q}}^t\left[\exp\left(-r_f t_\tau\right)C(t_\tau)\right] \\
c(t) &= \exp\left[-r_f(t_\tau - t)\right]E_{\mathbb{Q}}^t\{\max[0; S(t_\tau) - k(t_\tau)]\} \\
c(t) &= \exp\left(-r_f \tau\right)E_{\mathbb{Q}}^t\{\max[0; S(t_\tau) - k(t_\tau)]\}
\end{aligned}
$$

Fazendo $t = t_0 = 0$ tem-se:

$$
\begin{aligned}
c(0) &= \exp\left(-r_f t_\tau\right)E_{\mathbb{Q}}\{\max[S(t_\tau) - k(t_\tau)]\} \\
E_{\mathbb{Q}}\left\{\frac{\max[S(t_\tau) - k(t_\tau)]}{c(0)}\right\} &= \exp\left(-r_f t_\tau\right) \qquad . \\
E_{\mathbb{Q}}\left\{\frac{C(t_\tau)}{c(0)}\right\} &= \exp\left(-r_f t_\tau\right)
\end{aligned}
$$

Pela relação de paridade opção de compra e opção de venda (paridade *call-put*) tem-se:

$$S(t_\tau) - k(t_\tau) + \max[0; k(t_\tau) - S(t_\tau)] = \max[0; S(t_\tau) - k(t_\tau)].$$

Logo:

$$c(0) = \exp\left(-r_f t_\tau\right) E_{\mathbb{Q}} \left\{ S(t_\tau) - k(t_\tau) + \max[0; k(t_\tau) - S(t_\tau)] \right\}$$
$$c(0) = s(0) - \exp\left(-r_f t_\tau\right) k(t_\tau) + \exp\left(-r_f t_\tau\right) E_{\mathbb{Q}} \left\{ \max[0; k(t_\tau) - S(t_\tau)] \right\}$$

*2.1.4.2 Estratégia de proteção contra flutuações do preço do valor mobiliário (*hedging*)*

O *hedging* objetiva eliminar ou diminuir a variação dos resultados futuros por meio da transferência da incerteza por meio de seguro, replicação dos resultados futuros por meio de estratégia de carteira replicante dinâmica de valores mobiliários, imunização ou diversificação das posições financeiras.

O *hedging* perfeito reduz a taxa de desconto ajustada ao risco à taxa de juros certa. Ele também reduz a taxa de retorno esperada ou média à taxa de juros certa.

É uma estratégia usada para compensar investimentos com incerteza. Um *hedging* perfeito é aquele que elimina totalmente a possibilidade de flutuações do preço futuro do valor mobiliário.

2.1.4.2.1 Exemplo paradigmático de hedging

Conforme Downes & Goodman (1991), um acionista preocupado com a queda dos preços das ações, por exemplo, pode fazer um *hedging* de seus títulos por meio da compra de uma opção de venda de ações ou de venda de uma opção de compra. Alguém com 100 ações de XYZ que estão sendo vendidas a R$ 70,00 a ação pode fazer *hedging* de sua posição mediante a compra de uma opção de venda que lhe assegure o direito de vender 100 ações a R$ 70,00 por ação, a qualquer momento, nos meses seguintes. Esse investidor deve pagar uma determinada quantia, denominada **prêmio** (*premium*) por esses direitos. Se a ação XYZ cair durante esse período, ele pode exercer a opção – isto é, vender as ações a R$ 70,00 -, preservando desse modo o valor de R$ 70,00 de suas ações XYZ. O mesmo acionista de XYZ pode fazer *hedging* de sua posição por meio da venda de uma opção de compra. Nessa operação, o investidor vende o direito de compra de XYZ a R$ 70,00 por ação durante alguns meses. Em contrapartida, recebe um prêmio.

Se o preço da ação XYZ cair, o prêmio recebido compensará em certa medida a queda do valor da ação. A venda a descoberto (*short selling*) é outra técnica de *hedging* amplamente utilizada. Os investidores freqüentemente tentam realizar *hedging* contra a inflação, com a aquisição de ativos cujos valores subirão mais rapidamente que a inflação. Grandes sociedades comerciais que desejam assegurar o preço que será recebido ou pago por uma *commodity* farão *hedging* de suas posições via compra e venda simultâneas no mercado futuro. Por exemplo, uma companhia de chocolate fará *hedging* de seu fornecimento de cacau no mercado futuro para limitar o risco de um aumento nos preços do produto.♦

2.1.4.3 Investimento puro

Investimento puro é aquele que objetiva obter aumento de retorno futuro por meio da alocação de capital corrente. Esta é a motivação subjacente da maioria dos investimentos no mercado acionário, por exemplo. O problema de investimento que surge a partir desta motivação é referido como problema da seleção de portfólio, já que a questão é determinar onde alocar o capital disponível. A maioria das abordagens para problemas de investimento puro baseiam-se no princípio de aversão à incerteza. Segundo Luenberger (1998, p. 8), não há solução única para este tipo de problema de investimento.

2.1.4.4 Outros problemas

Os problemas de investimento podem combinar determinação de preço, *hedging* e investimento puro.

2.1.5 Relação entre o mundo com preço da incerteza neutralizado e a atitude do decisor diante da incerteza

Conforme Hull (1998, p. 260) a determinação do preço de derivativos pode ser feita neste mundo virtual, pois "qualquer valor mobiliário sobre outros valores mobiliários negociados pode ser avaliado sob a hipótese de que os decisores são neutros ao risco." Ele observa ainda que a determinação de preço racional, no mundo com preço da incerteza neutralizado, não exige que os investidores sejam neutros ao risco. O que ele estabelece é que os derivativos, tais como opções,

podem ter seus preços racionais determinados sob a hipótese de que os investidores são neutros ao risco. Isto significa que as preferências dos investidores não têm efeito sobre o valor da opção quando ela é expressa como uma função do preço de um valor mobiliário subjacente, como a ação. A determinação de preço racional, no mundo com preço da incerteza neutralizado, é um poderoso instrumento porque, neste mundo sintético, funcionam dois resultados particularmente simples: (i) a taxa de retorno esperada de todos os valores mobiliários é a taxa de juros certa e (ii) a taxa certa é a taxa de desconto apropriada para ser aplicada a qualquer fluxo de caixa futuro. Opções e derivativos que produzem um retorno num tempo específico podem ter preço determinado neste mundo livre de preço da incerteza. Este procedimento segue as etapas para a opção européia: (i) admitir que a taxa de retorno esperada do valor mobiliário subjacente é a taxa certa, (ii) calcular o retorno esperado da opção na sua data de exercício e (iii) calcular o valor presente, usando a taxa certa do valor esperado do retorno da opção na data de exercício.

Conforme Yang (1997, p. 4), no mundo com preço da incerteza neutralizado, na realidade, o investidor típico tem aversão à incerteza, isto é, é avesso à incerteza na terminologia usual da Teoria da Decisão. Felizmente, dado um universo econômico avesso à incerteza (com preço da incerteza positivo), é possível construir um outro universo econômico com preço da incerteza neutralizado equivalente. Dessa forma, a hipótese de mundo com preço da incerteza neutralizado simplifica a determinação do preço racional, eliminando a necessidade de conhecer os detalhes das preferências e funções de utilidade de cada investidor. Estes investidores virtuais demandam apenas a taxa certa como expectativa de taxa de retorno de um investimento.

Yang (1997, p. 4) se pergunta: "O que determina o preço de um ativo?" Ela responde: volatilidade.

2.1.6 Relação entre a hipótese de mercado completo, opções e a hipótese do mundo com preço da incerteza neutralizado

Há dois tipos de mercado completo. Um uniperíodo e outro multiperíodo. A existência do mercado completo garante a existência, unicidade e determinação da medida de probabilidade neutralizadora do preço da incerteza.

2.1.6.1 Mercado completo uniperíodo

Se os valores mobiliários (ativos financeiros) forem definidos como um conjunto finito de estados futuros de preços e se houver pelo menos a mesma quantidade de ativos financeiros linearmente independentes, então o mercado é denominado **completo uniperíodo** [adaptado de Luenberger (1998, cap. 9)].

2.1.6.1.1 Existência e unicidade da medida de probabilidade neutralizadora do preço da incerteza em mercado completo uniperíodo

"[...] a parcela inesperada da taxa de retorno, aquela que decorre de surpresas, é o risco autêntico de qualquer investimento. Afinal de contas, se tivéssemos conseguido o que esperávamos, não poderia haver qualquer risco ou incerteza."

Ross, Westerfield & Jaffe (1995, p. 233)

Imagine-se que um indivíduo esteja preocupado em investir hoje (t) e receber um retorno numa data futura (τ). No futuro (τ) há apenas três estados da natureza que poderão ocorrer: depressão (ω_1), normalidade (ω_2), prosperidade (ω_3). Imagine-se ainda que o mercado de valores mobiliários tem disponível apenas um ativo financeiro com risco, i.e., a ação 1 que retorna R$ 1,00 se ocorrer depressão, R$ 2,00 se ocorrer normalidade e R$ 3,00 se ocorrer prosperidade. Matricialmente, tem-se:

$$r_1(\tau) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

No instante t (agora) a ação vale $s_1(t)$ e é um escalar.

Ao comprar x ações está-se comprando o seguinte padrão de retornos:

$$x \times r_1(\tau) = \begin{bmatrix} 1x \\ 2x \\ 3x \end{bmatrix}$$

Imagine-se que o investidor deseja o padrão de retorno $\begin{bmatrix} 0 & 1 & 2 \end{bmatrix}^T$ (matriz transposta de $\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$). Este padrão é impossível de ser criado pela ação 1, pois não existe um valor de x que reproduza o padrão desejado pelo investidor.

Entretanto, se o mercado disponibilizar um ativo financeiro sem incerteza 2, então o padrão de retornos desejado poderá ser reproduzido. Um ativo financeiro sem incerteza possui o mesmo retorno para todos os estados da natureza.

$$\begin{bmatrix} r_1(\tau) & r_2(\tau) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}$$

Portanto, o padrão de retorno que os dois ativos financeiros podem gerar é:

$$x \times r_1(\tau) + y \times r_2(\tau) = \begin{bmatrix} x + y \\ 2x + y \\ 3x + y \end{bmatrix}$$

Logo, se $x = 1$ e $y = -1$, o padrão de retorno desejado é reproduzido.

Imagine-se agora que o padrão de retorno desejado é $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$. Não existem valores de x e y que possam gerar tal padrão de retornos. Isto ocorre pois um sistema com três equações e duas incógnitas geralmente não possui solução.

A inclusão de um terceiro ativo financeiro que seja apenas uma combinação linear da ação e do ativo sem incerteza não altera a situação anterior. Mas a inclusão de um terceiro ativo financeiro com incerteza, que tenha um padrão de retorno que não pode ser gerado pelos dois ativos anteriores, é capaz de resolver o problema da reprodução do padrão de retornos desejado.

$$
\begin{aligned}
r(\tau) &= \begin{bmatrix} r_1(\tau) & r_2(\tau) & r_3(\tau) \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ 3 & 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} r_{ij} \end{bmatrix}
\end{aligned}
$$

em que i representa um estado da natureza e j representa um valor mobiliário.

Então:

$$
x \times r_1(\tau) + y \times r_2(\tau) + z \times r_3(\tau) = \begin{bmatrix} x + y + z \\ 2x + y \\ 3x + y \end{bmatrix}.
$$

Logo, $(x;y;z) = (1;-2;1)$ reproduz o padrão de retornos desejado.

Note-se agora que os três vetores de preço são linearmente independentes. O posto da matriz $r(\tau)$ é três. Se $x = -1$, $y = 3$ e $z = -2$, então o padrão de retornos $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$ pode ser gerado também. Isto implica que qualquer padrão pode ser gerado com três padrões linearmente independentes. Então, pode-se utilizar a seguinte matriz de padrões de retornos no lugar da anterior:

$$
\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
$$

Desta forma torna-se evidente que qualquer padrão de retorno pode ser reproduzido, pois basta comprar x do primeiro ativo, y do segundo e z do terceiro.

Sempre que o número de ativos financeiros linearmente independentes for igual ao número de estados da natureza, qualquer padrão de retornos pode ser reproduzido e o mercado nesta situação chama-se **completo**. Ele é completo porque nenhum outro ativo financeiro com padrão de retorno diferente, i.e., que não é mera combinação linear dos retornos dos ativos anteriores, pode ser gerado.

Criar-se-á agora um valor mobiliário especial que retorna R$ 1,00 se um estado da natureza ocorrer. Ele será chamado de **direito contingencial do estado** (*state-contingent claim*) ou **valores mobiliários Arrow-Debreu** [Varian (1987, p. 58)]. Logo, se o mercado for completo, o espaço de retornos gerado pelos ativos reais

também pode ser gerado pelos direitos contingenciais $\pi_1(\tau)$, $\pi_2(\tau)$ e $\pi_2(\tau)$, um para cada estado da natureza:

$$\pi_1(\tau) = \begin{bmatrix} 1, \text{ se ocorrer } \omega_1 \\ 0, \text{ caso contrário} \\ 0, \text{ caso contrário} \end{bmatrix}$$

$$\pi_2(\tau) = \begin{bmatrix} 0, \text{ caso contrário} \\ 1, \text{ se ocorrer } \omega_2 \\ 0, \text{ caso contrário} \end{bmatrix}$$

$$\pi_3(\tau) = \begin{bmatrix} 0, \text{ caso contrário} \\ 0, \text{ caso contrário} \\ 1, \text{ se ocorrer } \omega_3 \end{bmatrix}$$

Logo:

$$\begin{aligned} \pi(\tau) &= \begin{bmatrix} \pi_1(\tau) & \pi_2(\tau) & \pi_3(\tau) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

Se o mercado for completo e for utilizado o princípio da não-arbitragem, é possível inferir os preços dos direitos contingenciais a partir dos preços dos ativos reais. Em termos de Álgebra Linear, deseja-se formar os preços da base ortogonal de vetores unitários do espaço de retornos.

Juntando os resultados anteriores e utilizando o princípio de não-arbitragem (dois ativos com o mesmo padrão de retorno devem ter o mesmo preço) o preço no instante t do ativo Arrow-Debreu, que ocorre no estado da natureza 1, deve ter o mesmo valor do preço do ativo 3, pois eles têm o mesmo padrão de retornos em τ (futuro). Logo:

$$\begin{aligned} \pi_1(t) &= 0 \times s_1(t) + 0 \times s_2(t) + 1 \times s_3(t) \\ \pi_2(t) &= -1 \times s_1(t) + 3 \times s_2(t) - 2 \times s_3(t) \\ \pi_3(t) &= 1 \times s_1(t) - 2 \times s_2(t) + 1 \times s_3(t) \end{aligned}$$

Portanto:

$$
\begin{aligned}
s_1(t) &= 1 \times \pi_1(t) + 2 \times \pi_2(t) + 3 \times \pi_3(t) \\
s_2(t) &= 1 \times \pi_1(t) + 1 \times \pi_2(t) + 1 \times \pi_3(t) \\
s_3(t) &= 1 \times \pi_1(t) + 0 \times \pi_2(t) + 0 \times \pi_3(t)
\end{aligned}
$$

Generalizando:

$$
\begin{aligned}
s(t) &= \begin{bmatrix} s_1(t) & s_2(t) & s_3(t) \end{bmatrix} \\
\pi(t) &= \begin{bmatrix} \pi_1(t) \\ \pi_2(t) \\ \pi_3(t) \end{bmatrix} \\
r(\tau) &= \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ 3 & 1 & 0 \end{bmatrix} \\
s(t) &= [\pi(t)]^T \, r(\tau) \\
[\pi(t)]^T &= s(t)[r(\tau)]^{-1} \\
\text{em que } A^T &= \text{matriz transposta} \\
\mathrm{A}^{-1} &= \text{matriz inversa}
\end{aligned}
$$

Se um ativo tiver a estrutura de retornos $\begin{bmatrix} 1 & 2 & 1 \end{bmatrix}^T$, seu preço racional será dado por $\pi_1(t) + 2\pi_1(t) + \pi_3(t)$.

Observe-se ainda que o preço de um valor mobiliário Arrow-Debreu é necessariamente menor que um, i.e.:

$$
\begin{aligned}
&0 < \pi_i(t) < 1 \\
&i = 1,2,3
\end{aligned}
$$

Como

$$
r_2(\tau) = \sum_{i=1}^{3} \pi_i(\tau),
$$

i.e., a soma dos retornos dos valores mobiliários Arrow-Debreu reproduz um ativo sem incerteza, então, por arbitragem, o preço do ativo sem incerteza deve ser igual à soma dos preços dos valores mobiliários Arrow-Debreu:

$$
s_2(t) = \sum_{i=1}^{3} \pi_i(t)
$$

Como o preço do ativo sem incerteza é o inverso de 1 mais a taxa certa ao período, então:

$$
\begin{aligned}
s_2(t) &= \sum_{i=1}^{3} \pi_i(t) \\
&= \frac{1}{1+r_f}
\end{aligned}
.
$$

Note-se ainda que multiplicando os dois lados por $1+r_f$ tem-se:

$$
\begin{aligned}
\left(1+r_f\right)\sum_{i=1}^{3} \pi_i(t) &= 1 \\
q_1(\tau)+q_2(\tau)+q_3(\tau) &= 1, \\
\text{sendo } q_i(\tau) &= \left(1+r_f\right)\pi_i(t) \\
&= vf_{r_f}\left[\pi_i(t)\right]
\end{aligned}
$$

Chamar-se-á $\mathbb{Q}(\tau) = \begin{bmatrix} q_1(\tau) \\ q_2(\tau) \\ q_3(\tau) \end{bmatrix}$ de medida de probabilidade sintética ou neutralizadora do preço da incerteza.

Logo, uma nova maneira de determinar o preço racional e único de um ativo surge:

$$
\begin{aligned}
s_j(t) &= \sum_{i=1}^{3} \pi_i(t) r_{ij}(\tau) \\
&= \sum_{i=1}^{3} \frac{q_i(\tau)}{\left(1+r_f\right)} r_{ij}(\tau) \\
&= \frac{1}{\left(1+r_f\right)} \sum_{i=1}^{3} q_i(\tau) r_{ij}(\tau) \\
&= \frac{1}{\left(1+r_f\right)} E_{\mathbb{Q}(\tau)}\left[R_j(\tau)\right] \\
&= vp_{r_f}\left\{E_{\mathbb{Q}(\tau)}\left[R_j(\tau)\right]\right\}
\end{aligned}
$$

Portanto, o preço racional de um valor mobiliário, num mercado completo, é o valor presente do valor esperado do retorno na medida de probabilidade neutralizadora do preço da incerteza $\mathbb{Q}(\tau)$.

A existência da medida de probabilidade neutralizadora do preço da incerteza garante a condição de não-arbitragem.

Note-se que a medida de probabilidade neutralizadora do preço da incerteza independe das probabilidades de ocorrência dos estados da natureza. Ela depende apenas das informações sobre os ativos básicos linearmente independentes.

Qual deveria ser a taxa de desconto se as probabilidades dos estados da natureza fossem utilizadas?

Como a taxa de desconto r deve ser maior que a taxa certa r_f, então o valor esperado do retorno, na medida de probabilidade real ou subjetiva, deve ser maior que o valor esperado do retorno, na medida de probabilidade neutralizadora do preço da incerteza , i.e.,

$$E_{\mathbb{P}(\tau)}\left[R_j(\tau)\right] > E_{\mathbb{Q}(\tau)}\left[R_j(\tau)\right].$$

Além disso, r_j é chamado de **taxa de desconto ajustada ao risco** do valor mobiliário j e pode ser expresso como

$$\begin{aligned} r_j &= r_f + r_{\text{preço da incerteza do valor mobiliario } j} \\ &= r_f + r_p^j \end{aligned}.$$

Uma discussão mais detalhada sobre esta decomposição da taxa de desconto encontra-se em Neftci (1996, cap. 2 e 14).

Por arbitragem, a segunda igualdade deve funcionar:

$$\begin{aligned} s_j(t) &= \frac{1}{\left(1+r_f\right)} E_{\mathbb{Q}(\tau)}\left[R_j(\tau)\right] \\ &= \frac{1}{\left(1+r_j\right)} E_{\mathbb{P}(\tau)}\left[R_j(\tau)\right] \\ &= \frac{1}{\left(1+r_f+r_p^j\right)} E_{\mathbb{P}(\tau)}\left[R_j(\tau)\right] \end{aligned}$$

Note-se que a parcela da taxa de desconto referente ao preço da incerteza foi incorporada pela medida de probabilidade neutralizadora do preço da incerteza $\mathbb{Q}(\tau)$!

2.1.6.2 Utilidade do mercado completo

Por que, desconsiderando custos de transação, os mercados completos são desejáveis?

Primeiramente, são desejáveis porque possibilitam a máxima flexibilidade aos investidores. Mas, mesmo num mercado incompleto, os investidores poderiam estar satisfeitos com os padrões de retorno existentes. A resposta mais precisa é a seguinte: um mercado completo é sempre Pareto-eficiente, independentemente da

natureza dos investidores, enquanto o mercado incompleto pode ser Pareto-ineficiente em algumas circunstâncias. Isto é, num mercado financeiro Pareto-eficiente, nenhuma mudança no mercado poderia ter aprovação unânime dos investidores.

É impossível, na prática, a existência de mercado completo devido aos custos de transação, regulamentação governamental etc. Poder-se-iam criar outros ativos financeiros que tornassem o mercado financeiro menos incompleto com não excessivos custos de transação? A resposta é sim, e estes ativos são as opções.

2.1.6.2.1 A contribuição das opções

Jogar geralmente refere-se à atividade de comprar uma recompensa monetária futura que pode:

1. Ser improdutiva do ponto de vista social;
2. Ter um valor futuro esperado negativo;
3. Ser altamente incerto;
4. Ter um curto tempo antes do resultado;
5. Surgir de diferenças de opinião entre vários participantes que apostam uns contra os outros.

Estas cinco características definem a aposta em corrida de cavalos. O mercado de opções, segundo Cox & Rubinstein (1985, p. 436), além de ter todas as características de jogo, possui também as características de investimentos úteis: obtenção de novos padrões de retorno, redução dos custos de obtenção dos padrões de retornos existentes e disseminação de informações úteis, como a volatilidade.

Suponha-se que exista apenas o ativo 1 no mercado. Imagine-se agora que duas opções de compra (*call*) sejam criadas sobre este único ativo, com preços de exercício $k_\tau = 1$ e $k_\tau = 2$. Portanto, lembrando que o resultado da opção de compra é o máximo entre zero e a diferença entre o preço do valor mobiliário e o preço de exercício, isto é, $S_1(\tau) - k_\tau$:

$$C\left(S_1(\tau);k_\tau = 1\right) = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

$$C\left(S_1(\tau);k_\tau = 2\right) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Logo, o mercado agora tem três ativos: um ativo base com risco e duas opções de compra.

$$\begin{bmatrix} S_1(\tau) & C\left(S_1(\tau);1\right) & C\left(S_1(\tau);2\right) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix}$$

Note-se que o posto da matriz acima é três. Portanto, o mercado foi completado pelas duas opções de compra sobre um único ativo. Note-se ainda que o ativo 1 é equivalente a uma opção de compra com valor de exercício igual a zero:

$$S_1(\tau) = C\left(S_1(\tau);0\right).$$

Suponha-se, no entanto, que o único ativo do mercado tenha o seguinte padrão de retornos:

$$S_1(\tau) = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$$

As opções ordinárias são insuficientes para completar o mercado.

Será que outras opções simples, i.e., com resultados não simplesmente lineares $C\left(S_1(\tau);k_\tau\right) = máx\left\{0;S_1(\tau) - k_\tau\right\}$, que dependem apenas do ativo base como $C\left(S_1(\tau);k_\tau\right) = máx\left\{S_1^2(\tau);\left(S_1(\tau) - k_\tau\right)^2\right\}$, por exemplo, completariam o mercado? A resposta é não. Isto é, as opções de compra e venda ordinárias (*vanilla*), conforme Ross (1976)[17], *apud* Cox & Rubinstein (1985), para um dado conjunto de ativos financeiros básicos, geram o mesmo espaço que todas as opções simples geram.

---

[17] Ross, S. 1976. "Options and efficiency." *Quarterly Journal of Economics* 90:75-89.

Existe algum outro tipo de opção que pudesse completar o mercado? Sim. Existem as opções complexas definidas sobre carteiras de ativos. Ross (1976) demonstra que as opções complexas completarão o mercado sempre que não houver duas linhas iguais de retornos dos ativos básicos.

2.1.7 Relação entre hedging perfeito e a hipótese do mundo com preço da incerteza neutralizado

Os exemplos paradigmáticos a seguir manifestam a existência da medida de probabilidade neutralizadora do preço da incerteza e sua independência da medida de probabilidade real ou subjetiva. Observe-se ainda que ela surge devido à operação contra perdas (*hedging*).

Este resultado é sintetizado por Baxter & Rennie (1998, p. 199-200) da seguinte maneira: *hedging* perfeito significa preço único, pois há apenas uma medida de probabilidade neutralizadora do preço da incerteza.

Além disso, conforme Baxter & Rennie (1998, p. 198-9), a medida de probabilidade neutralizadora do preço da incerteza significa impossibilidade de arbitragem, pois esta medida é realmente a essência da ausência de arbitragem, isto é, de um jogo justo.

Samperi (1998, p. 170-1) demonstra que o modelo de mercado é completo *se e somente se* há uma única medida de probabilidade neutralizadora do preço da incerteza. Neste caso todo derivativo possui um preço único e livre de arbitragem (racional), dado pela esperança matemática do valor intrínseco do derivativo, na medida de probabilidade neutralizadora do preço da incerteza.

A medida de probabilidade neutralizadora do preço da incerteza manifesta-se tanto em condição de incerteza técnica quanto em condição de incerteza de mercado. Condições diferentes de incerteza acarretam apenas formas distintas de *hedging*, i.e., seguro governamental, rateios em apostas ou carteira replicante. Os dois primeiros exemplos paradigmáticos (aposta em partidas de futebol e o da opção de abandono) referem-se a decisões de investimento em condições de incerteza técnica. O terceiro exemplo paradigmático refere-se à determinação do preço de uma aposta digital em condição de incerteza de mercado.

*2.1.7.1 Determinação de rateios (*odds*) por um bookmaker para hedgear uma aposta (estratégia de seguro dos resultados)*

Baxter & Rennie (1998), no início do livro, resolvem o problema do *bookmaker*. Ele está aceitando apostas para uma corrida de dois cavalos. Analisando os cavalos ele consegue determinar as probabilidades subjetivas de vitória deles. Em seguida, os autores analisam estratégias para administração do risco. Resolveu-se, no entanto, adaptar o problema substituindo a corrida de cavalos por uma partida de futebol de final de campeonato.

Uma casa de apostas está aceitando apostas em uma partida de futebol entre o São Paulo e o Corinthians. De acordo com o desempenho das equipes, o São Paulo tem 75% de vitória contra 25% do Corinthians. Respeitando-se as probabilidades, os rateios (*odds*) para cada time são:

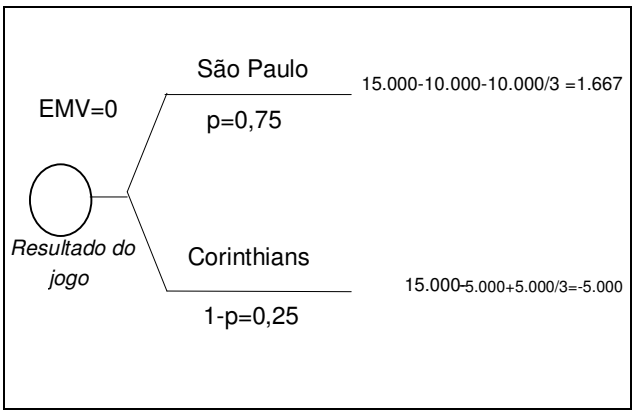
	São Paulo	Corinthians
Pr(vitória)	0,75=3/4	0,25=1/4
Vantagem	1 – 3	3 - 1

Rateio é uma maneira de expressar as apostas no formato *n*-*m*, sendo:

n: prêmio pago pela aposta;

m: valor apostado.

Suponha-se que a casa de apostas tenha recebido um volume nos valores de R$10 mil para o São Paulo e R$5 mil para o Corinthians. O lucro ou prejuízo da casa estão condicionados ao resultado da partida, podendo assumir os valores:



Os rateios derivaram das probabilidades subjetivas.

As características da situação são:

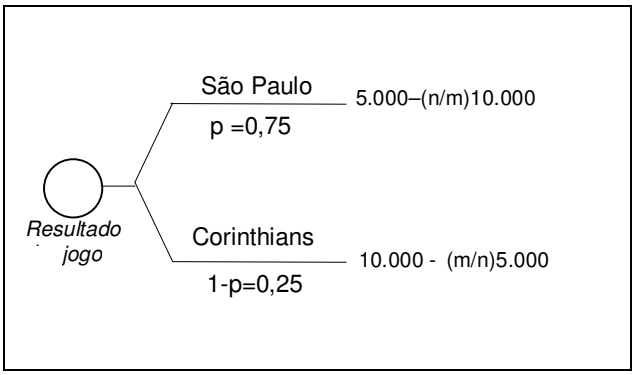
- A casa de apostas possui possibilidade de perda ao estabelecer tais apostas;
- O valor esperado deste resultado parte de algumas premissas:
  - as probabilidades subjetivas ou reais são válidas;
  - no longo prazo seu lucro médio será nulo;
  - a casa de apostas é neutra à incerteza.

Haveria alguma forma de evitar a completamente a possibilidade de perda da casa de apostas, para qualquer resultado da partida?

Suponha-se que a casa de apostas pudesse modificar os rateios de n-m para o São Paulo e m-n para o Corinthians, e as apostas continuassem nos mesmos valores:

	São Paulo	Corinthians
Pr(vitória)	0,75	0,25
Rateio	n-m	m-n

Qualquer que seja o resultado da partida, o lucro/prejuízo da casa de apostas pode ser representado pela árvore abaixo:



Observe-se que, se $m = 2$ e $n = 1$, o valor esperado é zero, quaisquer que sejam as probabilidades subjetivas. A possibilidade de perda foi totalmente eliminada, e a possibilidade de ganho também. Note-se que as probabilidades subjetivas não são mais necessárias.

Como estes valores foram obtidos?

Basta fazer cada resultado ser igual a zero. Portanto, os resultados devem ser iguais, independentemente do resultado, i.e.,

$$5000 - \frac{n}{m}10000 = 0$$
e
$$10000 - \frac{m}{n}5000 = 0$$

Logo, $\frac{m}{n} = \frac{1}{2}$.

Portanto, as probabilidades derivadas do rateio são:

$$\begin{aligned}\text{Pr(Sao Paulo vencer)} &= \frac{m}{n+m} \\ &= \frac{2}{3} \\ &\approx 0,67\end{aligned}$$

e

$$\begin{aligned}\text{Pr(Corinthians vencer)} &= \frac{n}{n+m} \\ &= \frac{1}{3} \\ &\approx 0,33\end{aligned}$$

Então:

$$\text{Pr(Sao Paulo vencer)} + \text{Pr(Corinthians vencer)} = 1$$

Se o *bookmaker* quiser ter um ganho certo de mil reais por aposta, então basta fazer:

$$5000 - \frac{n}{m}10000 = 1000$$
e
$$10000 - \frac{m}{n}5000 = 1000 \quad .$$

Da primeira equação, tem-se que $\frac{n}{m} = \frac{2}{5}$ e

$$\begin{aligned}\text{Pr(Sao Paulo vencer)} &= \frac{5}{7} \\ &\approx 0,71\end{aligned}$$

e da segunda tem-se que $\frac{n}{m} = \frac{5}{9}$ e

$$\begin{aligned}\Pr(\text{Corinthains vencer}) &= \frac{5}{14} \\ &\approx 0,36\end{aligned}$$

Observe-se agora que a soma das probabilidades derivadas dos rateios é maior que um, i.e.,

$$\begin{aligned}\Pr(\text{Sao Paulo vencer}) + \Pr(\text{Corinthians vencer}) &= \frac{15}{14} \\ &\approx 1,07\end{aligned}$$

O *bookmaker*, além de eliminar a possibilidade de perda, operando os rateios, consegue ainda ter um ganho certo de mil reais líquidos por aposta, independentemente do resultado da partida!

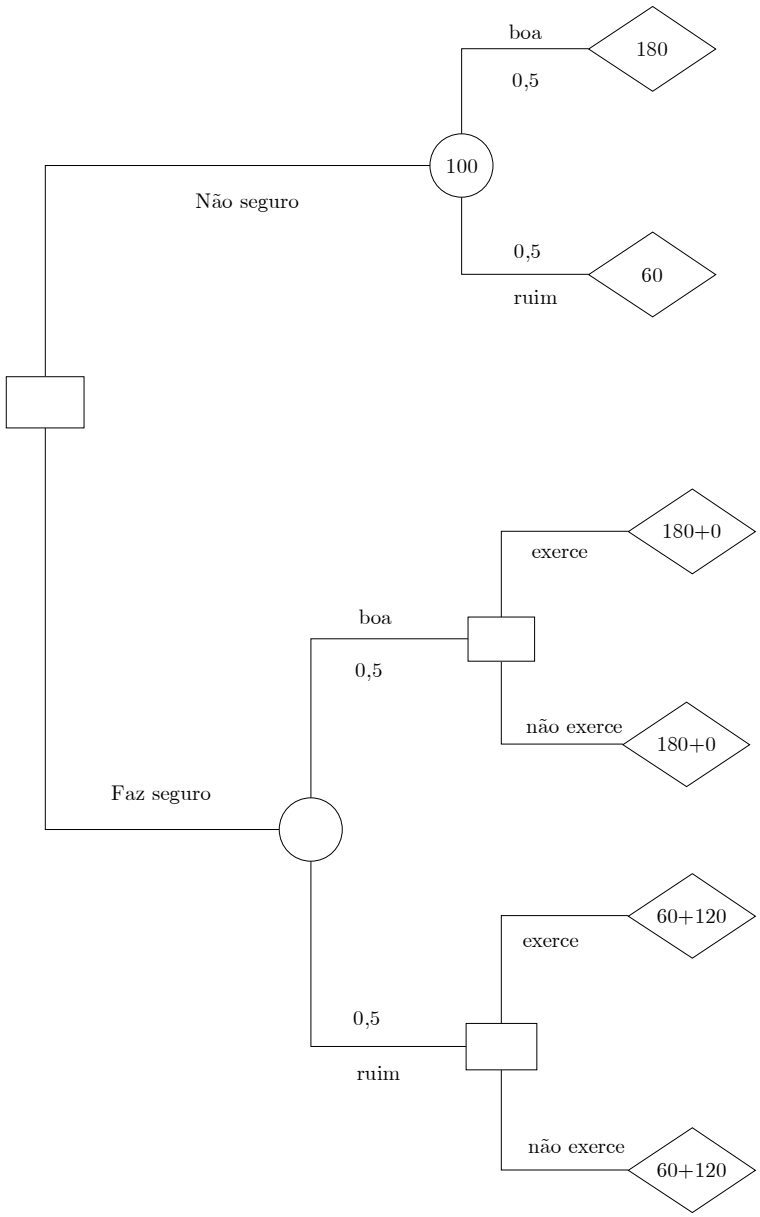
Este é um exemplo simples de arbitragem.

Sassatani (1999) estende este exemplo para dois *bookmakers* e relaciona os ativos Arrow-Debreu com as probabilidades sintéticas.

2.1.7.2 Cálculo do preço de uma opção de abandono de projeto com cobertura completa de risco por meio de seguro governamental

O problema a seguir encontra-se em Trigeorgis (1998, p. 4-7).

Uma empresa deseja avaliar uma oportunidade de investir num projeto (por exemplo, num esforço de pesquisa e desenvolvimento para descobrir uma nova droga) que, em um ano, irá gerar 180 milhões de dólares se as condições forem boas ($v^+ = 180$) e 60 milhões de dólares se as condições forem ruins ($v^- = 60$). Os dois resultados têm a mesma probabilidade de ocorrência ($p = 0,5$). O governo, querendo apoiar este projeto, oferece uma garantia (seguro) de comprar a produção inteira por 180 milhões de dólares se a situação ruim ocorrer. Sem a garantia do governo, o fluxo de caixa do projeto possui uma taxa de retorno esperada (ou taxa de retorno ajustada ao risco) de $r = 0,20$ ao ano, e uma taxa de juros certa $r_f = 0,08$ ao ano. Qual é o valor deste projeto (v) e o valor da opção de abandono do projeto proporcionada pela garantia do governo (g)?



A garantia (seguro) oferecida pelo governo funciona como uma opção de venda (*put)* européia ordinária, dando à empresa o direito de "vender" o valor do projeto e receber o valor da garantia (ou o preço de exercício) de 180 milhões de dólares. Sob a condição favorável, o valor da garantia será nulo ($g^{+} = 0$). Sob a condição desfavorável, o valor da garantia será de 120 milhões de dólares ($g^{-} = 180 - 60 = 120$).

Utilizando a técnica tradicional de fluxo de caixa descontado (FCD), o valor presente do projeto sem a garantia governamental (seguro) é:

$$
\begin{aligned}
v &= \frac{E_{\mathbb{P}}\left[C_1\right]}{1+r} \\
&= \frac{p \times v^{+} + (1-p) \times v^{-}}{1+r} \\
&= \frac{0,5 \times 180 + 0,5 \times 60}{1+0,20} \\
&= 100
\end{aligned}
$$

Já o valor do projeto com a garantia é:

$$
\begin{aligned}
v^{*} &= \frac{E_{\mathbb{P}}\left[C^{*}{}_1\right]}{1+r} \\
&= \frac{p \times \left(v^{+} + g^{+}\right) + (1-p) \times \left(v^{-} + g^{-}\right)}{1+r} \\
&= \frac{0,5 \times 180 + 0,5 \times (60 + 120)}{1+0,20} \\
&= \frac{0,5 \times 180 + 0,5 \times 180}{1,20} \\
&= 150
\end{aligned}
$$

O valor da garantia (seguro) governamental (opção de venda de abandono) é:

$$
\begin{aligned}
g &= v^{*} - v \\
&= 150 - 100 \\
&= 50
\end{aligned}
$$

A avaliação, por este método, supõe que os resultados da opção de venda (garantia) possuem o mesmo risco e podem ser descontados pela mesma taxa de desconto utilizada no projeto sem a garantia, isto é:

$$
\begin{aligned}
g &= \frac{E_{\mathbb{P}}\left[G_1\right]}{1+r} \\
&= \frac{p \times g^{+} + (1-p) \times g^{-}}{1+r} \\
&= \frac{0,5 \times 0 + 0,5 \times 120}{1+0,20} \\
&= 50
\end{aligned}
$$

No entanto, a aplicação deste método (FCD) está claramente errada, pois a possibilidade de abandonar o projeto, daqui a um ano, por um preço garantido, altera o risco do projeto e, consequentemente, a taxa de desconto. De fato, neste caso, com preço de exercício $\left(k = 180\right)$ igual ao resultado do projeto na situação favorável $\left(v^{+} = 180\right)$, a garantia do governo elimina totalmente o risco do projeto, já que a empresa irá receber 180 milhões, independentemente da situação ser favorável ou desfavorável! Já que o risco foi totalmente eliminado do projeto, há

um fluxo de caixa determinístico de 180 daqui um ano, que deverá ser descontado apenas pela taxa certa de 8% ao ano. Portanto, fazendo corretamente o cálculo, o valor do projeto com a garantia é:

$$\begin{aligned} v^* &= \frac{180}{1+0,08} \\ &= 166,7 \end{aligned}$$

Logo, o valor da opção de venda de abandono é:

$$\begin{aligned} g &= v^* - v \\ &= 166,7 - 100 \\ &= 66,7 \end{aligned}$$

Note-se que este valor é consistente com um valor negativo *ad hoc* da taxa de desconto para a opção de compra de $r_h = -0{,}10$ ao ano (já que a garantia funciona como um *hedging*). Isto se confirma pela seguinte equação:

$$\begin{aligned} v &= \frac{E_{\mathbb{P}}[G_1]}{1+r_h} \\ &= \frac{p \times v^+ + (1-p) \times v^-}{1+r_h} \\ &= \frac{0,5 \times 0 + 0,5 \times 120}{1+(-0{,}10)} \\ &= 66,7 \end{aligned}$$

O método FCD subestima o valor da opção de venda de abandono do projeto.

Mesmo neste exemplo canônico, se o valor do preço de exercício fosse diferente do valor do resultado do projeto em condição favorável ($v^+ = 180$), por exemplo $k = 100$, a taxa de desconto correta não poderia ser prontamente determinada. O emprego da taxa certa r poderia ser inconsistente com o uso das probabilidades reais ou subjetivas ($p = 0{,}5$) se o risco não fosse completamente eliminado. No entanto, como se verá adiante, pode-se ainda aplicar a taxa certa se a medida de probabilidade for alterada. Esta medida chamar-se-á **medida de probabilidade neutralizadora do preço da incerteza** ou **equivalente certo** ($\mathbb{Q}$). Utilizando $q = 0,4$ neste caso, obtém-se que o valor correto tanto do projeto sem garantia quanto do projeto com garantia é:

$$
\begin{aligned}
v &= \frac{E_{\mathbb{Q}}[C_1]}{1+r_f} \\
&= \frac{q \times v^+ + (1-q) \times v^-}{1+r_f} \\
&= \frac{0,4 \times 180 + 0,6 \times 60}{1+0,08} \\
&= 100
\end{aligned}
$$

$$
\begin{aligned}
g &= \frac{E_{\mathbb{Q}}[G_1]}{1+r_f} \\
&= \frac{q \times g^+ + (1-q) \times g^-}{1+r_f} \\
&= \frac{0,4 \times 0 + 0,6 \times 120}{1+0,08} \\
&= 66,7
\end{aligned}
$$

Portanto,

$$
\begin{aligned}
v^* &= v + g \\
&= 100 + 66,7 \\
&= 166,7
\end{aligned}
$$

Mesmo se a garantia tiver um preço de exercício (k) diferente de 180 milhões de dólares (portanto, há risco no projeto), o valor da opção de abandono pode ser calculada prontamente por meio do valor presente, descontado pela taxa certa do valor esperado, na medida de probabilidade neutralizadora do preço da incerteza, isto é (para $k = 100$):

$$
\begin{aligned}
g &= \frac{E_{\mathbb{Q}}[G_1]}{1+r_f} \\
&= \frac{q \times g^+ + (1-q) \times g^-}{1+r_f} \\
&= \frac{0,4 \times 0 + 0,6 \times (100-60)}{1+0,08} \\
&= \frac{0,4 \times 0 + 0,6 \times 40}{1,08} \\
&= 22,2
\end{aligned}
$$

2.1.7.3 Exemplo paradigmático do apreçamento racional de uma aposta digital

O exemplo paradigmático a seguir do apreçamento de uma aposta digital foi adaptado de Baxter & Rennie (1998, p. 15).

No mercado há um ativo de renda fixa (*bond*) com taxa de juros nula ($r_f = 0$), e uma ação cujos valores hoje são iguais a um, isto é, $b = 1$ e $s = 1$. Ao final do período (na data τ), a ação pode valer R$2,00 ou R$0,50, isto é, $s_1(\tau) = 2$ e $s_2(\tau) = \frac{1}{2}$. Qual é o valor de uma aposta f, que paga R$1,00, se a ação subir, e nada (R$0,00), se ela cair, isto é, $f_1(\tau) = 1$ e $f_2(\tau) = 0$?

Tem-se inicialmente a possibilidade construir uma carteira no instante inicial, que permite a reprodução exata dos valores da aposta:

$$\begin{aligned}\pi &= \alpha \times s + \beta \times b \\ &= \alpha \times 1 + \beta \times 1 \\ &= \alpha + \beta\end{aligned}$$

em que π é o valor da carteira no instante inicial, α é a quantidade de ações e β é a quantidade de *bonds* na carteira replicante.

Os valores de alfa e beta devem satisfazer as seguintes equações para replicar os resultados da aposta:

$$\Pi(\tau) = \alpha \times S(\tau) + \beta \times b = F(\tau)$$

$$\pi_1(\tau) = \alpha \times s_1(\tau) + \beta \times b = f_1(\tau)$$

$$\pi_2(\tau) = \alpha \times s_2(\tau) + \beta \times b = f_2(\tau)$$

$$1 = \alpha \times 2 + \beta \times 1$$

$$0 = \alpha \times \frac{1}{2} + \beta \times 1$$

Portanto:

$$\begin{aligned}\alpha^* &= \frac{2}{3} \\ \beta^* &= -\frac{1}{3}\end{aligned}$$

Deve-se comprar 2/3 de uma ação e tomar empréstimo de 1/3 de um *bond*. O custo desta carteira replicante é:

$$
\begin{aligned}
\pi^* &= \alpha^* + \beta^* \\
&= \frac{2}{3} - \frac{1}{3} \quad . \\
&= \text{R\$ } \frac{1}{3}
\end{aligned}
$$

Como esta carteira replica exatamente os resultados da aposta, o preço da aposta deve ser igual ao custo da carteira, pelo princípio da não-arbitragem. Isto é:

$$
f = \pi^* .
$$

Logo, o preço desta aposta é R$1/3.

Note-se novamente que as probabilidades reais ou subjetivas da opção subir ou cair são irrelevantes para o cálculo do preço da aposta.

Intuitivamente, se fossem conhecidas as probabilidades reais ou subjetivas de subida e de descida do preço da ação (medida de probabilidade $\mathbb{P}$), seria possível imaginar que o preço da aposta deveria ser calculado pelo valor esperado dos resultados da aposta:

$$
\begin{aligned}
f &= E_{\mathbb{P}}\left[F(\tau)\right] \\
&= p_{\text{subir}} \times 1 + p_{\text{descer}} \times 0 \\
&= \text{R\$ } p_{\text{subir}}
\end{aligned}
$$

No entanto, no mercado haveria possibilidade de arbitragem.

Seria interessante calcular o preço da aposta utilizando também a idéia de valor esperado (operador esperança matemática).

Como será mostrado mais adiante, é possível calcular o preço da aposta desta forma, porém é preciso construir uma nova medida de probabilidade denominada **medida de probabilidade neutralizadora do preço da incerteza** $\mathbb{Q}$.

Para o cálculo da probabilidade neutralizadora do preço da incerteza de subida da ação, basta utilizar os valores (informações disponíveis da ação) da ação:

$$
\begin{aligned}
q_{\text{subir}} &= \frac{s - s_2(\tau)}{s_1(\tau) - s_2(\tau)} \\
&= \frac{1 - 0,5}{2 - 0,5} \\
&= \frac{1}{3}
\end{aligned}
$$

Portanto, o preço único e livre de arbitragem da aposta é:

$$\begin{aligned} f &= E_{\mathbb{Q}}\left[F(\tau)\right] \\ &= q_{\text{subir}} \times f_1(\tau) + q_{\text{descer}} \times f_2(\tau) \\ &= q_{\text{subir}} \times 1 + q_{\text{descer}} \times 0 \\ &= q_{\text{subir}} \\ &= \text{R\$ } \frac{1}{3} \end{aligned} .$$

Note-se que s é o preço *spot* e os valores $s_1(\tau)$ e $s_2(\tau)$ correspondem à informação sobre o domínio do preço da ação, isto é, informação de onde o preço variará. Estas duas informações constituem o conjunto informacional do preço da ação no instante zero. A medida de probabilidade $\mathbb{P}$, medida de probabilidade subjetiva ou real, é uma informação especulativa sobre como o preço variará.

De modo mais geral, tem-se que ($r_f > 0$ e constante no período considerado):

$$\begin{aligned} \pi_1(\tau) &= \alpha s_1(\tau) + \beta b \exp(r_f \tau) = f_1(\tau) \\ \pi_2(\tau) &= \alpha s_2(\tau) + \beta b \exp(r_f \tau) = f_2(\tau) \end{aligned}$$

Logo, o custo da carteira replicante é:

$$\begin{aligned} \alpha^* &= \frac{f_1(\tau) - f_2(\tau)}{s_1(\tau) - s_2(\tau)} \\ \beta^* &= b^{-1} \exp(-r_f \tau) \left(\frac{f_2(\tau) s_1(\tau) - f_1(\tau) s_2(\tau)}{s_1(\tau) - s_2(\tau)} \right) \\ \pi^* &= \alpha^* s + \beta^* b \end{aligned}$$

Portanto, o preço do derivativo é o valor presente pela taxa de desconto certa da esperança do resultado (*payoff*) do derivativo, na medida de probabilidade $\mathbb{Q}$.

$$\begin{aligned} f &= \pi^* \\ &= \exp(-r_f \tau) \left\{ \left(\frac{s \exp(r_f \tau) - s_2(\tau)}{s_1(\tau) - s_2(\tau)} \right) f_1(\tau) + \left[1 - \left(\frac{s \exp(r_f \tau) - s_2(\tau)}{s_1(\tau) - s_2(\tau)} \right) \right] f_2(\tau) \right\} \\ &= \exp(-r_f \tau) \left[q f_1(\tau) + (1 - q) f_2(\tau) \right] \\ &= vp_{r_f} \left\{ E_{\mathbb{Q}}\left[F(\tau)\right] \right\} \end{aligned}$$

em que:

$$q = \frac{s \exp(r_f \tau) - s_2(\tau)}{s_1(\tau) - s_2(\tau)}$$

Outra forma de definir mercado completo é: um mercado é completo se e somente se existe apenas uma medida de probabilidade neutralizadora do preço da incerteza para valor mobiliário subjacente [Samperi (1998)].

2.2 Modelo de Black-Scholes-Merton para determinação do preço racional da opção européia simples ordinária sobre ação

"Os derivativos não são transações envolvendo ações ou taxas de juros, vidas humanas, casas vulneráveis a incêndios ou hipotecas imobiliárias.
O produto das transações com derivativos é a própria incerteza".

Bernstein (1997, p. 315)

2.2.1 A opção e suas propriedades

As opções, contratos futuros e outros tipos de derivativos existem em virtude da imprevisibilidade dos preços dos ativos em uma economia. Se estes fossem previsíveis, não haveria risco, não seria possível especular sobre os preços futuros e não haveria necessidade de proteção contra perdas (*hedging*).

Um contrato de opção pode ser definido como um direito de comprar (ou vender) um determinado ativo por um preço prefixado k_τ até (ou na) data de exercício τ. Participam deste contrato duas partes que se denominam **titular** e **lançador**. O detentor deste direito, chamado **titular da opção**, paga um prêmio ao lançador da opção que passa a ter uma obrigação. A avaliação de um contrato de opção significa determinação de um preço justo para o prêmio.

As opções podem ser classificadas quanto ao tipo em:

- opção de compra (*call option*);
- opção de venda (*put option*).

Quanto ao momento do exercício, as opções se classificam em:

- européias: o titular da opção pode exercer seu direito de compra (ou venda) somente na data de vencimento;
- americanas: o titular da opção pode exercer seu direito de compra (ou venda) a qualquer momento entre a data da operação e o dia do vencimento.

A avaliação das opções americanas é mais complexa, pois a cada instante é preciso determinar se o exercício antecipado produz benefícios ao titular da opção. Os valores intrínsecos de uma opção de compra (*call*) e uma opção de venda (*put*) no vencimento devem ser representados por:

$$C_\tau = \max(0; S_\tau - k_\tau)$$

$$P_\tau = \max(0; k_\tau - S_\tau)$$

C_τ : preço da *call* no instante τ (vencimento);

P_τ : preço da *put* no instante τ;

S_τ : valor do ativo em τ;

k_τ : preço de exercício em τ.

Observe-se que as expressões acima mostram o valor terminal das opções no instante τ. O problema reside na avaliação do preço em momentos anteriores ao vencimento (t, para $t \leq \tau$). $t = 0$ será considerado como o momento do fechamento do contrato da opção e o seu preço será c_0. Neste caso, S_τ passa a ser uma variável aleatória, e é necessário considerar uma hipótese sobre a distribuição de probabilidade dos preços (S_τ) ou taxas de retorno futuros do valor mobiliário subjacente. Além disso, é preciso determinar uma taxa de desconto apropriada para antecipar os resultados financeiros da opção (valor intrínseco).

2.2.1.1 Condição de contorno

Um elemento fundamental na avaliação dos contratos de opções é a chamada **condição de contorno** (*boundary condition*) ou **valor intrínseco**. Pode-se dizer que a condição de contorno é o núcleo da expressão matemática que define uma opção. Merton (1973) definiu algumas relações importantes envolvendo a precificação de opções para o investidor racional, entre as quais:

a) $c(S_t; t; k) \geq 0$: uma opção de compra européia tem sempre valor positivo;

b) $c(S_t; t; k) \leq S_t$: uma opção de compra européia não pode custar mais que o preço do valor mobiliário subjacente;

c) $c(S_t;t;k_1) \geq c(S_t;t;k_2)$: uma opção de compra com preço de exercício menor (k_1) deve ter um valor igual ou superior a outra com menor preço de exercício (k_2), isto é, $k_1 \geq k_2$;

d) $c(S_t;t;k_\tau) \geq \max\{0;S_t - k_\tau \exp[-r_f(\tau - t)]\}$: o preço de uma opção de compra européia deve ser igual ou superior ao seu valor intrínseco;

e) $c(S_t;t = \infty;k = 0) = S_t$: o preço do valor mobiliário subjacente corresponde a uma opção de compra perpétua com preço de exercício igual a zero.

Cada espécie de opção possui uma condição de contorno característica baseada no resultado financeiro final proporcionado pelo direito de exercício. A opção de compra européia simples ordinária possui como condição de contorno, o valor $\max\{0;S_\tau - k_\tau\}$. A opção de venda, por sua vez, tem seu valor como $\max\{0;k_\tau - S_\tau\}$.

2.2.2 História das fórmulas de determinação do preço racional da opção européia simples ordinária sobre ação

A história do desenvolvimento das fórmulas de formação do preço da opção européia simples ordinária foi adaptada de Yang (1997).

2.2.2.1 Modelos pré-Black-Scholes-Merton (1900-1964)

2.2.2.1.1 Modelo de Bachelier (1900)

Conforme Yang (1997, p. 8), o modelo de formação do preço da opção mais antigo de que se tem conhecimento é o de Louis Bachelier de 1900. Ao adotar um movimento browniano com $\mu = 0$ (*drift*) e variância a termo (σ^2) constante para o processo do preço da ação, ele obteve a seguinte fórmula para determinar o valor esperado da opção de compra européia simples ordinária:

$$c_0 = (s_0 - k_\tau)\Phi\left(\frac{s_0 - k_\tau}{\sigma\sqrt{\tau}}\right) + \sigma\sqrt{\tau}\phi\left(\frac{s_0 - k_\tau}{\sigma\sqrt{\tau}}\right)$$

em que s_0, k_τ e τ representam o preço do valor mobiliário subjacente, o preço de exercício e o prazo, respectivamente, $\Phi(\cdot)$ é a função de distribuição da normal padronizada e $\phi(\cdot)$ é a função de densidade de probabilidade da normal padronizada.

As objeções ao modelo de formação do preço da opção de Bachelier são:

1. Claramente o valor esperado da opção não foi trazido a valor presente.
2. A hipótese de movimento browniano aritmético, como um descritor do movimento do preço do valor mobiliário subjacente, permite que haja uma probabilidade de ocorrerem preços negativos e preços de opções maiores que seus respectivos preços dos valores mobiliários subjacentes para um período grande (τ grande).
3. A hipótese de que a variação esperada de preço do valor mobiliário subjacente é zero sugere taxas de juros nulas, portanto, ignora o valor positivo do dinheiro no tempo e aversão à incerteza.

Apesar dessas objeções, a fórmula funciona bem para formar o preço de opção com prazo curto e taxa de desconto pequena.

2.2.2.1.2 Modelo de Sprenkle (1961)

Sprenkel assumiu a distribuição lognormal para o preço do valor mobiliário subjacente com média e variância constantes. Portanto, as duas últimas objeções ao modelo de Bachelier são eliminadas. O valor esperado da opção de compra européia ordinária na data de exercício é:

$$c_0 = s_0 \exp(\alpha\tau)\Phi\left[\frac{\ln\left(\frac{s_0}{k_\tau}\right) + \left(\alpha + \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}\right] - k_\tau\,(1-\pi)\,\Phi\left[\frac{\ln\left(\frac{s_0}{k_\tau}\right) + \left(\alpha - \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}\right]$$

em que α é a taxa de retorno esperada da ação e π é um ajuste devido ao preço de alavancagem do mercado. Note-se novamente que o valor da opção não foi trazido a valor presente.

2.2.2.1.3 Modelo de Boness (1964)

Esta abordagem utiliza uma medida de probabilidade subjetiva e uma taxa de desconto ajustada ao risco também subjetiva.

Conforme Lee (1997, p. 31), Boness em 1964 assumiu também uma distribuição lognormal como sendo a distribuição do preço da ação e que a taxa de desconto

esperada para a opção de compra pode ser aproximada pela taxa de desconto esperada da ação. A fórmula do preço da opção é:

$$c_0 = s_0\Phi\left[\frac{\ln\left(\frac{s_0}{k_\tau}\right)+\left(\alpha+\frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}\right] - k_\tau \exp(-\alpha\tau)\Phi\left[\frac{\ln\left(\frac{s_0}{k_\tau}\right)+\left(\alpha-\frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}\right]$$

Note-se que esta fórmula torna-se a de Black-Scholes-Merton se $\alpha = r_f$, isto é, se a taxa de desconto for a taxa certa.

Esta fórmula fornece o preço da opção de clarividência (valor esperado da informação perfeita) do preço futuro da ação para um decisor neutro à incerteza. Como a clarividência só antecipa a resolução da incerteza, a taxa de desconto esperada continua sendo a da ação.

Este preço da ação é chamado também de **preço da opção pelo método do fluxo de caixa descontado modificado** (*modified discount cash flow – MDFC*) (Lee (1997, p. 30-3)). Este preço será denominado c_{MDFC}.

Conforme Merton (1973), *apud* Lee (1997, p. 43), quanto maior a taxa de desconto, maior o preço da opção. Portanto, o c_{MDFC} é o limite superior do preço da opção para um decisor neutro à incerteza. E, quando há possibilidade de *hedging*, o preço da opção atinge seu valor mínimo, fazendo $\alpha = r_f$, que coincide com o valor dado pela fórmula de Black-Scholes-Merton.

Esta abordagem foi utilizada por Lee (1997, p. 1-3) para introduzir a formação de preços de opções não-hedgeáveis, pois, quando as opções não são hedgeáveis, elas não podem ter seus preços formados pelas técnicas convencionais.

2.2.2.2 Modelo de Black-Scholes-Merton (1973)

> *"Black later said: 'A key part of the option paper I wrote with Myron Scholes was the arbitrage argument for deriving the formula. Bob [Merton] gave us that argument. It should probably be called the Black-Merton-Scholes paper."*
>
> Amram & Kulatilaka (1999, p. 32)

Conforme Rubinstein (1994, p. 772), o modelo para formação do preço da opção européia simples ordinária de Black-Scholes-Merton, publicado em 1973, é visto como um dos modelos que obtiveram maior sucesso nas Ciências Sociais e talvez

a sua fórmula de apreçamento da opção européia tenha sido a mais amplamente utilizada na história da humanidade.

Ao acrescentar a estratégia da carteira replicante (*hedging* perfeito dinâmico) às hipóteses de Boness, tem-se a fórmula de Black-Scholes-Merton:

$$c_0 = s_0\Phi\left[\frac{\ln\left(\frac{s_0}{k_\tau}\right)+\left(r_f+\frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}\right] - k_\tau \exp\left(-r_f\tau\right)\Phi\left[\frac{\ln\left(\frac{s_0}{k_\tau}\right)+\left(r_f-\frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}\right]$$

Surge, em substituição à medida de probabilidade subjetiva ou real, a medida de probabilidade neutralizadora do preço da incerteza, que é única em mercado completo.

Esta fórmula funciona razoavelmente bem para determinar preços de opções do tipo européias simples ordinárias sobre ações, por exemplo. No entanto, ela apresenta vários problemas para formar preço de opção sobre taxa de juros.

Para ilustrar estas dificuldades será proposto o seguinte problema: uma opção de compra européia ordinária de três meses sobre um ativo de renda fixa (*bond*), sem cupom de três anos. O valor na maturidade do ativo de renda fixa é R$100,00 e o preço atual do ativo-base (*bond*) é R$75,13; a taxa certa ao ano é de 10% no período de três anos e a volatilidade esperada do preço é 4%. Suponha-se que o preço de exercício é R$120,00. Qual o valor da opção sobre a taxa de juros?

Claramente, não é necessário nenhum modelo de formação do preço da opção para obter a resposta certa. Este ativo de renda fixa jamais atingirá um valor superior a R$100,00, pois este é o seu valor na maturidade. Como o preço de exercício é R$120,00, então a opção nunca será exercida, e seu valor é zero. No entanto, se um comprador de opção estiver armado com a fórmula de Black-Scholes-Merton, ele irá obter o preço de R$5,16!

Considerando as hipóteses do modelo de Black-Scholes-Merton, esse erro não deve causar surpresa. Há três hipóteses subjacentes ao modelo de Black-Scholes-Merton, que limitam seu uso na formação do preço de opções sobre taxas de juros.

1. A medida de probabilidade neutralizadora do preço da incerteza do preço do valor mobiliário subjacente é lognormal, permitindo, assim, que o preço tenha

alguma probabilidade de assumir qualquer valor positivo. A única maneira de permitir que o preço do *bond* seja maior que seu valor na maturidade é permitir valores negativos para a taxa de juros. Como isso não é possível, na prática, preços sem sentido de opções podem surgir.

2. A segunda hipótese do modelo de Black-Scholes-Merton, normalmente utilizada, é que a taxa de juros de curto prazo é constante durante a vida da opção (existência do banco ideal constante). No entanto, o preço de uma opção sobre taxa de juros muda quando a taxa de juros muda. Logo, tratar a taxa de juros como constante e simultaneamente admitir que o preço do ativo de renda fixa é uma variável aleatória ocasiona uma inconsistência lógica.

3. A terceira hipótese é que a variância dos preços é constante durante a vida da opção. É bem conhecido o fato de que a volatilidade do preço de ativo de renda fixa (*bond*) diminui com o decorrer do tempo porque o preço do *bond* deve ser igual ao seu valor de face na data de sua maturidade. A volatilidade do preço do ativo de renda fixa tende a zero quando a data se aproxima da de maturidade.

O resultado mostra basicamente que o prêmio único e livre de arbitragem da opção européia simples ordinária é composto por valor tempo e valor intrínseco.

O valor tempo é resultante da preferência temporal, considerando a possibilidade de *hedging* perfeito do valor intrínseco que implica o uso da taxa de juros certa para o cálculo do valor presente.

2.2.2.3 Modelos pós-Black-Scholes-Merton (até 1998)

Segue abaixo uma relação de teses de doutorado recentes sobre determinação de preço de opções com abordagens mais gerais e alternativas:

- Gulko (1998) – Yale University e Yang (1997) – The Florida State University

  Os autores aplicam a Teoria da Informação de C. E. Shannon e o princípio da máxima entropia de E. T. Jaynes na formação do preço de opções.

Gulko (1998) e Buchen & Kelly (1996) utilizam o princípio da máxima entropia da Teoria da Informação para obter a medida de probabilidade neutralizadora do preço da incerteza. A diferença entre os dois trabalhos reside no fato de Gulko (1998) impor como conhecidos os dois primeiros momentos do preço do ativo-base como restrições, enquanto o outro impõe como conhecidos os preços das opções de compra como restrições. Gulko (1998) exibe um modelo com expressões exatas para as opções européias simples ordinárias sobre ativos de renda variável (ação) e fixa (*bond*) e introduz um método alternativo de extração da medida de probabilidade neutralizadora do preço da incerteza.

Yang (1997) obtém expressões quase exatas cujas aproximações são as fórmulas encontradas por Gulko (1998). Yang (1997) elimina completamente a necessidade do cálculo estocástico para a obtenção de fórmulas (quase) exatas de formação de preços de opções dos tipos européia e americana.

- Cheung (1997) – Stanford University

O seu modelo macroscópico substitui a estimação da volatilidade pelo volume negociado para captar a componente subjetiva do mercado e mostra, em sua tese, que os resultados são melhores que os obtidos pelo modelo baseado apenas no componente objetivo do mercado.

O modelo macroscópico de Cheung mostra como as naturezas objetivas e subjetivas da tomada de decisão humana afetam a formação do preço de valores mobiliários. Desta forma o autor cria uma caracterização mais precisa do processo de formação do preço da ação para desenvolver um modelo de formação de preço de opções mais robusto e melhor que o de Black-Scholes-Merton. O autor decompõe o retorno do ativo num componente objetivo e subjetivo derivado da distinção dos conceitos de risco e incerteza de Knight (1964) e da hipótese de ruído de transação (*trading noise hypothesis*) de French & Roll (1986). O modelo incorpora o volume transacionado, relacionando-o com o componente subjetivo do retorno. O volume transacionado possui o mesmo papel da volatilidade no modelo tradicional,

com a vantagem de aquele ser observável. O modelo suporta não-normalidade dos retornos, autocorrelação dos retornos, persistência da volatilidade dos retornos, relação contemporânea positiva entre volume negociado e volatilidade, e relação contemporânea inversa entre volume negociado e correlação serial dos retornos, e retornos negativos na segunda-feira. Portanto, o modelo, segundo o autor, é útil para a gestão do risco e previsão de volatilidade, e uma alternativa metodológica em Finanças. O modelo necessita da hipótese de mercado completo.

O modelo continua necessitando da abordagem de processo estocástico e da hipótese de mercado completo.

Além disso, para calcular o preço da opção européia, são necessárias duas etapas. A primeira utiliza programação matemática para estimar a volatilidade. Na segunda etapa, é necessário utilizar o método de simulação de Monte Carlo.

- Lee (1997) – Stanford University

O autor apresenta o conceito de opção de novo negócio (*joint ventures*, fusões, aquisições etc.) como um tipo especial de opção de crescimento. A tese apresenta uma teoria de formação de preço destas opções reais e analisa as decisões estratégicas de escolhê-las. Com algumas modificações, a teoria de opções compostas é utilizada para esta análise estratégica.

- Herath (1998) – Auburn University

A autora segue a abordagem bayesiana de formação do preço da opção européia simples ordinária sem levar em conta a condição de não-arbitragem. Por outro lado, ela enfatiza a aplicação do conceito de clarividência (valor esperado da informação perfeita) e do valor esperado da informação imperfeita na formação do preço do que ela designa de **quase-opção** ou de **opção real**. Os conceitos de oportunidade de replicação e de revisão seqüencial junto ao valor da informação proporcionam a criação de uma abordagem bayesiana de formação do preço da opção real. A derivação da fórmula de Black-Scholes-Merton, a partir desta abordagem, encontra-se em sua tese. A tese desta

autora está grandemente baseada nas obras de H. Bierman Jr., R. Schlaifer e L. Trigeorgis.

- Samperi (1998) – New York University

  O autor formaliza o conceito de mercado completo, medida de probabilidade neutralizadora do preço da incerteza , utilidade, risco etc. a partir da Teoria da Informação. O conceito de divergência tem um papel central em sua teoria e dele praticamente todos os resultados são derivados.

- Yu (1998) – Princeton University

  O autor mostra como obter o preço livre de arbitragem de um ativo, mesmo quando o mercado não é perfeito, isto é, quando se considera que há custos de transação.

Escolheu-se analisar nesta tese os resultados de Gulko (1998 e 1998a) e Yang (1997) devido à conexão, neles presentes, entre a Teoria da Informação e a Ciência do Investimento.

2.2.3 Demonstração da fórmula de Black-Scholes-Merton

As hipóteses normalmente utilizadas para a formação do preço de derivativos são: mercado perfeito, mercado eficiente (movimento browniano geométrico do preço da ação), mercado completo e investidores racionais. A condição da ausência de arbitragem deriva diretamente da condição de mercado completo. Há classicamente dois tipos de mercado completo. Um uniperíodo e outro multiperíodo. A demonstração da fórmula de Black-Scholes-Merton depende da existência de uma estratégia replicante dinâmica e autofinanciável de todos os possíveis valores intrínsecos do derivativo. Esta carteira replicante completa o mercado de forma dinâmica. Por isso, o mercado é chamado de **dinamicamente completo** ou **completo multiperíodo**. No entanto, para a determinação da medida de probabilidade neutralizadora do preço da incerteza é necessário a introduzir a hipótese da eficiência de mercado, que acrescenta uma medida de probabilidade real ou subjetiva. Isso claramente viola o princípio da parcimônia, pois, ao longo da tese, foi mostrado que a medida de probabilidade neutralizadora do preço da

incerteza independe da distribuição de probabilidade real ou subjetiva do valor mobiliário subjacente na data de exercício. A suposição de que o preço do valor mobiliário subjacente tem o comportamento regido pelo passeio aleatório (versão discreta) ou movimento browniano geométrico (versão contínua) introduz indevidamente a medida de probabilidade real ou subjetiva na solução do problema. Quando essa suposição é introduzida, a medida de probabilidade real ou subjetiva do preço do valor mobiliário ação é a lognormal. Por meio de uma mudança da medida de probabilidade, a medida equivalente também é lognormal. Apesar da demonstração conter estes dois artifícios matemáticos para a determinação da medida de probabilidade neutralizadora do preço da incerteza, sua demonstração tradicional completa será feita para que sejam entendidos exatamente todos os passos e, desta forma, torne-se possível uma crítica fundamentada do modelo de B-S-M.

2.2.3.1.1 Mercado completo multiperíodo ou dinamicamente completo

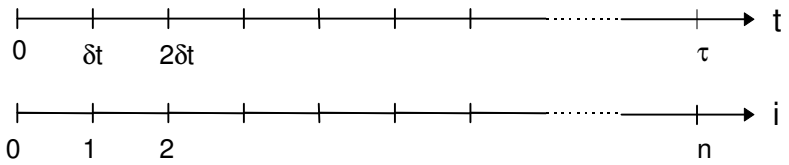
A possibilidade de negociação seqüencial dos ativos da carteira replicante (abordagem multiperíodo com possibilidade de reestruturação da carteira a cada período) para realização do *hedging* perfeito dos valores intrínsecos da opção pode, em alguma medida, servir como substituto para um conjunto completo de ativos de Arrow-Debreu, conforme Cox & Rubinstein (1985, cap. 8).

O livro intitulado *Financial Calculus*, de Baxter & Rennie (1998), possui uma notação matemática inconsistente para determinar o preço dos derivativos a partir de uma estrutura binária. Para facilitar a compreensão do texto, preferiu-se mudar alguns elementos. Há agora uma distinção clara entre valor numérico, variável aleatória e processo estocástico. Melhorou-se a identificação do posicionamento na árvore. Eliminou-se a abordagem de martingale[18] para a demonstração do teorema da representação binária e carteira replicante.

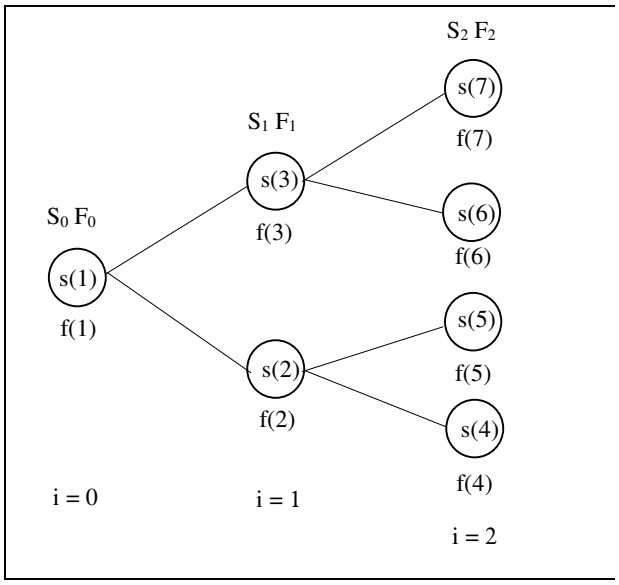
---

[18] **Martingale** é o nome genérico de qualquer sistema para ganhar na roleta. O termo provém do nome de uma rédea (**gamarra**, em português) usada para impedir o cavalo de virar a cabeça. Assim, o martingale evitaria o desastre, ao

Seja um intervalo de tempo δt, formado pelo instante 0, em que surge o derivativo e o instante τ, que é sua data de vencimento. Dividindo-se este intervalo em n subintervalos de δt duração, tem-se n segmentos δt. O final de cada intervalo será um instante de tempo i, em que i varia de 0 a n, segundo uma outra escala de tempo:



Em cada instante i, ocorrem 2^i estados da natureza representando o número de estados possíveis de preços do ativo objeto. Cada estado da natureza é descrito por um "nó" na estrutura binária, que terá como indicador a letra j. Portanto, $s(j)$ será o preço do ativo no nó j:



---

reparar as perdas [Seabra (1992, p. 35-9)]. Esta idéia foi formalizada matematicamente por J. Doob no seu livro *Stochastic Processes*, publicado pela John Wiley & Sons (New York), em 1953. Uma formalização atualizada encontra-se em Ross (1996). Baxter & Rennie (1998) aplicam martingale na formalização da formação do preço de derivativos.

Portanto, de acordo com a árvore acima, observa-se:

S_i: variável aleatória preço do ativo no instante i;

$s(j)$: preço do ativo objeto no nó j.

E para cada instante i e em cada nó j, têm-se os preços dos derivativos associados:

F_i: variável aleatória preço do derivativo no instante i;

$f(j)$: preço do derivativo no nó j.

O conjunto de todas as variáveis aleatórias indexadas no instante i é denominado processo estocástico. Para um intervalo de tempo entre 0 e n, verificam-se todas as trajetórias possíveis de preços do ativo $s(j)$ e do derivativo $f(j)$. O processo descrito pela variável preço pode ser representado por:

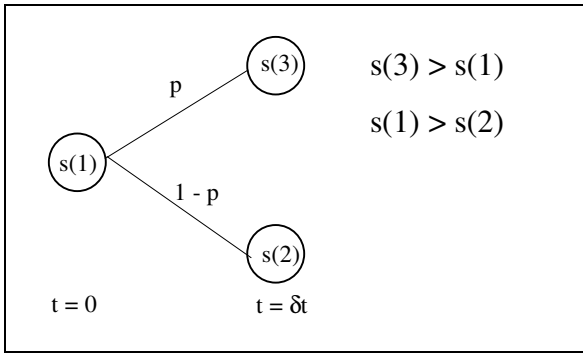
$$\{S_i; i = 0,1,2,\ldots,n\}$$

Dado um nó j qualquer, é possível descobrir o instante de tempo em que se encontra, pois j varia de acordo com o intervalo:

$$j \in \left[2^i \ldots 2^{i+1} - 1\right]$$

2.2.3.1.1.1 Ramo binário em um processo discreto

Parte-se do pressuposto que os preços dos ativos apresentam dois valores somente. O preço pode subir para $s(3)$ ou cair para $s(2)$, decorrido um intervalo de tempo δt.



Existe um título sem incerteza b_0, remunerado a uma taxa r_f em um processo contínuo de capitalização.

2.2.3.1.1.2 Construção de derivativos para um período a partir de uma carteira formada por ações e bonds

De acordo com Baxter & Rennie (1998), é possível encontrar o preço único e livre da arbitragem de um derivativo a partir de uma carteira formada pelo ativo-objeto e *bonds* que repliquem o valor deste derivativo na data de vencimento. Seja uma carteira π_0 composta por $\phi(1)$ ativos e $\psi(1)$ *bonds* no nó $j = 1$, de que resulte um valor exatamente igual ao do derivativo $f(j)$ nos nós $j = 2$ ou $j = 3$:

$$f(2) = \phi(1)s(2) + \psi(1)b_1$$
$$f(3) = \phi(1)s(3) + \psi(1)b_1$$

Sabe-se, também, que o valor da carteira replicante é expressa pela seguinte fórmula:

$$v_0(\pi_0) = \phi(1)s(1) + \psi(1)b_0$$

A simples existência da carteira replicante de um derivativo em cada nó do processo permite inferir, pelo princípio da não-arbitragem, que o preço do derivativo, na data de lançamento, é único e deve ser o próprio valor da carteira replicante nessa data. Logo, esse é o preço racional do derivativo.

Pelo princípio da arbitragem, o valor inicial v_0 desta carteira deve ser exatamente igual ao preço do derivativo $f(1)$:

$$f(1) = v_0(\pi_0).$$

Resolvendo o sistema de equações dos preços $f(2)$ e $f(3)$, podemos encontrar as quantidades $\phi(1)$ e $\psi(1)$ que representam as quantidades de ações e *bonds* da carteira π_0 que sintetizam o valor do derivativo:

$$\phi(1) = \frac{f(3) - f(2)}{s(3) - s(2)} \text{ e } \psi(1) = \frac{1}{b_1}\left[f(3) - \frac{f(3) - f(2)}{s(3) - s(2)}s(3)\right] \text{ (exp. 1)}$$

Portanto, ao adquirir $\phi(1)$ ações e $\psi(1)$ *bonds*, quaisquer que sejam os preços finais do ativo-objeto S_1, a carteira reproduzirá o valor do derivativo para cada estado da natureza.

2.2.3.1.1.3 A medida de probabilidade neutralizadora do preço da incerteza $\mathbb{Q}$

A partir da fórmula do preço do derivativo v_0, pode-se substituir os termos $\phi(1)$ e, $\psi(1)$ pelos resultados encontrados:

$$\begin{aligned} f(1) &= \phi(1)s_0 + \psi(1)b_0 \\ &= s(1)\frac{f(3)-f(2)}{s(3)-s(2)} + \exp(-r_f\delta t)\left[f(3) - \frac{f(3)-f(2)}{s(3)-s(2)}s(3)\right] \end{aligned}$$

Seria interessante se houvesse uma maneira de escrever esta expressão de uma maneira simplificada. Escrevendo todos os termos sobre um mesmo denominador, tem-se:

$$f(1) = \frac{s(1)\exp(r_f\delta t)[f(3)-f(2)] - s(3)[f(3)-f(2)] + f(3)[s(3)-s(2)]}{\exp(r_f\delta t)[s(3)-s(2)]}$$

Na expressão anterior, será acrescido e subtraído um termo com valor igual a $s(2)[f(3)-f(2)]$:

$$f(1) = \frac{\left[s(1)\exp(r_f\delta t) - s(2)\right][f(3)-f(2)] + [s(2)-s(3)][f(3)-f(2)] + f(3)[s(3)-s(2)]}{\exp(r_f\delta t)[s(3)-s(2)]}$$

Para simplificar a expressão acima, será criada uma variável q descrita por:

$$q = \frac{s(1)\exp(r_f\delta t) - s(2)}{s(3)-s(2)}$$

Substituindo, na expressão anterior, e simplificando os demais termos, tem-se:

$$\begin{aligned} f(1) &= \exp(-r_f\delta t)q[f(3)-f(2)] + \exp(-r_f\delta t)f(2) \\ &= \exp(-r_f\delta t)[(1-q)f(2) + qf(3)] \end{aligned}$$

Observe-se que o valor do derivativo passou a ser representado como uma esperança dos possíveis preços do derivativo, ponderado por uma probabilidade q, trazidos a valor presente por uma taxa certa. A esperança com as probabilidades q e $1-q$ pode ser representada pelo operador esperança na medida de probabilidade $\mathbb{Q}$. Logo, o valor do derivativo na data 0 nada mais é do que o valor presente pela taxa certa da esperança da variável F_1 na medida $\mathbb{Q}$:

$$f(1) = \exp(-r_f\delta t)E_{\mathbb{Q}}[F_1].$$

Note-se que este é o mesmo resultado do preço racional quando se impõe a existência do mundo com preço da incerteza neutralizado.

2.2.3.1.1.4 Propriedades da medida de probabilidade neutralizadora do preço da incerteza $\mathbb{Q}$

Existe alguma razão para acreditar que q possa ter as características de uma probabilidade? Se o valor de q estiver entre o intervalo 0 e 1, é possível imaginá-lo como uma probabilidade.

Partindo da expressão de q anterior e sabendo que $s(3) > s(2)$, por definição, devem-se considerar as seguintes hipóteses:

a) $q < 0$. Se $q < 0 \Rightarrow s(1)\exp(r_f \delta t) < s(2)$. Se o menor dos preços é superior ao custo do financiamento, é possível ter um lucro certo. Portanto, os investidores vão se endividar cada vez mais para adquirir o ativo s. Logo, $q \geq 0$.

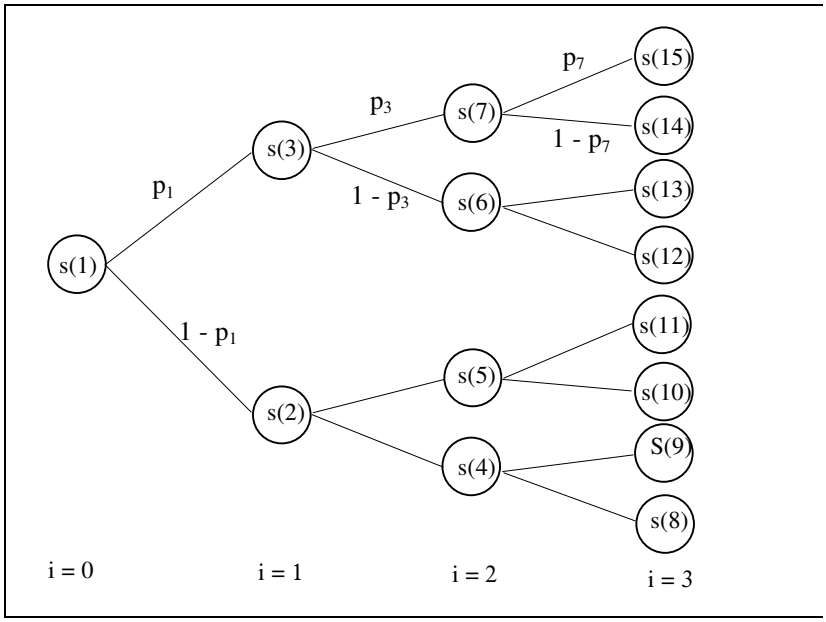
b) $q > 1$. Se $q > 1 \Rightarrow s(1)\exp(r_f \delta t) > s(3) > s(2)$. Se qualquer um dos preços é inferior ao preço da ação corrigido, é possível vender o ativo a descoberto e aplicar os recursos em *bonds*. Qualquer que seja o preço final, o investidor terá um ganho certo. Portanto, $q \leq 1$.

Como resultado final, sabe-se que q é um número que se encontra no intervalo compreendido entre 0 e 1.

2.2.3.1.1.5 A árvore binária

A representação de um derivativo requer mais de um estágio ou período de tempo se se quiser generalizar o modelo de formação de preço para um intervalo de tempo qualquer. Uma maneira de expandir o ramo binário é acrescentar ramos aos novos valores que surgem a cada etapa:

Figura 1: Árvore binária



Na árvore binária acima, podem ser observadas as seguintes características:

- preço inicial $s(1)$;
- a cada período surgem dois novos valores em cada nó j;
- 2^i estados da natureza para cada instante i;
- p_j: probabilidade do preço subir associada ao nó j.

Embora a estrutura da árvore pareça simples, à medida que δt tende a zero, o modelo vai se aproximando de um modelo mais preciso.

Cada nó j representa um estado da natureza diferente para o ativo S_i. E, para cada preço $s(j)$, existe um preço do derivativo representado por $f(j)$.

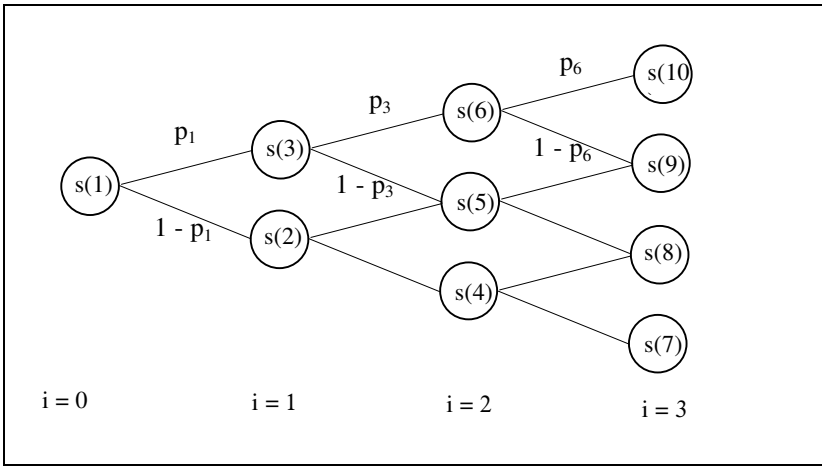
2.2.3.1.1.6 A árvore binária recombinante

Uma forma de simplificar a árvore binária é utilizar o conceito de árvore recombinante. Neste tipo de modelo, a quantidade de nós que surgem são exatamente iguais aos da árvore binária. Contudo, como particularidade, o valor do nó inferior, que se ramifica de um nó $j+1$ coincide com o valor do nó superior

proveniente do nó j. Portanto, a quantidade de nós com valores diferentes em cada etapa i é inferior ao de uma árvore binomial comum.

Na árvore recombinante, os nós crescem em progressão aritmética, conforme visto na figura abaixo:

Figura 2: Árvore binária recombinante

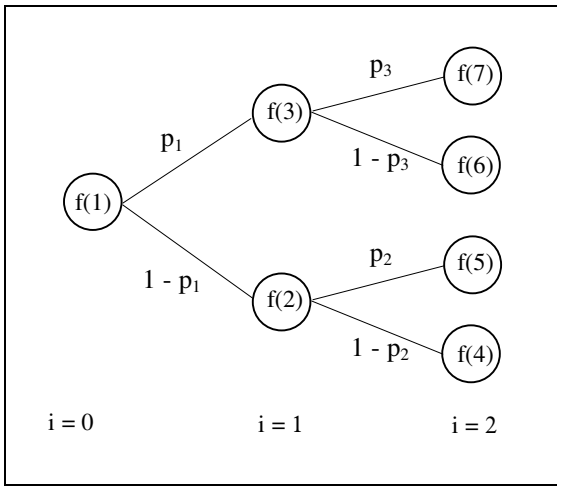


Existe uma relação entre o número do nó do ativo $s(j)$ e sua posição na árvore recombinante. O valor de j varia de acordo com i, conforme a expressão:

$$j \in \left[1 + \frac{i(i+1)}{2} \ldots \frac{(i+1)(i+2)}{2}\right]$$

2.2.3.1.1.7 Indução de trás para diante

No exemplo anterior, o valor do derivativo foi calculado a partir dos seus preços finais $f(2)$ e $f(3)$ e da medida $\mathbb{Q}$. Expandindo este resultado para mais de um período, pode-se chegar no valor inicial do derivativo por meio de um processo denominado **indução de trás para diante**:



Generalizando o resultado, o valor do derivativo em cada nó j pode ser escrito como:

$$f(j) = \exp(-r_f \delta t)\big[q_j f(2j+1) + (1-q_j) f(2j)\big]$$

Em cada nó j existe uma probabilidade q_j específica:

$$q_j = \frac{s(j)\exp(r_f \delta t) - s(2j)}{s(2j+1) - s(2j)}$$

Observe-se que pela figura anterior e pelas expressões, é preciso conhecer a distribuição dos preços ao longo do tempo. Sem o valor dos ativos $s(j)$ em cada nó j, não se calculam os valores q_j e, por conseqüência, $f(j)$.

É possível obter o valor do derivativo no nó inicial em função dos preços finais do derivativo. A lógica é a mesma que a empregada para se calcular o derivativo em um ramo. O valor do derivativo é a esperança dos preços finais na medida $\mathbb{Q}$, trazida a valor presente por uma taxa certa:

$$\begin{aligned} f(1) &= \exp(-r_f 2\delta t)\big[q_1 q_3 f(7) + q_1 \overline{q}_3 f(6) + \overline{q}_1 q_2 f(5) + \overline{q}_1 \overline{q}_2 f(4)\big] \\ &= \exp(-r_f 2\delta t) E_{\mathbb{Q}(2\delta t)}\big[F_2\big] \end{aligned}$$

2.2.3.1.1.8 Carteira replicante auto-financiável

Seja uma carteira replicante formada pela composição de um ativo sem incerteza e *bonds*. No período posterior, o valor da carteira replicante assume a seguinte expressão:

$$F_1 = V_1(\pi_0) = \phi(1)S_1 + \psi(1)b_1$$

Substituindo as quantidades ϕ e ψ determinadas anteriormente, ter-se-á:

$$V_1(\pi_0) = S_1 \frac{f(3) - f(2)}{s(3) - s(2)} + \frac{1}{b_1}\left[f(1) - \frac{f(3) - f(2)}{s(3) - s(2)} s(1)\right] b_1$$

Simplificando a expressão, o valor da carteira pode ser expresso como:

$$V_1(\pi_0) = \frac{[S_1 - s(1)][f(3) - f(2)]}{s(3) - s(2)} + f(1)$$

Observe-se que o valor do ativo S pode assumir dois valores na data $i = 1$. Quando o ativo S for igual a $s(3)$ e $v_1(\pi_0) = f(3)$:

$$v_1(\pi_0) = [s(3) - s(1)]\phi(1) + f(1) \Rightarrow f(3) - f(1) = [s(3) - s(1)]\phi(1)$$

E, quando o ativo S for igual a $s(2)$ e $v_1(\pi_0) = f(2)$:

$$v_1(\pi_0) = [s(2) - s(1)]\phi(1) + f(1) \Rightarrow f(2) - f(1) = [s(2) - s(1)]\phi(1)$$

Portanto:

$$\Delta F_i = \Phi_i \Delta S_i$$

Desenvolvendo a expressão acima para diversos períodos de tempo, pode-se expressar a variação do valor do derivativo como (Teorema da Representação Binária):

$$F_i = F_0 + \sum_{k=1}^{i} \phi_k \Delta S_k$$

No instante $i = 1$, a carteira precisa ser modificada para replicar o valor do derivativo no instante posterior. Logo, é preciso alterar as quantidades ϕ e ψ para que a carteira replique o valor do derivativo:

$$V_1(\Pi_1) = \Phi_2 S_1 + \Psi_2 b_1 \ \text{(exp. 2)}$$

Utilizando a expressão 1 para determinar as quantidades ϕ e ψ no instante $i = 1$, vê-se que:

$$\Phi_2 = \frac{F_2 - F_1}{S_2 - S_1} \ \text{e} \ \Psi_2 = \frac{1}{b_1}\left[F_1 - \frac{F_2 - F_1}{S_2 - S_1} S_1\right]$$

Substituindo as quantidades acima na expressão 2, tem-se:

$$
\begin{aligned}
V_1(\Pi_1) &= \frac{F_2 - F_1}{S_2 - S_1} S_1 + \left[F_1 - \frac{F_2 - F_1}{S_2 - S_1} S_1 \right] \\
&= F_1
\end{aligned}
$$

Note-se que o resultado acima é surpreendente. A carteira replicante formada no instante $i = 1$ possui o mesmo valor da carteira replicante formada em $i = 0$, ajustada pela valorização de seus ativos, isto é:

$$V_1(\pi_0) = V_1(\Pi_1) .$$

2.2.3.1.1.9 Comportamento do preço do ativo em uma árvore recombinante: a abertura para modelos contínuos

Seja uma árvore recombinante formada pelas trajetórias dos preços do ativo-objeto em quatro estágios do tempo. A adoção deste tipo de árvore binária deve-se a suas vantagens com computacionais [Dos Anjos & Paiva (1998, p. 34-5)].

$$s(3) = s(1)u$$

ou

$$s(2) = s(1)d$$

em que:

$$u > \exp(r_f \delta t) > d > 0$$

Os valores de u e d não são únicos mesmo na ausência de arbitragem [Dos Anjos & Paiva (1998, p. 30)].

Conforme Avellaneda (1995-6), *apud* Dos Anjos e Paiva (1998, p. 30), tem-se que os valores permitidos para u e d são:

$$u = \frac{\exp(\kappa\sqrt{\delta t} + r_f \delta t)}{(1-q)\exp(-\kappa\sqrt{\delta t}) + q\exp(\kappa\sqrt{\delta t})}$$

$$d = \frac{\exp(-\kappa\sqrt{\delta t} + r_f \delta t)}{(1-q)\exp(-\kappa\sqrt{\delta t}) + q\exp(\kappa\sqrt{\delta t})}$$

onde

$$\sigma = \sqrt{\frac{\operatorname{var}\left[\ln\left(\frac{S_\tau}{s(1)}\right)\right]}{\tau}}$$

$$q = \frac{1}{2}\left[1 \pm \sqrt{1 - \left(\frac{\sigma}{\kappa}\right)^2}\right]$$

$$\kappa \geq \sigma$$

Estes valores de u e d fazem com que o valor assintótico de $E\left[\frac{\ln\left(\frac{S_\tau}{s(1)}\right)}{\tau}\right] = \mu$ (*drift*) seja $r_f - \frac{\sigma^2}{2}$. Lembrar que $E_{\mathbb{P}}\left[\ln\left(\frac{S_\tau}{s(1)}\right)\right] = \exp\left[\left(\nu + \frac{\sigma^2}{2}\right)\tau\right]$ se o preço da ação possui a distribuição lognormal com parâmetros ν e σ^2 na data de exercício e $E_{\mathbb{Q}}\left[\ln\left(\frac{S_\tau}{s_0}\right)\right] = \exp\left(r_f \tau\right)$ sob a hipótese de mundo com preço da incerteza neutralizado.

Supondo que o preço inicial seja $s(1)$ e, no instante seguinte, o preço do ativo assuma dois possíveis valores, conforme Baxter & Rennie (1998):

$$s(3) = s(1)\exp(\mu\delta t + \sigma\sqrt{\delta t})$$

ou

$$s(2) = s(1)\exp(\mu\delta t - \sigma\sqrt{\delta t})$$

sendo

$\mu = E\left[\frac{\ln\left(\frac{S_\tau}{s(1)}\right)}{\tau}\right]$ (*drift* ou esperança da taxa de retorno ao período) e $\sigma = \sqrt{\frac{\operatorname{var}\left[\ln\left(\frac{S_\tau}{s(1)}\right)\right]}{\tau}}$ (volatilidade a termo ou desvio-padrão da taxa de retorno ao período).

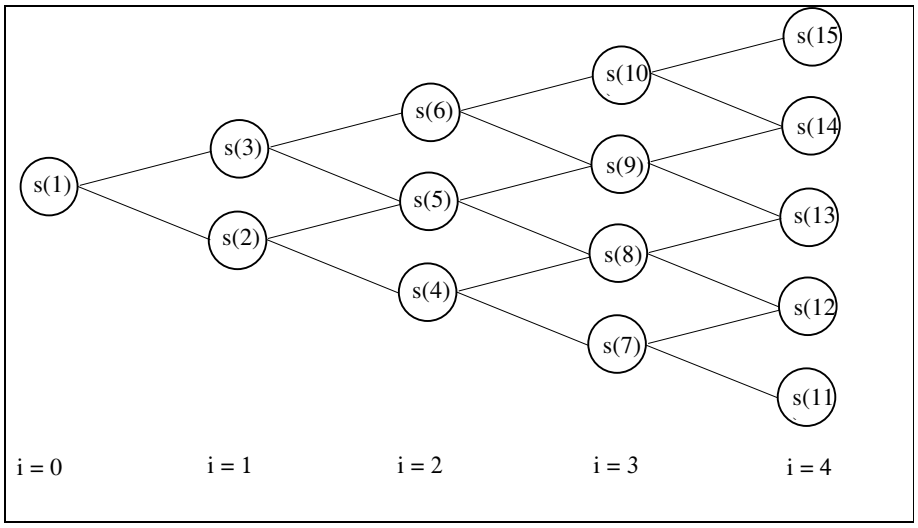
Note-se que a exponenciação evita que haja valor do preço negativo ou nulo e representa melhor a forma de crescimento real do preço. Além disso, observe-se que $\sigma\sqrt{\delta t}$ é muito maior que $\mu\delta t$. Por isso alguns autores utilizam o fator de

subida $u = \exp(\sigma\sqrt{\delta t})$ e de descida $d = \exp(-\sigma\sqrt{\delta t})$ [Luenberger (1998, p. 298)]. Isso equivale a fazer $ud = 1$.

Para simular um passeio aleatório dos preços do ativo, a probabilidade de subida será igual a ½.

Estas probabilidades reais ou subjetivas do comportamento do preço da ação derivam da hipótese de mercado eficiente.

Dessa forma, a árvore recombinante pode ser ilustrada como:



Os preços do ativo S_4 na data final, isto é, $i = 4$, podem assumir os seguintes valores:

Preços $\left(\delta t = \frac{\tau}{4}\right)$	**Probabilidade**
$s(15) = s(1)\exp(\mu\tau + 2\sigma\sqrt{\tau})$	$(1/2)^4 \,.\, 1$
$s(14) = s(1)\exp(\mu\tau + \sigma\sqrt{\tau})$	$(1/2)^4 \,.\, 4$
$s(13) = s(1)\exp(\mu\tau)$	$(1/2)^4 \,.\, 6$
$s(12) = s(1)\exp(\mu\tau - \sigma\sqrt{\tau})$	$(1/2)^4 \,.\, 4$
$s(11) = s(1)\exp(\mu\tau - 2\sigma\sqrt{\tau})$	$(1/2)^4 \,.\, 1$

Observe-se que S_4 é uma variável aleatória que possui cinco valores distintos e com probabilidades definidas por uma distribuição binomial:

$$S_4 = s(1)\exp\left[\mu\tau - \sigma\sqrt{\tau}\left(\frac{2X_4 - 4}{\sqrt{4}}\right)\right]$$

Na expressão acima X_4 é o número de vezes que o preço subirá. Podendo ser definido como uma soma de variáveis aleatórias independentes com distribuição de Bernoulli ($p = \frac{1}{2}$):

$Y_i = 1$, se o preço subir;

$Y_i = 0$, se o preço descer.

Portanto, a variável X_4 será escrita como:

$$X_4 = \sum_{i=1}^{4} Y_i$$

Logo, X_n tem uma distribuição binomial com os seguintes parâmetros:

$$X_n \sim \text{binomial}\left(\frac{n}{2};\frac{n}{4}\right)$$

Padronizando a variável X_n, tem-se:

$$Z_n = \frac{X_n - n/2}{\sqrt{n/4}} \sim N(0;1)$$

A expressão acima é válida quando n tende ao infinito. Para $n > 30$, Z_n é assintoticamente normal.

Logo, $\ln(S_n)$ tem distribuição normal com os seguintes parâmetros:

$$\ln(S_\tau) \sim \text{normal}\left[\ln(s(1)) + \mu\tau; \sigma^2\tau\right]$$

Seja uma nova variável W_τ igual ao logaritmo natural do preço S_τ. Portanto, a distribuição de S_τ pode ser definida como:

$$S_\tau = e^{W_\tau} \sim \text{lognormal}\left[E\left[e^{W_\tau}\right]; Var\left[e^{W_\tau}\right]\right]$$

Lembrando que:

$$E[\exp(aW_\tau)] = \exp(aE[W_\tau] + a^2 Var[W_\tau]/2),$$

tem-se:

$$S_\tau = \exp(W_\tau) \sim \text{lognormal}\left[s(1)\exp((\mu + \sigma^2/2)\tau); s^2(1)\exp((2\mu + \sigma^2)\tau)\left(\exp(\sigma^2\tau) - 1\right)\right]$$

2.2.3.1.1.10 Criando a medida de probabilidade neutralizadora do preço da incerteza $\mathbb{Q}$

De acordo com Baxter & Rennie (1998), a probabilidade neutralizadora do preço da incerteza q anteriormente definida pode ser expressa como:

$$q = \frac{s(1)\exp(r_f\delta t) - s(2)}{s(3) - s(2)}$$

Substituindo os valores dos preços do ativo por valores genéricos em qualquer nó da árvore binária recombinante, tem-se:

$$q(j) = \frac{s(j)\exp(r_f\delta t) - s(j)\exp(\mu\delta t - \sigma\sqrt{\delta t})}{s(j)\exp(\mu\delta t + \sigma\sqrt{\delta t}) - s(j)\exp(\mu\delta t - \sigma\sqrt{\delta t})}$$

É importante ressaltar que embora a medida $q(j)$ seja expressa como uma probabilidade de subida no nó j, ela é única em toda a árvore. Portanto, $q(j)$ será representado somente por q.

$$q(j) = \frac{s(j)\exp(r_f\delta t) - s(j)\exp(\mu\delta t - \sigma\sqrt{\delta t})}{s(j)\exp(\mu\delta t + \sigma\sqrt{\delta t}) - s(j)\exp(\mu\delta t - \sigma\sqrt{\delta t})}$$

Utilizando a aproximação de Taylor, obtém-se um valor aproximado de q definido como:

$$q \cong \frac{1}{2}\left[1 - \sqrt{\delta t}\left(\mu + \frac{\sigma^2}{2} - r_f\right)\frac{1}{\sigma}\right]$$

Conforme Dos Anjos e Paiva (1998, p. 35-6) e Luenberger (1998, p. 376-7), se forem utilizados os fatores de subida e de descida simplificados (sem o *drift*) [implementação de Cox, Ross & Rubinstein (1979)]:

$$q = \frac{\exp(r_f \delta t) - \exp(-\sigma\sqrt{\delta t})}{\exp(\sigma\sqrt{\delta t}) - \exp(-\sigma\sqrt{\delta t})}$$

$$q \cong \frac{1}{2}\left[1 - \sqrt{\delta t}\left(\frac{1}{2}\sigma^2 - r_f\right)\frac{1}{\sigma}\right]$$

Observe-se que a probabilidade de subida q é diferente de $\frac{1}{2}$ nos dois casos. Quando $\delta t \to 0$, $q \to \frac{1}{2} = p$. Note-se ainda que esta última equação só é válida $(1 > q > 0)$ se $\sigma \geq r_f\sqrt{\delta t}$, isto é, δt não deve ser muito grande. Além disso, a volatilidade local não é constante, isto é, depende de δt.

2.2.3.1.1.11 Fórmula do movimento do preço do ativo-base

X_n continua sendo binomialmente distribuído com média nq e variância $nq(1-q)$.

É importante calcular agora a esperança e a variância de Z_n na medida $\mathbb{Q}$. De acordo com os resultados obtidos:

$$E_{\mathbb{Q}}[Z_n] = -\sqrt{\tau}\left(\mu + \tfrac{1}{2}\sigma^2 - r_f\right)/\sigma$$

Quando δt tende a zero ou n tende ao infinito, observa-se que a variância de Z_n na medida $\mathbb{Q}$ tende a 1.

$$\begin{aligned} Var_{\mathbb{Q}}[Z_n] &= E_{\mathbb{Q}}\left[Z_n^2\right] - E_{\mathbb{Q}}^2[Z_n] \\ &= 1 - \left[\sqrt{\delta t}\left(\mu + \tfrac{1}{2}\sigma^2 - r_f\right)/\sigma\right]^2 \\ &= 1 - \frac{E_{\mathbb{Q}}^2[Z_n]}{n} \end{aligned}$$

Desta forma:

$$Z_{\mathbb{Q}} = \frac{Z_n^{\mathbb{Q}} - E_{\mathbb{Q}}[Z_n]}{\sqrt{Var_{\mathbb{Q}}[Z_n]}} \xrightarrow[n\to\infty]{} \text{normal}(0;1).$$

Também tem-se que:

$$Z_n^{\mathbb{Q}} = Z_{\mathbb{Q}} - E_{\mathbb{Q}}[Z_n].$$

Cumpre lembrar que as distribuições de W_τ (logaritmo natural de S_τ) e de S_τ são, respectivamente:

$$W_\tau \sim \text{normal}\left[\ln(s(1)) + (r_f - \tfrac{1}{2}\sigma^2)\tau; \sigma^2\tau\right]$$

$$S_\tau \sim \text{lognormal}\left[s(1)\exp(r_f\tau); s^2(1)\exp(2r_f\tau)\left[\exp(\sigma^2\tau) - 1\right]\right].$$

Então:

$$\begin{aligned}
S_\tau^{\mathbb{Q}} &= s(1)\exp\left[\mu\tau - \sigma\sqrt{\tau}Z_n^{\mathbb{Q}}\right] \\
&= s(1)\exp\left[\mu\tau - \sigma\sqrt{\tau}\left(Z_{\mathbb{Q}} + E_{\mathbb{Q}}[Z_n]\right)\right] \\
&= s(1)\exp\left[\mu\tau - \sigma\sqrt{\tau}\left(Z_{\mathbb{Q}} + \sqrt{\tau}\left(\mu + \tfrac{1}{2}\sigma^2 - r_f\right)/\sigma\right)\right] \\
&= s(1)\exp\left[\left(r_f - \tfrac{1}{2}\sigma^2\right)\tau - \sigma\sqrt{\tau}Z_{\mathbb{Q}}\right]
\end{aligned}$$

A partir do desenvolvimento das expressões do preço do ativo, é possível determinar o modelo de preço do ativo para qualquer momento do tempo.

2.2.3.1.1.12 Fórmula de Black-Scholes-Merton de determinação do preço racional da opção de compra européia simples ordinária

Uma vez definida a fórmula do preço do ativo na data futura, é possível determinar o preço de uma opção européia na data inicial:

$$\begin{aligned}
f(1) &= b_\tau^{-1}E_{\mathbb{Q}}[F_\tau] \\
&= \exp\left(-r_f\tau\right)E_{\mathbb{Q}}[\max\{0; S_\tau - k_\tau\}] \\
&= \exp\left(-r_f\tau\right)E_{\mathbb{Q}}\left[\max\left\{0; s(1)\exp\left[\left(r_f - \tfrac{1}{2}\sigma^2\right)\tau - \sigma\sqrt{\tau}Z\right] - k_\tau\right\}\right] \\
&= E_{\mathbb{Q}}\{\max[0; s(1)\exp\left(\sigma\sqrt{\tau}Z - \tfrac{1}{2}\sigma^2\tau\right) - k_\tau\exp(-r_f\tau)]\}
\end{aligned}$$

F_τ: preço terminal do derivativo ou condição de contorno $\max\{S_\tau - k\}$.

O resultado desta integração é a tradicional fórmula de Black-Scholes-Merton, obtida em 1973, para opção de compra européia ordinária.

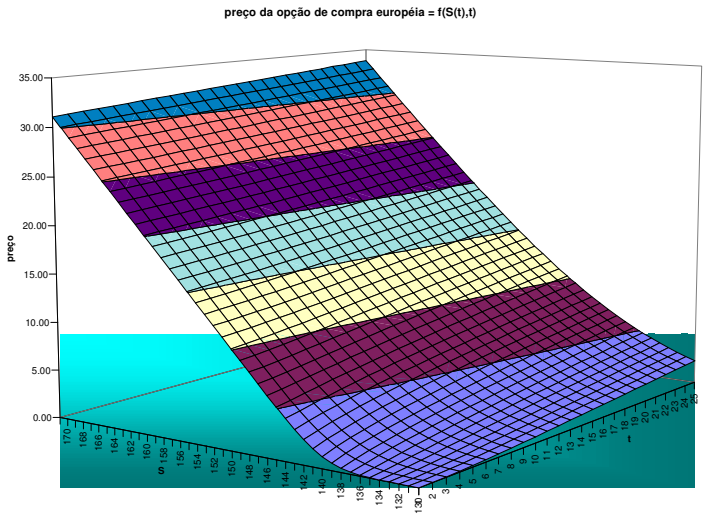
$$\begin{aligned}
f(1) &= s(1)\mathbb{Q}(d_1) - \exp\left(-r_f\tau\right)k_\tau\mathbb{Q}(d_1 - \sigma\sqrt{\tau}) \\
&= s(1)\mathbb{Q}(d_1) - \exp\left(-r_f\tau\right)k_\tau\Pr\left(S_\tau^{\mathbb{Q}} \geq k_\tau\right)
\end{aligned}$$

sendo:

- τ com dimensão igual ao tempo
- r_f com dimensão igual ao inverso do tempo
- $\sqrt{Var_{\mathbb{Q}}\left[\ln\left(\frac{S_\tau}{s(1)}\right)\right]} = \sigma\sqrt{\tau}$ é o desvio-padrão ou a volatilidade da taxa de retorno composta continuamente (adimensional); portanto, conforme Wilmott *et al.*

(1997), σ tem dimensão igual ao inverso da raiz quadrada do tempo $(t^{-\frac{1}{2}})$ e é chamada de volatilidade instantânea ou a termo [Baxter & Rennie (1998)]

- $d_1 = \dfrac{\ln\left(\dfrac{s(1)}{k_\tau}\right) + \left(r_f + \dfrac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}}$ (adimensional)
- $\mathbb{Q}(\bullet)$ é a função de distribuição acumulada de Z_τ na medida $\mathbb{Q}$, isto é, $\mathbb{Q}(x) = \Pr(Z_{\mathbb{Q}} \leq x)$ (adimensional)
- $\Pr\left(S_\tau^{\mathbb{Q}} \geq k_\tau\right)$ é a probabilidade de exercício da opção na medida $\mathbb{Q}$



Conforme Pauli *et al.* (1980, cap. 21) o princípio da homogeneidade dimensional estabelece que toda equação que traduz um fenômeno físico ou econômico verdadeiro é, necessariamente, homogênea do ponto de vista dimensional. Uma equação que traduz uma lei física ou econômica é dita homogênea se as parcelas que constituem os dois membros da igualdade apresentam os mesmos símbolos dimensionais tendo, respectivamente, os mesmos expoentes. O princípio exprime uma condição necessária, porém não suficiente. Com efeito, uma equação pode ser dimensionalmente homogênea e, apesar disso, não traduzir um fenômeno físico ou econômico verdadeiro.

A equação de B-S-M é, portanto, dimensionalmente homogênea.

O resultado final obtido por Black & Scholes (1973) e Merton (1973) demonstra que é possível avaliar uma opção de compra a partir de alguns parâmetros

determinísticos, como preço à vista do ativo, volatilidade (desvio-padrão da taxa de retorno), prazo, taxa de juros certa e preço de exercício.

Gulko (1998, p. 1-38:39) afirma que o único parâmetro livre do modelo é a volatilidade. Afirma ainda que a volatilidade é mensurável, enquanto os outros parâmetros são observáveis. Ele observa ainda que o modelo de Black-Scholes-Merton possui apenas uma fonte de incerteza (o preço do valor mobiliário subjacente) e que, portanto, para permitir o *hedging* perfeito (ou equivalentemente compor um mercado dinamicamente completo por meio de uma estratégia replicante) são necessários dois valores mobiliários linearmente independentes.

É preciso enfatizar que a medida de probabilidade real ou subjetiva do preço do valor mobiliário subjacente é irrelevante para a determinação do preço único e livre de arbitragem da opção. Isto é, a medida de probabilidade neutralizadora do preço da incerteza independe da medida de probabilidade real ou subjetiva. Note-se que S_τ (preço da ação) possui distribuição lognormal na medida de probabilidade neutralizadora do preço da incerteza. Outro resultado importante é a irrelevância da consideração da atitude do decisor diante da incerteza, isto é, este resultado é livre de preferência.

2.2.3.1.1.13 Simplificação do cálculo do preço da opção

Conforme Cox & Rubinstein (1985, p. 157-8) a fórmula de Black-Scholes-Merton para opção de tipo européia ordinária possui a seguinte propriedade:

$$c(\lambda s_t;\tau - t;\lambda k_\tau) = \lambda c(s_t;\tau - t;k_\tau)$$

em que λ é uma constante positiva.

Portanto, pode-se escrever a fórmula de Black-Scholes-Merton da seguinte maneira:

$$c(vpl;\tau - t;1) = vpl\mathbb{Q}(d_1) - \mathbb{Q}(d_2)$$

$$vpl = \frac{s_t}{k_\tau \exp\left[-r_f(\tau - t)\right]}$$

$$\frac{\mathbb{Q}(d_1)}{\mathbb{Q}(d_2)} > 1$$

Note-se que, quanto maior o *vpl* (acima de 1) e a volatilidade acumulada no período $\sigma\sqrt{\tau - t}$ (o que implica no aumento do valor da razão das distribuições acumuladas $\frac{\mathbb{Q}(d_1)}{\mathbb{Q}(d_2)}$), mais atraente é o projeto (opção).

A vantagem desta representação é a possibilidade do uso de tabela cujas entradas são *vpl* e $\sigma\sqrt{\tau - t}$, e o resultado é um valor percentual que deve ser multiplicado por $k_\tau \exp\left[-r_f(\tau - t)\right]$ (valor presente do preço de exercício ou custo de implementação da opção). No entanto, Brealey & Myers (desde a quarta edição de 1991), Luehrman (nos seus dois artigos da HBR de 1998 e um caso de 1995 pela HBS) e Edleson (caso publicado em 1994 pela HBS) ensinam que o valor percentual deve ser multiplicado por s_t (valor do ativo-base na data t). A origem do erro deve-se ao livro de Brealey & Myers que se equivocaram nesta passagem do texto de Cox & Rubinstein (1985). Luehrman faz referência à quarta edição de Brealey & Myers de 1991. Em Brealey & Myers (2000) esse erro de cálculo foi eliminado.

2.2.3.1.2 Abordagem contínua para o cálculo do preço de opções européias

2.2.3.1.2.1 Introdução

Uma maneira de representar matematicamente a dinâmica do retorno do ativo, que é encontrada em Merton (1973), Neftci (1996), Hull (1997), Wilmott *et al.* (1997), pode ser descrita pelo modelo abaixo[19]:

---

[19] Num intervalo de tempo dt podem ser definidas duas maneiras de cálculo da taxa de retorno. Na verdade, não importa se a taxa de retorno (variação percentual do preço) é definida com capitalização discreta ou contínua, pois a solução das equações diferenciais estocásticas abaixo são idênticas.

Taxa de retorno composta continuamente em dt: $\frac{d\ln(S_t)}{dt} = \left(\mu - \frac{\sigma^2}{2}\right) + \frac{dz_t}{dt}$

Taxa de retorno composta discretamente em dt: $\frac{\frac{dS_t}{dt}}{S_t} = \mu + \sigma\frac{dz_t}{dt}$.

$$\frac{dS_t}{S_t} = \mu dt + \sigma dZ_t \text{, sendo}$$

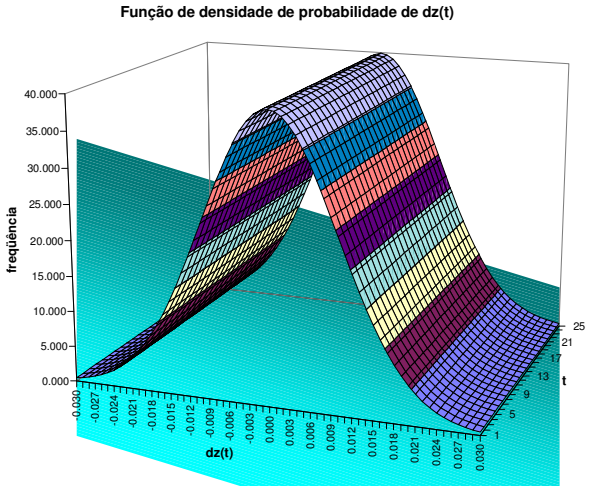
$\frac{dS_t}{S_t}$: taxa de retorno discreto do ativo

dt: intervalo de tempo infinitesimal

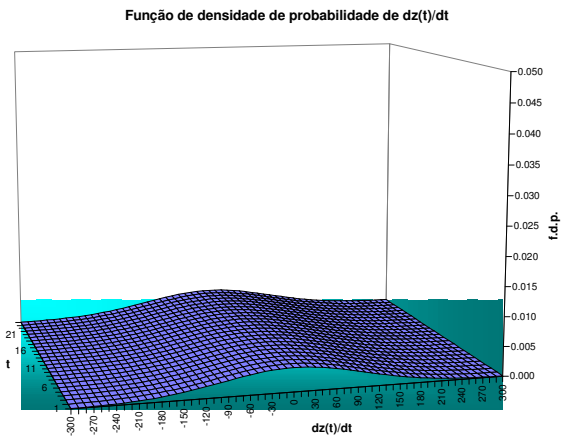
μ: valor esperado instantâneo da taxa de retorno composta discretamente do ativo[20]

σ: volatilidade a termo

dZ_t: variável aleatória com distribuição normal $\mathrm{N}(0;dt)$



[20] A taxa de retorno composta continuamente é $\mu - \frac{\sigma^2}{2}$, pois $\eta = \frac{1}{\tau - t}\ln\left(\frac{S_t}{s_0}\right)$. Como $\ln\left(\frac{S_t}{s_0}\right)$ tem distribuição normal com esperança $\left(\mu - \frac{\sigma^2}{2}\right)(\tau - t)$, então η tem distribuição normal com esperança $\mu - \frac{\sigma^2}{2}$ [Hull (1997, p.230-2)].



A equação acima também é usualmente conhecida como *equação diferencial estocástica*. Ela mostra que o retorno $\frac{dS_t}{S_t}$ do ativo, no intervalo dt, é composto por um termo determinístico μdt e por uma variável aleatória σdZ_t, que segue uma distribuição normal com média 0 e variância $\sigma^2 dt$. Portanto, a taxa de retorno do ativo segue um processo de Wiener generalizado com distribuição normal $\mathrm{N}(\mu dt; \sigma^2 dt)$.

Um aspecto interessante da equação diferencial estocástica diz respeito à dimensão das variáveis envolvidas. A taxa de retorno do ativo $\frac{dS_t}{S_t}$ não possui unidade, ou seja, é adimensional. Isso decorre do fato da taxa de retorno esperada do ativo μ e do desvio-padrão da taxa de retorno σ terem como dimensões t^{-1} e $t^{-½}$, respectivamente [Wilmott *et al.* (1997, p. 81)]. Como o intervalo de tempo tem unidade t e a variável aleatória possui dimensão $t^{½}$, o produto das variáveis gera uma variável adimensional.

2.2.3.1.2.2 Processo de Wiener

Os modelos de retornos de ações são expressos em termos do que se conhece como **processo de Wiener**. Conforme afirma Hull (1997), este tipo de processo é utilizado na física para descrever o movimento de uma partícula sujeita a choques com outras partículas, fato também denominado **movimento browniano**. Na fórmula anterior, o termo dZ_t representa a componente estocástica, responsável pela produção de variações inesperadas. A construção dos modelos de

determinação do preço de derivativos baseados em processos estocásticos exige que sejam conhecidas as propriedades da variável aleatória.

Diz-se que uma variável Z_t segue um processo de Wiener padrão ou movimento browniano se seu diferencial dZ_t possuir duas propriedades:

a) dZ_t pode ser expresso pela equação $dZ_t = \mathrm{E}_t\sqrt{dt}$, sendo E_t uma variável aleatória obtida com distribuição normal padronizada N(0;1). Logo, a variável dZ_t também apresenta uma distribuição normal com os parâmetros de média e variância iguais a 0 e dt, respectivamente;

b) as variáveis dZ_t para dois intervalos de tempo (dt) distintos são independentes.

Ross[21] (1997, p. 524) ainda complementa que a variável Z_t segue movimento browniano padrão se:

a) o processo iniciar do 0: $z_0 = 0$;

b) o conjunto dos possíveis valores de Z_t, para $t \geq 0$, possuem incrementos independentes e estacionários;

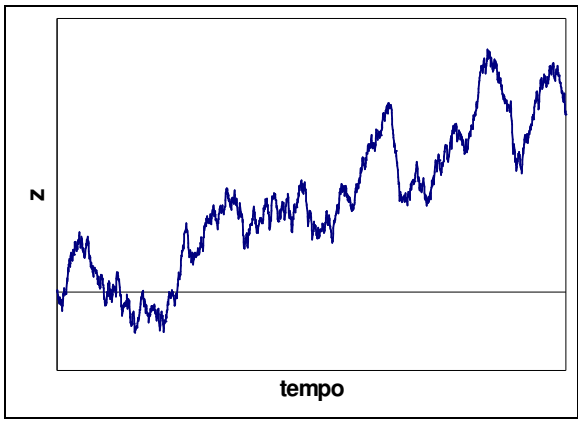
c) para cada instante tempo $t > 0$, Z_t tem distribuição normal com média 0 e variância *t*.

A figura a seguir ilustra uma possível trajetória das variáveis Z_t no tempo, que segue um processo de Wiener padrão:

---

[21] O texto original de Ross (1997) utiliza a expressão $X(t)$ para representar a variável que compõe o movimento browniano. Para uniformizar a notação e dar continuidade às demonstrações, foi feita a substituição por Z_t.

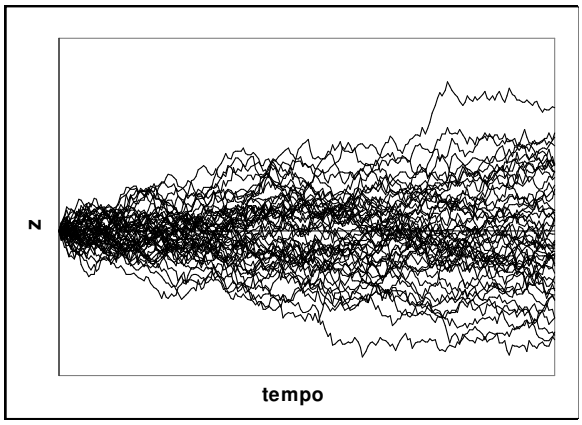
Figura 3: Simulação de uma trajetória do movimento browniano



Observando-se o comportamento do processo estocástico $\{ Z_t ; t \geq 0 \}$ pelo gráfico anterior, nota-se uma série de "bicos" produzidos pelo processo aleatório. Uma função desse tipo não pode ser integrada ou derivada por métodos convencionais do cálculo newtoniano.

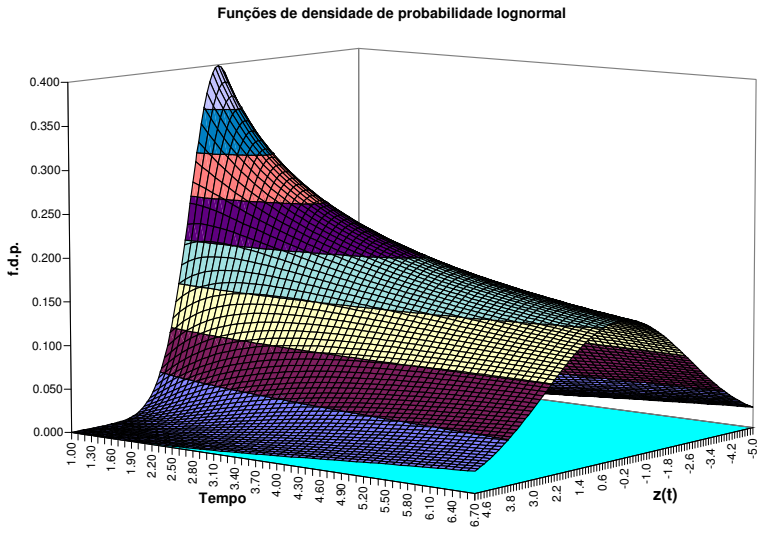
Um processo estocástico seria o conjunto de todas as trajetórias percorridas pelas variáveis Z_τ como ilustra a figura a seguir:

Figura 4: Processo estocástico $\{ Z_t ; t \geq 0 \}$



A seguinte figura representa a função de densidade de probabilidade da variável Z_t em função do tempo. Observa-se que, à medida que o tempo passa, a variância de Z_t aumenta e passa a ter valores mais extremos, diminuindo a concentração em torno de sua média.

Figura 5: Funções densidade de probabilidade do processo estocástico $\{Z_t ; t \geq 0\}$



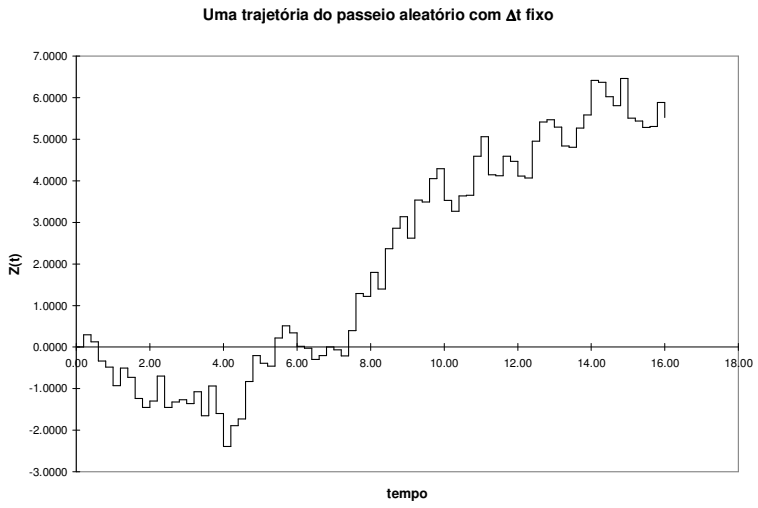
O resultado obtido no gráfico anterior também foi demonstrado por Hull (1997, p. 211). Analisando o comportamento de Z_t durante um período relativamente longo de tempo τ, pode-se descrever a variação como sendo $Z_\tau - z_0$. A variação de Z_τ é a somatória dos incrementos de Z_τ em n intervalos pequenos de tempo δt, sendo:

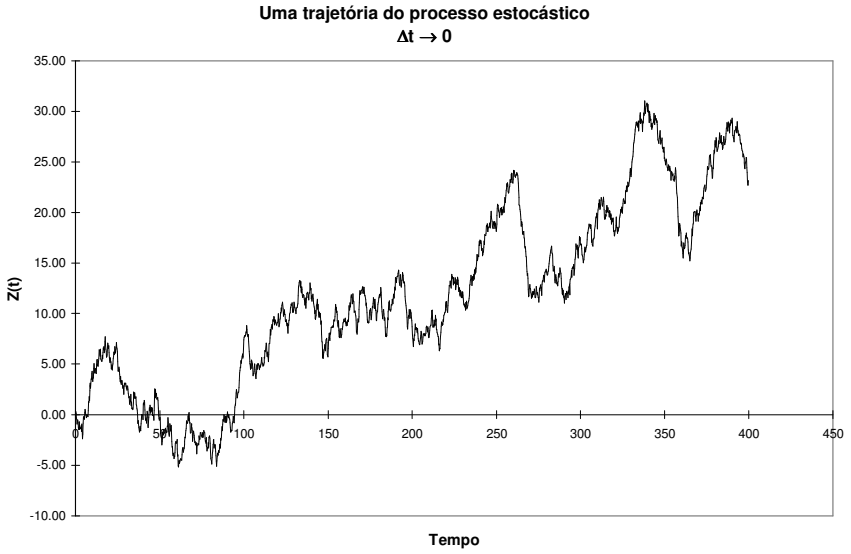
$$n = \frac{\tau}{\delta t}$$

Logo, a variação de Z_τ pode ser escrita como:

$$Z_\tau - z_0 = \sum_{i=1}^{n} \mathrm{E}_i \sqrt{\delta t}$$

E_i: são variáveis aleatórias com N(0;1) e são independentes entre si.

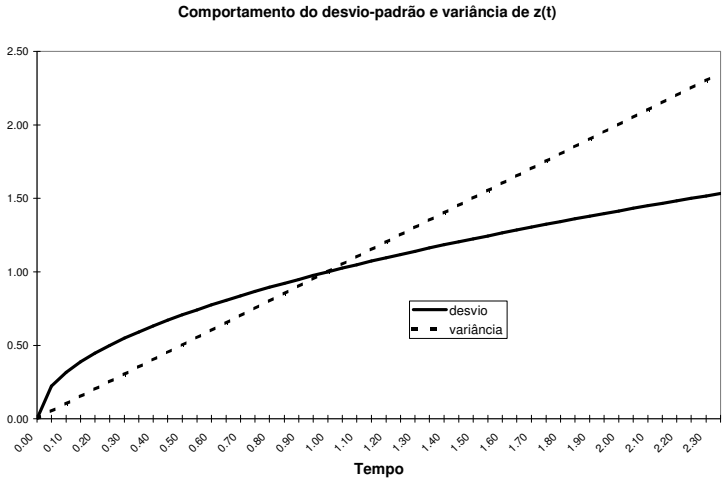




Note-se que, a partir da equação anterior, pode-se inferir que $Z_\tau - z_0$ também tem distribuição normal com parâmetros $\mathrm{N}(0;\tau)$. A soma de diversas variáveis aleatórias normais independentes é uma variável aleatória com distribuição normal.

Observe-se que, para um dado instante de tempo τ, a variável Z_τ também possui uma distribuição normal com média 0 e variância τ:

$$Z_\tau = \sum_{i=1}^{n} E_i \sqrt{\delta t} \text{ , pois } z_0 = 0 \text{ .}$$



Portanto, em qualquer intervalo de tempo τ, o incremento no valor de uma variável que segue o processo de Wiener segue uma distribuição normal com média 0 e desvio-padrão $\tau^{\frac{1}{2}}$.

2.2.3.1.2.3 Fórmula do preço do ativo

No item anterior, a taxa de retorno do ativo foi representada por uma equação diferencial estocástica do tipo:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dZ_t \text{, em que } dZ_t \sim \mathrm{N}(0; dt) \text{ \{eq. 1\}}$$

Se a expressão acima fosse uma equação diferencial ordinária (com $\sigma = 0$), seria fácil encontrar uma solução pela integração. Dessa forma, seria possível isolar a variável S_t e obter uma função determinística $S(\mu; t)$, como a dedução a seguir:

$$\int_0^t \frac{dS_u}{S_u} = \int_0^t \mu du$$

$$\ln|S_u|_0^t = \mu t \Rightarrow S_t = s_0 \exp(\mu t)$$

Analogamente, incluindo a componente estocástica na equação diferencial e integrando, ter-se-ia:

$$\int_0^t \frac{dS_u}{S_u} = \int_0^t \mu du + \int_0^t \sigma dZ_u$$

Contudo, a integração de uma equação diferencial estocástica é bastante complexa e não segue as regras básicas do cálculo newtoniano. Para efetuar esta

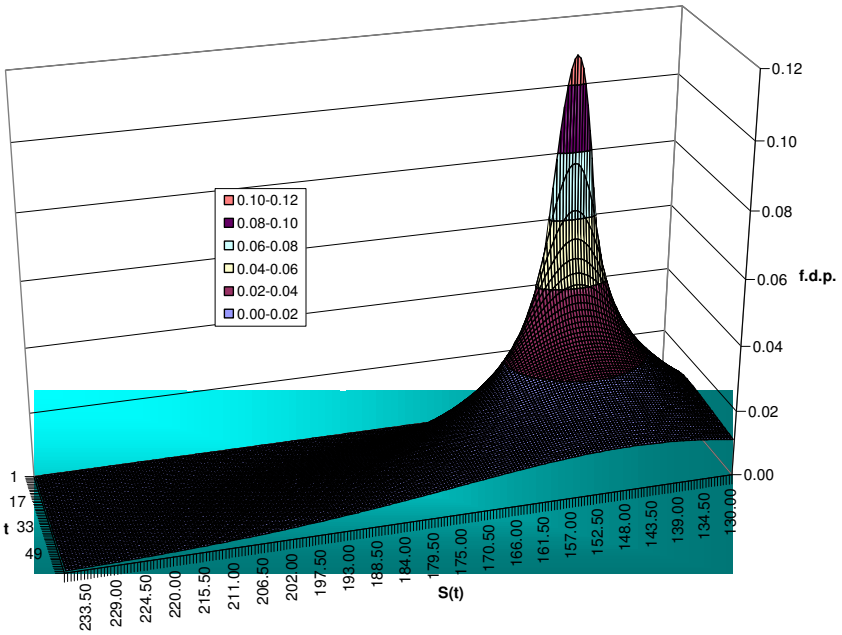
operação, seria necessário aplicar a integral de Ito, que é um instrumento matemático útil para lidar com variáveis aleatórias. Neftci (1996, p. 172) afirma que a "integral de Ito é uma maneira de somar os inúmeros e imprevisíveis incrementos aleatórios durante o tempo".

O resultado final desta integração segue a fórmula abaixo:

$$S_t = s_0 \exp[(\mu - \sigma^2/2)t + \sigma Z_t] \ \{\text{eq. } 2\}$$

A expressão acima é a fórmula para o preço do ativo, cujas taxas de retorno instantâneas possuem uma distribuição normal. Ela é fundamental para modelar o comportamento do ativo no cálculo das opções. A variável aleatória S_t segue uma distribuição lognormal, pois o seu logaritmo é uma variável aleatória com distribuição normal. De acordo com Luenberger (1998, p. 301), dada uma variável $U(t) = \exp(W_t)$, se W_t for uma normal, então diz-se que $U(t)$ segue uma distribuição lognormal.

Figura 6: Lognormal $(\mu = 0{,}0009\,;\sigma = 0{,}025\,;s_0 = 147\,;t\,;Z_t)$



Na expressão 2, a variável S_t segue um movimento browniano geométrico e possui os seguintes parâmetros, supondo que o processo inicie em $t = 0$ e s_0 seja conhecido:

$$E[S_t] = s_0 \exp(\mu t)$$

$$Var[S_t] = s_0^2 \exp(2\mu t)[\exp(\sigma^2 t) - 1]$$

2.2.3.1.2.4 Lema de Ito e sua aplicação sobre o diferencial total estocástico do ativo

Uma maneira de descobrir se a equação 2 é a solução da integração da equação diferencial 1 consiste em aplicar o Lema de Ito. O Lema de Ito estabelece uma expressão matemática que possibilita a obtenção do diferencial total de uma função cujas variáveis seguem um processo de Wiener. Suponha-se o diferencial de uma variável S_t que varia em função do tempo (dt) e de um fator aleatório (dZ_t):

$$dS_t = A(S_t;t)dt + B(S_t;t)dZ_t \quad \{\text{eq. } 3\}$$

De acordo com este lema, se houver uma função G que varia em função de S_t e t, isto é $G(S_t;t)$, seu diferencial dG pode ser expresso da seguinte forma:

$$dG = \left(\frac{\partial G}{\partial S_t} A(S_t,t) + \frac{\partial G}{\partial t} + \frac{1}{2}\frac{\partial^2 G}{\partial S_t^2} B(S_t,t)\right)dt + \frac{\partial G}{\partial S_t} B(S_t,t)dZ_t \quad \{\text{eq. } 4\}$$

A expressão acima nada mais é do que a expansão de Taylor para uma função que possui variáveis independentes determinísticas e aleatórias. A demonstração deste resultado pode ser obtida em Hull (1997, p. 225).

O Lema de Ito será aplicado sobre a expressão 2 para provar que se trata de um resultado possível do diferencial do preço do ativo. A fórmula do preço do ativo obtida anteriormente pode ser considerada uma função das variáveis Z_t e t, logo:

$$\begin{aligned} S_t &= G(z_t;t) \\ &= s_0 \exp[(\mu - \sigma^2/2)t + \sigma Z_t] \end{aligned}$$

O diferencial da função G será expresso em relação a Z_t e t. Efetuando a devida substituição no Lema de Ito descrito na equação 4, tem-se:

$$dG = \left(\frac{\partial G}{\partial Z_t} A(Z_t,t) + \frac{\partial G}{\partial t} + \frac{1}{2}\frac{\partial^2 G}{\partial Z_t^2} B(Z_t,t)\right)dt + \frac{\partial G}{\partial Z_t} B(Z_t,t)dZ_t \quad \{\text{eq. } 5\}$$

Para simplificar as expressões das derivadas parciais de G em relação a Z_t e t a seguir, substituiu-se o termo $\mu - \frac{\sigma^2}{2}$ por α:

$$\begin{aligned}\frac{\partial G}{\partial t} &= \alpha s_0 \exp(\alpha t + \sigma Z_t) \\ &= \alpha S_t\end{aligned}$$

e

$$\begin{aligned}\frac{\partial G}{\partial Z_t} &= \sigma s_0 \exp(\alpha t + \sigma Z_t) \\ &= \sigma S_t\end{aligned}$$

e

$$\frac{\partial^2 G}{\partial Z_t^2} = \sigma^2 S_t \,.$$

Substituindo as derivadas parciais na expressão 5 e admitindo que $A(Z_t;t) = 0$ e $B(Z_t;t) = 1$, pois $dS_t = dZ_t$ conforme equação 3, tem-se:

$$dG(Z_t,t) = \left(\sigma S_t 0 + \alpha S_t + \frac{1}{2}\sigma^2 S_t 1\right)dt + \sigma S_t 1 dZ_t \Rightarrow$$

$$\begin{aligned}dS_t &= S_t\left[\left(\alpha + \frac{1}{2}\sigma^2\right)dt + \sigma dZ_t\right] \Rightarrow \\ \frac{dS_t}{S_t} &= \left(\alpha + \frac{1}{2}\sigma^2\right)dt + \sigma dZ_t\end{aligned}$$

Como $\alpha = \mu - \frac{\sigma^2}{2}$, segue-se que:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dZ_t \sim N(\mu dt; \sigma^2 dt)$$

Outra aplicação do Lema de Ito é a resolução do diferencial do logaritmo dos preços ($d \ln S_t$). A vantagem de se utilizar o logaritmo dos preços é que os termos multiplicativos passam a funcionar como uma simples soma. Seja uma função $F(Z_t;t) = \ln S_t$. Pode-se escrever a partir da expressão 2 que:

$$\begin{aligned}\ln S_t &= F(Z_t,t) \\ &= \ln s_0 + \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma Z_t\end{aligned}$$

As derivadas parciais de $F(Z_t;t)$ em t e Z_t são:

$$\frac{\partial F}{\partial t} = \mu - \frac{\sigma^2}{2}$$

e

$$\frac{\partial F}{\partial Z_t} = \sigma$$

e

$$\frac{\partial^2 F}{\partial Z_t^2} = 0$$

Considerando que $A(Z_t;t) = 0$ e $B(Z_t;t) = 1$ e substituindo no Lema de Ito:

$$d\ln S_t = (\mu - \frac{\sigma^2}{2})dt + \sigma dZ_t \sim N\left[\left(\mu - \frac{\sigma^2}{2}\right)dt;\sigma^2 dt\right] \text{ \{eq. 6\}}$$

Observa-se que a expressão anterior é semelhante à obtida pelo cálculo $\frac{dS_t}{S_t}$, exceto pelo fator de ajuste $\frac{\sigma^2}{2}$. Segundo Luenberger (1998, p. 310), este fator de ajuste decorre da transformação da variável S_t para $\ln S_t$ e é igual ao ajuste necessário quando se calcula a esperança de uma variável aleatória com distribuição lognormal. Existe uma relação entre os diferenciais de preços e do logaritmo dos preços dada pela expressão:

$$\frac{dS_t}{S_t} = d\ln S_t + \frac{\sigma^2}{2}$$

Observe-se que S_t dado pela equação 2 também é solução da equação diferencial estocástica 6. Se S_t fosse uma variável determinística, pela regra da cadeia, ter-se-ia:

$$\frac{d}{dt}\ln(S_t) = \frac{1}{S_t}\frac{d}{dt}S_t \Rightarrow d\ln(S_t) = \frac{dS_t}{S_t}$$

Em um instante de tempo t, tal que $t \leq \tau$, segundo Ross (1997, p.538), Hunter e Stowe (1992, p. 32), Hull (1997, p. 241) e Luenberger (1998, p. 357), o preço de uma opção de compra européia simples ordinária pode ser descrito como o valor esperado do valor intrínseco, na medida de probabilidade subjetiva ou real de S_τ:

$$c_{t;\tau}^{\mathbb{P}} = \frac{E_{\mathbb{P}}\{\max[0;S_\tau - k]\}}{\exp[\mu(\tau - t)]}$$

Substituindo o termo S_τ pelo modelo do preço do valor mobiliário subjacente, dado que o processo se inicie a partir da data $t = 0$ e considerando uma avaliação num mundo com preço da incerteza neutralizado, pode-se substituir μ pela taxa certa r_f, pois o excesso de retorno esperado do valor mobiliário subjacente $\mu - r_f$ deve ser subtraído de μ devido à eliminação do risco (*hedging* perfeito devido à estratégia replicante), isto é, $\mu - (\mu - r_f) = r_f$:

$$c_{0;\tau}^{\mathbb{P}} = \frac{E_{\mathbb{P}}\left\{\max\left[0; s_0 \exp[(\mu - \sigma^2/2)\tau + \sigma Z_\tau] - k_\tau\right]\right\}}{\exp(\mu\tau)}$$

$$c_{0;\tau}^{\mathbb{Q}} = \frac{E_{\mathbb{Q}}\left\{\max\left[0; s_0 \exp[(r_f - \sigma^2/2)\tau + \sigma Z_\tau] - k_\tau\right]\right\}}{\exp(r_f\tau)}$$

O teorema de Cameron-Martin-Girsanov (Baxter & Rennie (1998, p. 98) e Neftci (1996, cap. 14)) fundamentalmente estabelece que apenas o *drift* da equação diferencial estocástica deve ser corrigido pela eliminação do excesso de retorno esperado devido ao preço da incerteza para a obtenção da medida de probabilidade neutralizadora do preço da incerteza equivalente[22] à medida de probabilidade real ou subjetiva para "compensar" esta transformação. Em Neftci (1996, cap. 6, 14 e 15), encontram-se também excelentes explicações sobre estas transformações e suas aplicações na formação de preço de derivativos.

Este é o modelo para formação do preço de opções de compra européias em um contexto sem incerteza, para ativos cujos preços seguem um movimento browniano geométrico. Ross (1997, p. 538) provou (sem o devido cuidado de mostrar que a taxa de desconto deve ser a taxa certa) que, a partir da equação anterior, o preço racional da opção pode ser expresso da seguinte forma:

$$c_{0;\tau}^{\mathbb{Q}} = \exp(-r_f\tau)\int_{-\infty}^{\infty} \max\{0; s_0 \exp(y) - k_\tau\} \frac{1}{\sqrt{2\pi\tau\sigma^2}} \exp\{-[y - (r_f - \sigma^2/2)\tau]^2 / 2\tau\sigma^2\} dy$$

em que $c_{0;\tau}^{\mathbb{Q}}$: preço racional da opção de compra européia simples ordinária no instante $t = 0$ com exercício em τ e $y = \left(r_f - \frac{\sigma^2}{2}\right)\tau + \sigma Z_\tau$.

Resolvendo a equação, chega-se à mesma equação obtida por Black & Scholes (1973):

$$c_{0;\tau}^{\mathbb{Q}} = s_0 \mathbb{Q}(d_1) - \frac{k_\tau}{\exp(r_f\tau)} \mathbb{Q}(d_1 - \sigma\sqrt{\tau})$$

sendo

$\mathbb{Q}(\cdot)$: função de probabilidade acumulada de uma normal padronizada.

---

[22] Conforme Neftci (1996, p. 283), a expressão **medidas de probabilidade equivalentes** significa apenas que elas são atribuídas ao mesmo domínio (conjunto de apoio) da variável aleatória.

$$d_1 = \frac{\ln(s_0 / k_\tau) + (r_f + \sigma^2 / 2)\tau}{\sigma\sqrt{\tau}}$$

Observe-se que, substituindo o instante inicial $t = 0$ por $t \leq \tau$ e τ por $\tau - t$, obtém-se:

$$c_{t;\tau}^{\mathbb{Q}} = s_t \mathbb{Q}(d_1) - \frac{k}{\exp[r_f(\tau - t)]} \mathbb{Q}(d_1 - \sigma\sqrt{\tau - t})$$

$$d_1 = \frac{\ln(s_t / k_\tau) + (r_f + \sigma^2 / 2)(\tau - t)}{\sigma\sqrt{\tau - t}}$$

2.3 Elementos da Teoria da Informação

"Quando Deus criou o mundo e as pessoas para viver nele - obra que segundo a ciência moderna, levou muito tempo - eu poderia muito bem imaginar que raciocinou da seguinte forma: "Se faço tudo previsível, estes seres humanos os quais eu dotei de cérebros bastante bons, sem dúvida aprenderão a prever tudo, e em conseqüência não terão estímulo para fazer nada, porque reconhecerão que o futuro está totalmente determinado e nele nenhuma ação humana poderia influenciar. Por outro lado, se faço tudo imprevisível, gradualmente descobrirão que não há base racional para nenhuma decisão e em conseqüência, como no primeiro caso, não terão motivos para fazer nada. Nenhum destes projetos tem sentido. Criarei, em conseqüência, uma mistura dos dois. Algumas coisas previsíveis e outras imprevisíveis. Terão então, entre muitas outras coisas, a importante tarefa de saber qual é qual."

Schumacher, E. F. 1977. *O negócio é ser pequeno: um estudo de economia que leva em conta as pessoas.* R.J.: Zahar.

A Teoria da Informação baseia-se em dois conceitos matemáticos: probabilidade e logaritmo.

Serão formalizados matematicamente quantificadores do conceito de surpresa ou incerteza. Estes quantificadores formam a base da Teoria da Informação. Esta teoria permite uma abordagem alternativa que unifica e expande a capacidade de formulação e resolução de problemas de natureza estocástica. Os dois quantificadores fundamentais da Teoria da Informação são o informacional e a entropia. As unidades de medida destes quantificadores também são aqui definidas. A entropia concorre com a variância para medir variabilidade.

Os conceitos e modelos de independência Estatística, cadeia de Markov, função utilidade, análise de componentes principais e distinção entre processos caóticos e aleatórios desempenham um importante papel na modelagem de problemas das

Ciências Sociais. Neste capítulo serão formulados matematicamente estes conceitos e modelos utilizando a Teoria da Informação. Esta abordagem possibilita a unificação e a expansão da capacidade de lidar com problemas de natureza estocástica.

2.3.1 Um pouco de história

Conforme Lima (1985, p. 2), o logaritmo foi inventado simultaneamente por Jost Bürgi (1552-1632), suíço, relojoeiro, matemático e inventor, e por John Napier (1550-1617), um nobre escocês, teólogo e matemático. Os dois inventores desconheciam inteiramente o trabalho um do outro. Ainda conforme Lima (1985, p. 3), esta invenção tornou-se decisiva para a ciência e a tecnologia, e está estreitamente ligada à diversas leis matemáticas e vários fenômenos naturais e mesmo sociais.

Para Bernstein (1997, p. 53), o *Livro dos Jogos de Azar*, iniciado em 1525 e reescrito em 1565, de Girolamo Cordano (1500-1571), médico e jogador inveterado, foi o primeiro esforço de pôr a medição a serviço do risco. No entanto, para Stigler (1986, p. 63), o pai da quantificação da incerteza é Jacob Bernoulli (1654-1705), matemático, filósofo e professor da Universidade de Basel a partir de 1687, contemporâneo e rival ocasional de Isaac Newton. Conforme Bernstein (1997, cap. 7), por meio de sua obra póstuma *A Arte da Conjectura*, publicada em 1713, ocorre o início da formalização do conceito de probabilidade, evento, probabilidade prévia (*a priori*) e posterior (*a posteriori*) e a lei dos grandes números. Seu livro procura demonstrar onde termina a arte de pensar objetivamente e começa a arte da conjectura, considerando que a conjectura é o processo de estimar o todo a partir das partes.♦

2.3.2 O conceito físico de entropia discreta de Boltzmann

A entropia, na Mecânica Estatística, é definida como o grau de desordem de um sistema (Yang (1997, p. 27)). Quanto maior a desordem de um sistema, maior a entropia [Sears (1955, p. 286)].

O desenvolvimento a seguir foi baseado em Golan *et al.* (1996, p. 6-11), Sears (1955, p. 277-90) e Wolf (1996).

O sistema mais simples que se pode considerar, de um ponto de vista estatístico, é um gás monoatômico contido num sistema fechado. Uma especificação completa do estado do gás, a partir de um ponto de vista molecular, reclama a definição da posição e da velocidade de cada uma das moléculas. Isto é, seis quantidades devem ser especificadas para cada molécula: x, y, z, v_x, v_y e v_z.

Este hiperespaço de dimensão seis é chamado de espaço de fase.

Este espaço de fase é dividido em pequenos elementos de volume, chamados células, cujos lados possuem comprimento dx, dy, dz, dv_x, dv_y e dv_z. Estes diferenciais são pequenos quando comparados com as dimensões do espaço de fase (sistema) e a faixa de velocidades das moléculas, porém grande o suficiente para que cada célula contenha um grande número de moléculas (pontos). O volume desta célula ou o produto das seis quantidades acima será representado por v. Cada molécula de gás tem sua representação na forma de um ponto com seis coordenadas e, por comodidade, será chamado de ponto de fase. Imaginem-se as células 1, 2, ..., *i*, ... e com $n_1, n_2, ..., n_i, ...$ pontos de fase, respectivamente.

Uma especificação completa das seis coordenadas de cada molécula de um sistema, com os limites das dimensões da célula onde seus pontos jazem, é dito um microestado do sistema. Tal especificação determina onde cada molécula está, dentro dos limites de dx, dy e dz e quão rápida e em que direção ela está se movendo, dentro dos limites dos diferenciais de velocidade. Esta descrição detalhada é desnecessária para determinar as propriedades observáveis do gás. Por exemplo, a densidade (no conceito simples de massa por unidade de volume) é a mesma se há o mesmo número de moléculas por elemento de volume do espaço, desconsiderando quais moléculas jazem nos elementos de volume. De forma semelhante, a pressão exercida depende apenas da quantidade de moléculas com determinadas velocidades, sendo irrelevante o conhecimento de quais moléculas possuem aquelas velocidades. Em outras palavras, as propriedades observáveis dependem apenas de quantos pontos de fase estão em cada célula do

espaço de fase. A especificação do número de pontos de fase em cada célula do espaço de fase (n_i) define um macroestado do sistema.

Uma das hipóteses fundamentais da mecânica Estatística é que todos os microestados são igualmente prováveis. Isto é, durante um longo período de tempo, qualquer microestado ocorre tão freqüentemente quanto qualquer outro.

O número de microestados correspondendo a qualquer dado macroestado é chamado de **probabilidade termodinâmica relativa do macroestado** [Golan *et al.* (1996, p. 23)] e é representado por *w*. Em geral, *w* é um número muito grande.

$p_{w_i} = \dfrac{w_i}{\sum_{i=0}^{m} w_i}$ é a probabilidade termodinâmica absoluta do macroestado (Golan *et al.* (1996, p. 23)).

Suponha-se que há apenas duas células no espaço de fase, *i* e *j*, e quatro pontos de fase, *a, b, c* e *d*. Sejam n_i e n_j as quantidades de pontos de fase das respectivas células. Os possíveis macroestados são

n_i	4	3	2	1	0
n_j	0	1	2	3	4

Isto é, há cinco macroestados. Para cada um dos macroestados há, em geral, um número diferente de microestados. São quatro os microestados correspondentes ao particular macroestado $n_i = 3$ e $n_j = 1$.

Célula *i*	b c d	c d a	d a b	a b c
Célula *j*	a	b	c	d

Portanto, o número de microestados de um dado macroestado é:

$$\begin{aligned} w &= \frac{n!}{n_1!n_2!} \\ &= \frac{n!}{n_1!(n-n_1)!} \\ &= \binom{n}{n_1} \\ &= \frac{4!}{3!1!} \\ &= 4 \end{aligned}$$

No caso geral de haver n pontos de fase e com mais de duas células, o número de microestados relativos a um macroestado ou a probabilidade termodinâmica relativa do macroestado é:

$$\begin{aligned} w &= \frac{n!}{n_1!\,n_2!\,n_3!\ldots} \\ &= \frac{n!}{\prod n_i!} \end{aligned}$$

Retornando ao exemplo canônico:

$$\begin{aligned} w\left(n_i = 4; n_j = 0\right) &= \frac{4!}{4!0!} = 1 \\ w\left(n_i = 3; n_j = 1\right) &= \frac{3!}{3!1!} = 4 \\ w^*\left(n_i = 2; n_j = 2\right) &= \frac{2!}{2!2!} = 6 \\ w\left(n_i = 1; n_j = 3\right) &= \frac{1!}{1!3!} = 4 \\ w\left(n_i = 0; n_j = 4\right) &= \frac{0!}{0!4!} = 1 \end{aligned}$$

Logo, as probabilidades termodinâmicas absolutas dos macroestados são:

$$\begin{aligned} p_{w_4} &= \frac{w\left(n_i = 4; n_j = 0\right)}{\sum_{k=0}^{4} w\left(n_i = 4-k; n_j = k\right)} = \frac{1}{16} \\ p_{w_3} &= \frac{4}{16} \\ p_{w_2} &= \frac{6}{16} = p^* \\ p_{w_1} &= \frac{4}{16} \\ p_{w_0} &= \frac{1}{16} \end{aligned}$$

Há ao todo 16 microestados possíveis gerados pelos cinco macroestados. Sob a hipótese de que todos os microestados são igualmente prováveis, o primeiro e o quinto macroestados ocorrerão, no longo prazo, 1/16 das vezes, o segundo e quarto macroestados ocorrerão ¼ das vezes e o terceiro, 3/8 das vezes. Note-se que o terceiro macroestado é o mais provável. Além disso, note-se que este macroestado é o mais equilibrado de todos em termos de quantidade de pontos de fase por célula, isto é, os pontos de fase distribuem-se uniformemente entre as células do sistema fechado. Este macroestado especial pode ser chamado, desta

forma, de estado de equilíbrio ou de máxima probabilidade termodinâmica. Se o sistema não está em equilíbrio, mudanças ocorrerão nos sistema, até que o estado de equilíbrio tenha sido atingido. Portanto, a Mecânica Estatística interpreta o aumento da entropia num sistema fechado como conseqüência natural de um sistema tender de um estado menos provável para um mais provável. Quanto maior a desordem de um sistema, maior a sua probabilidade termodinâmica e maior a entropia. O mais alto grau de ordem de um sistema é ter todos os pontos de fase concentrados numa única célula. Neste estado, a probabilidade termodinâmica relativa (w) é 1 e a entropia é nula. O estado de máxima entropia ou de equilíbrio não é um estado estático, pois os pontos de fase estão em permanente movimento no sistema. Logo, ocasionalmente o sistema diferirá do estado de equilíbrio e, consequentemente, a entropia deixará de ser máxima. Pequenas mudanças são mais prováveis do que as grandes, mas estas não são impossíveis, apenas altamente improváveis. Este fenômeno é estudado na Teoria das Flutuações [Sears (1955, cap. 15)].♦

É preciso retornar ao problema de determinar o w do sistema de gás monoatômico, em que o número n e todos os n_k, $k = 1,2,...,m$ são grandes. O fatorial de um número grande pode ser obtido com boa precisão pela aproximação assintótica de Stirling:

$$\begin{aligned}
\log(n!) &= \sum_{k=1}^{n} \ln k \\
&= \sum_{k=2}^{n} \ln k \\
&\approx \int_{1}^{n} \ln z dz \\
&= n \ln n - n + 1 \\
&\approx n \ln n - n
\end{aligned}$$

Aplicando o logaritmo natural em ambos os lados da equação e a aproximação de Stirling:

$$\begin{aligned}
w &= \frac{n!}{n_1! n_2! ... n_m!} \\
&= \frac{n!}{\prod_{k=1}^{m} n_k!}
\end{aligned}$$

obtém-se:

$$\begin{aligned}
\ln w &= \ln n! - \sum_{k=1}^{m} \ln n_k! \\
&\approx n \ln n - n - \sum_{k=1}^{m} n_k \ln n_k + \sum_{k=1}^{m} n_k \\
&= n \ln n - \sum_{k=1}^{m} n_k \ln n_k
\end{aligned}$$

Fazendo $p_k = \dfrac{n_k}{n}$ quando $n \to \infty$ tem-se:

$$\begin{aligned}
h &= \frac{\ln w}{n} \\
&= -\sum_{k=1}^{m} p_k \ln p_k
\end{aligned}$$

em que $0 \ln 0 = 0$.

Note-se também que:

$$w = \exp(nh).$$

Em termos matriciais:

$$\begin{aligned}
h(p) &= -\sum_{k=1}^{m} p_k \ln p_k \\
&= -p^T \ln p
\end{aligned}$$

em que $p^T 1 = 1$.

Esta última equação é a famosa medida de entropia de Shannon.

A máxima entropia (h^*) ocorre quando o sistema está no estado de equilíbrio (w^*) e vale m (número de células).

$$\begin{aligned}
h^* &= \frac{\ln w^*}{n} \\
&= -\sum_{k=1}^{m} \frac{1}{m} \ln \frac{1}{m} \\
&= m
\end{aligned}$$

Outra forma interessante de derivar a fórmula da entropia discreta é apresentada por Wolf (1996, note12.html).

Conforme Wolf (1996), a entropia de algo é a média assintótica do logaritmo do número de modo que este algo acontece. Então, se há dois geradores de eventos independentes

$$m = \{m_1 = \text{cara}; m_2 = \text{coroa}\}$$

e

$$d = \{d_1 = \text{face } 1; d_2 = \text{face } 2; \ldots; d_6 = \text{face } 6\}\ )$$

As medidas de probabilidade são uniformes, isto é,

$$\Pr(m_i) = \frac{1}{2},\ i = 1,2$$
$$\Pr(d_i) = \frac{1}{6},\ i = 1,2,\ldots,6$$

Há $n_m = 2$ e $n_d = 6$ eventos em cada espaço amostral. Portanto, pelo princípio fundamental da contagem, há $n_m n_d = 2 * 6 = 12$ combinações de eventos equiprováveis dos dois geradores.

Isto é, há a criação de um novo espaço amostral equiprovável s resultante do produto cartesiano dos dois espaços amostrais:

$$\begin{aligned} s &= m \times d \\ &= \boxed{\begin{matrix} \{m_1;d_1\} & \{m_1;d_2\} & \{m_1;d_3\} & \{m_1;d_4\} & \{m_1;d_5\} & \{m_1;d_6\} \\ \{m_2;d_1\} & \{m_2;d_2\} & \{m_2;d_3\} & \{m_2;d_4\} & \{m_2;d_5\} & \{m_2;d_6\} \end{matrix}} \\ \Pr\left(\left\{m_i;d_j\right\}\right) &= \frac{1}{2} * \frac{1}{6} = \frac{1}{12} \end{aligned}$$

Considere-se que são produzidos n eventos de s. Portanto, há $(n_m * n_d)^n$ modos dos eventos de s ocorrerem. Imagine-se que $n = 2$. Logo, $(n_m * n_d)^n = (n_m)^n (n_d)^n = 12^2 = 144$. Um novo espaço amostral é criado $s^2 = s \times s$.

A entropia de $m^n = \prod_{i=1}^{n} m$ é dada por:

$$\begin{aligned} h(m) &= \frac{1}{n} \log_2 n_m^n \\ &= \log_2 n_m \end{aligned}$$

A entropia de $d^n = \prod_{i=1}^{n} d$ é dada por:

$$\begin{aligned} h(d) &= \frac{1}{n} \log_2 n_d^n \\ &= \log_2 n_d \end{aligned}$$

A entropia discreta deste processo conjunto ou do espaço amostral $s^n = \prod_{i=1}^{n} s$ é:

$$
\begin{aligned}
h(m,d) &= \frac{1}{n}\log_2\left(n_m n_d\right)^n \\
&= \log_2 n_m n_d \\
&= \log_2 n_m + \log_2 n_d \\
&= h(m) + h(d)
\end{aligned}
$$

Portanto, a entropia conjunta do espaço amostral com base em dois geradores independentes é a soma das entropias dos espaços amostrais de cada gerador (propriedade aditiva da entropia).

No entanto, quando os geradores não são independentes, dá-se que:

$$h(m,d) \leq h(m) + h(d)\,.$$

2.3.3 O conceito lógico de entropia discreta de Shannon

"*Os dois aspectos da matéria cômica e as duas impressões que dela recebemos devem ser irreconciliáveis, e quanto mais díspares forem entre si, maior será a comicidade da situação. A surpresa é o resultado da aceitação simultânea dos dois aspectos irreconciliáveis da situação.*"

Élie Auborim. 1948. *Técnica e psicologia do cômico.*
(citado por Luján, N. 1979. *O humorismo.* Rio de Janeiro: Salvat, p. 19)

2.3.3.1 Quantificação da surpresa no exemplo da eleição para governador do Distrito Federal

Nas últimas eleições (do segundo turno) estaduais e do Distrito Federal, uma das surpresas observadas foi a vitória do atual Governador do Distrito Federal, Sr. Joaquim Roriz, posto que as pesquisas eleitorais apontavam como vencedor o seu oponente, o Prof. Buarque. Tais pesquisas eleitorais apresentavam aproximadamente os seguintes resultados em termos de porcentagem de intenção de voto (aproximação da probabilidade de vitória) destes candidatos: Roriz com 45% e Buarque, 55%. A surpresa, ao ter conhecimento do resultado, de acordo com a Teoria da Informação de Shannon (TI), pode ser quantificada antes da ocorrência de um evento, isto é, pode ser medida tomando-se o logaritmo na base dois do inverso da probabilidade de vitória de cada candidato. A unidade deste resultado é bit. Se a base do logaritmo for e, isto é, $\ln(x)$ (logaritmo natural), a unidade será nit. No presente exemplo seriam:

$$\text{Surpresa(Roriz vencer)}=\log_2(1/0{,}45)=-\log_2(0{,}45)\approx 1{,}15 \text{ bit}$$

e

$$\text{Surpresa(Buarque vencer)}=\log_2(1/0{,}55)=-\log_2(0{,}55)\approx 0{,}86 \text{ bit.}$$

Ao ter conhecimento dos resultados de opinião eleitorais, a surpresa associada ao candidato Roriz é maior do que a surpresa associada ao seu oponente. Portanto, há aproximadamente 34% (1,15/0,86) a mais de surpresa de Roriz vencer a eleição em relação à vitória de Buarque.

Estatisticamente, neste caso, temos uma variável aleatória chamada Vitória Eleitoral (V), que está associada a dois eventos exaustivos e mutuamente excludentes: ω_1=vitória de Roriz e ω_2=vitória de Buarque com probabilidades $\Pr(\omega_1) = p_1 = 0,45$ e $\Pr(\omega_2) = p_2 = 0,55$, respectivamente. Note-se que a soma das probabilidades é um, isto é, $p_1 + p_2 = 1$, e que elas são estritamente positivas, isto é, $p_1 > 0$ e $p_2 > 0$. Este espaço amostral pode ser representado de forma tabular:

Evento ω_i	*Probabilidade* $\Pr(\omega_i) = p_i$
ω_1=vitória de Roriz	$\Pr(\omega_1) = 0,45$
ω_2=vitória de Buarque	$\Pr(\omega_2) = 0,55$

2.3.3.2 Auto-informação ou informacional de um evento

Desta forma, a surpresa provocada pela vitória de Roriz pode ser formalizada da seguinte maneira:

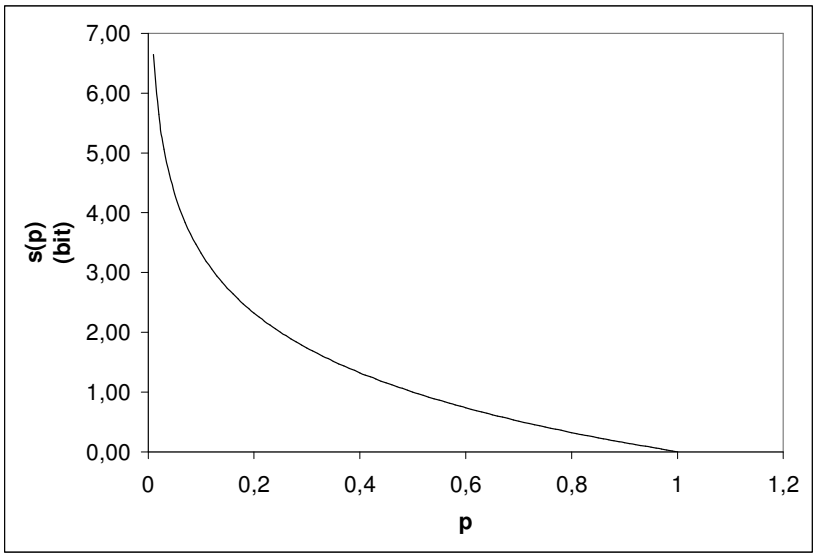
$$\begin{aligned} s(p_i) &= \log_2 \frac{1}{p_i} \\ &= \log_2 p_i^{-1} \\ &= -\log_2 p_i \text{ bit} \end{aligned}$$

A medida da surpresa $s(p_i)$ é também chamada de **auto-informação do evento** ω_i ou **informacional** de ω_i.

Evento ω_i	*Probabilidade* $\Pr(\omega_i) = p_i$	*Valor (bit)* $x_i = -\log_2 p_i$
ω_1=vitória de Roriz	$\Pr(\omega_1) = 0,45$	$x_1 = -\log_2 p_1 = 1{,}15$
ω_2=vitória de Buarque	$\Pr(\omega_2) = 0,55$	$x_2 = -\log_2 p_2 = 0,86$

Seria possível tomar os logaritmos em qualquer base, mas é conveniente utilizar a base dois (da mesma forma que se utiliza o metro para medir comprimento), pois é freqüente medir a quantidade de surpresa, incerteza ou informação de um evento em bit.

O informacional é usado como medida da surpresa de um evento por várias razões. A primeira delas é que a probabilidade é um número que varia no intervalo (0;1] e pode ser considerado como uma medida que quantifica o grau de expectativa de ocorrência de um evento. Logo, o logaritmo do inverso da probabilidade é decrescente quando a probabilidade cresce. Este resultado está de acordo com o conceito de surpresa estabelecido. Num extremo, o informacional de um evento certo $(p = 1)$ é nulo, ou seja, não há nenhuma surpresa com o resultado de um evento certo. Isto é, $p \to 1 \Leftrightarrow s(p) \to 0$. Noutro extremo, quando a probabilidade de um evento tende para zero, o informacional tende para o infinito, isto é, $p \to 0 \Leftrightarrow s(p) \to +\infty$. Logo, é incomensurável a surpresa provocada pelo evento com probabilidade ínfima de ocorrência.



Logo, o informacional varia entre zero e o infinito positivo.

Se for considerado o informacional associado a dois eventos simultâneos e independentes ω_1 e ω_2, isto é, $s(p_1 p_2)$ é de esperar que essa quantidade de surpresa seja igual à soma das quantidades de surpresa $s(p_1)$ e $s(p_2)$, isto é:

$$\begin{aligned} s(p_1 p_2) &= -\log_2(p_1 p_2) \\ &= -(\log_2 p_1 + \log_2 p_2) \\ &= -\log_2 p_1 - \log_2 p_2 \\ &= s(p_1) + s(p_2) \end{aligned} .$$

2.3.3.3 Dedução axiomática do informacional

"O Hóspede - Na direção da parte a que tende nossa investigação, patenteiam-se-nos dois caminhos: um mais rápido, no caso de contrapormos uma pequena porção ao conjunto maior; o outro, de acordo, com o princípio adotado, de, sempre que possível, dividir pelo meio, é , de fato, mais longo. Depende de nós decidirmo-nos por um deles. [...] Mas esses pequeninos cortes, meu caro, não inspiram confiança; é de muito maior proveito seccionar firme pelo meio. Só assim chegaremos às idéias, que é o que importa na presente investigação."

Platão. 1980. *Diálogos*: *Político* ou *Da Realeza*, 262bc e 265a.
Tradução de Carlos Alberto Nunes. Universidade Federal do Pará.

A demonstração de que a função logarítmica é a única que se adapta ao conceito de informacional está em Ross (1998, p. 436-9).

O que representa afinal o informacional $s(p) = -\log_2 p$?

Para uma explicação mais simplificada, convém lembrar inicialmente que a TI baseia-se também em algumas observações simples para, em seguida, transformá-las em axiomas básicos, como se verá adiante.

Seja um evento ω_i, que pode ocorrer quando um experimento é executado. Quão surpresos se estaria ao ouvir que ω_i ocorreu ? Parece razoável supor que a quantidade de surpresa gerada pela informação que ω_i ocorreu depende da probabilidade de ω_i. Qual dos dois eventos resultantes do lançamento de dois dados canônicos causa maior surpresa ao ocorrer? ω_1 representa o evento "a soma dos pontos dos dados é par", cuja probabilidade é 1/2, e ω_2 representa o evento "a soma dos pontos dos dados é 12", cuja probabilidade é 1/36. Certamente a resposta é ω_2, pois sua chance de ocorrer é menor.

É possível quantificar o conceito de surpresa?

Para responder a esta pergunta impõe-se primeiramente admitir que a surpresa que se sente, ao saber que um evento ω_i tenha ocorrido, depende somente da probabilidade de ω_i, isto é, $\mathrm{P}(\omega_i) = p_i$. Indica-se esse fato com $s(p_i)$, a surpresa evocada pela ocorrência de um evento que tenha probabilidade p_i.

Determinou-se a forma funcional de $s(p_i)$ primeiramente concordando sobre uma série de condições razoáveis (axiomas) que $s(p_i)$ deve satisfazer e, então, provando que estes axiomas requerem que $s(p_i)$ tenha uma forma especificada. Admite que $s(p_i)$ seja definida para todo $0 < p_i \leq 1$, mas não para eventos impossíveis de ocorrer ($p_i = 0$).

A primeira condição consiste no estabelecimento de que não haverá nenhuma surpresa ao saber que um evento certo de fato tenha ocorrido, portanto:

- **Axioma 1**: $s(1) = 0$

A segunda condição estabelece que quanto mais improvável é a ocorrência de um dado evento, tanto maior é a surpresa evocada pela sua ocorrência.

- **Axioma 2** : $s(p_i)$ é uma função estritamente decrescente, isto é, se $p_i < q_i \Leftrightarrow s(p_i) > s(q_i)$.

A condição seguinte estabelece que, se p_i varia um pouco, isso corresponde a uma pequena variação de $s(p_i)$.

- **Axioma 3**: $s(p_i)$ é uma função contínua de p_i.

Sejam dois eventos independentes ω_i e ω_j com probabilidades p_i e p_j respectivamente. Como $\mathrm{P}(\omega_i \cap \omega_j) = \mathrm{P}(\omega_i)\mathrm{P}(\omega_j) = p_i p_j$, a surpresa evocada pela informação que ω_i e ω_j tenham ocorrido é $s\left[\mathrm{P}(\omega_i \cap \omega_j)\right] = s\left(p_i p_j\right)$. Suponha-se que inicialmente se receba a informação de que ω_i tenha ocorrido e, em seguida, que ω_j também tenha ocorrido. Como $s(p_i)$ é a surpresa evocada quando se é informado da ocorrência de ω_i, então segue-se que $s\left(p_i p_j\right) - s\left(p_i\right)$ representa a surpresa adicional evocada quando se é informado de que ω_j também ocorreu. Como ω_j é independente de ω_i, o conhecimento que ω_i tenha ocorrido não altera a probabilidade ω_j, logo a surpresa adicional é exatamente $s(p_j)$.

Esta condição dá origem ao último axioma.

- **Axioma 4**: $s\left(p_i p_j\right) = s\left(p_i\right) + s\left(p_j\right)$.

Prova-se que a única função $s(p_i)$ que satisfaz os quatro postulados é, conforme Ross (1998, p. 438):

$$s(p_i) = -c \log_2 p_i .$$

c é um número inteiro positivo qualquer. Conforme Ross (1998, p. 438), é usual adotar o valor 1 para esta constante arbitrária.

A base do logaritmo pode ser qualquer como, por exemplo, a base 2. Quando isso é feito diz-se que a surpresa associada a p_i é medida em bit. Esta é a unidade de surpresa mais comumente usada, da mesma forma que o metro o é para medidas de comprimento. A razão disso será vista em seguida.

Estabeleça-se a suposição de que se possui um conjunto de símbolos de 1 até 8 e que se escolhe um deles, por exemplo o 4. Admita-se que uma outra pessoa utilize a estratégia de seleção seletiva (a mais eficiente), em vez de uma técnica de seleção aleatória, para adivinhar qual o símbolo que foi selecionado pelo outro. A técnica consiste, portanto, em dividir ao meio o conjunto ordenado de símbolos {1, 2, 3, 4, 5, 6, 7, 8} indagando "o símbolo está antes do 5?". Se a resposta for

"sim", continuará procedendo da mesma forma , isto é, dividindo ao meio o restante dos símbolos do subconjunto {1, 2, 3, 4} e indagará "o símbolo está antes do 3?" e assim por diante.

O quadro abaixo mostra uma das possíveis seqüências de perguntas e respostas.

Pergunta	Divisão	Resposta
1. Antes do 5?	{1, 2, 3, 4} - {5, 6 7, 8}	Sim
2. Antes do 3?	{1, 2} - {3, 4}	Não
3. Antes do 4?	{3} - {4}	Não

Solução: 4

Observe-se que três perguntas do tipo sim-não (binárias) seriam suficientes para a pessoa saber qual o símbolo selecionado pelo outro. Aqui, portanto, foi determinado o número de respostas binárias para adivinhar um símbolo dentro de um conjunto ordenado {1, 2, 3, 4, 5, 6, 7, 8} cujos elementos possuem a mesma probabilidade igual a 1/8, isto é, são eqüiprováveis.

Se existir um conjunto com apenas quatro símbolos {1, 2, 3, 4}, adivinhar-se-ia o símbolo selecionado com apenas duas respostas binárias.

Suponha-se agora esse último conjunto de símbolos, mas seus elementos ocorrem na proporção 4:2:1:1. Em outras palavras, com probabilidades: p_1=1/2, p_2=1/4, p_3=1/8, p_4=1/8. Continuando a manter igual número de símbolos em torno do ponto central, a primeira pergunta seria:

Pergunta	Divisão	Resposta
1. Antes do 3?	{1, 2} - {3, 4}	Sim

Em termos de probabilidades, a escolha seria entre 1/2+1/4=3/4 e 1/8+1/8=1/4. Isso evidentemente não descreve um estado de iguais probabilidades como anteriormente. Para manter um estado de equilíbrio das probabilidades dos símbolos deve-se considerar uma lista em que cada símbolo apareça proporcionalmente a sua probabilidade, como por exemplo, o conjunto {1, 1, 1, 1, 2, 2, 3, 4}.

Agora será possível agir de acordo com a técnica da seleção seletiva. Se for escolhido, por exemplo, o símbolo 1, a primeira pergunta poderá ser:

Pergunta	Divisão	Resposta
1. Antes do 2?	{1, 1, 1, 1} - {2, 2, 3, 4}	Sim

Solução: 1

Agora uma única resposta foi suficiente para adivinhar o símbolo escolhido.

Verifica-se que, se o símbolo escolhido fosse 2, bastariam duas perguntas e, se o símbolo escolhido fosse o 3 ou 4, bastariam três perguntas para a pessoa adivinhar qual símbolo foi o escolhido. Agora é fácil compreender que o fato de restabelecer o equilíbrio das probabilidades é notório quando se pretende manter o interesse em tornar o número médio de perguntas tão reduzido quanto possível. Assim, em se colocando pura e simplesmente igual número de símbolos de cada lado do ponto central, o número médio de perguntas para adivinhar o símbolo escolhido é dois. Por outro lado, caso se restabeleça o equilíbrio das probabilidades pode-se calcular o número médio de respostas como segue:

Solução	Número de respostas	Fator de ponderação
1	1	1/2
2	2	1/4
3	3	1/8
4	3	1/8

A média ponderada de respostas ou o valor esperado do número de respostas binárias será 1,75 bit/símbolo selecionado:

Existe, portanto, uma redução de 1/4 em relação à situação de eqüiprobabilidade.

Observe-se no cálculo acima que o número de respostas binárias nada mais é do que os informacionais das probabilidades, isto é, o número de respostas para descobrir o símbolo selecionado.

Numa situação de eqüiprobabilidade, o número médio seria:

Solução	Número de respostas	Fator de ponderação
1	$2 = s(p_1) = -\log_2 p_1$	$p_1 = 1/4$
2	2	1/4
3	2	1/4
4	2	1/4

Neste caso, o valor da média do número de respostas coincide com o valor dos informacionais, isto é, 2.

O número coincide, mas as unidades são diferentes, pois os informacionais são medidos em bit e a média ou valor esperado do número de respostas ou entropia deve ser medido em bit/evento; no caso presente, o evento é o símbolo solução.

Portanto, generalizando em termos mais formais, a entropia de um espaço amostral é:

$$\begin{aligned} E_{\mathbb{P}}(X_{\mathbb{P}}) &= h_{\mathbb{P}}(X_{\mathbb{P}}) \\ &= -\sum_{i=1}^{n} p_i \log_2 p_i \text{ bit/evento} \end{aligned}$$

Conforme Edwards (1971, p. 58), fixou-se uma medida de incerteza com base no índice de dificuldade existente para identificar um dos elementos de certo conjunto. O método empregado, o de reduzir ao meio as probabilidades, corresponde, em essência, ao apresentado por Fano[23] (1949 e 1950).

2.3.3.4 Entropia

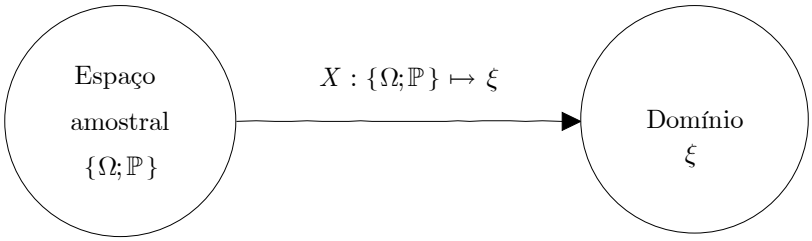
Generalizando, existe uma variável aleatória discreta X associada a n eventos mutuamente excludentes e coletivamente exaustivos $\Omega = \{\omega_i; i = 1,2,...,n\}$, que adquirem valores $\xi = \{x_i \in \mathbb{R}; i = 1,2,...n\}$ do domínio ou conjunto de apoio e que ocorrem com probabilidades $\mathbb{P} = \{p_i; i = 1,2,...n\}$, respectivamente. $\mathbb{P}$ será chamada

[23] Fano, R. M. 1949 (1950). "The transmition of information I (II)" *M.I.T. Technical Report*, 65 (149). Cambridge: Research Lab. of Electronics.

de medida de probabilidade. A variável aleatória discreta X pode ser representada de forma tabular:

Evento ω_i	*Probabilidade* $\Pr(\omega_i) = p_i$	*Valor* x_i
ω_1	$\Pr(\omega_1) = p_1$	x_1
ω_2	$\Pr(\omega_2) = p_2$	x_2
$\vdots$	$\vdots$	$\vdots$
ω_n	$\Pr(\omega_n) = p_n$	x_n

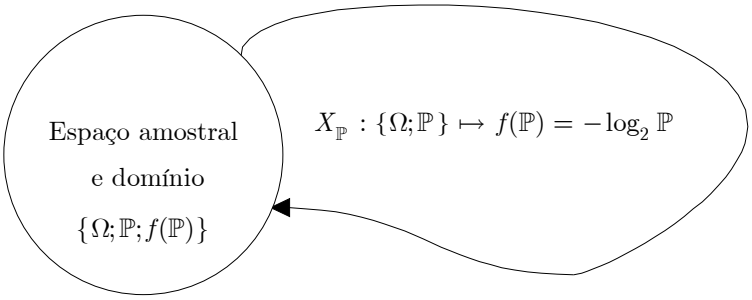
Graficamente:



As quantidades de surpresa correspondentes a cada um dos eventos são dadas pelos informacionais $f(\mathbb{P}) = \{s(p_i) = -\log_2 p_i; i = 1,2,...n\}$. Considerar-se-á a existência de uma autovariável aleatória discreta $X_{\mathbb{P}}$ cujo domínio é constituído pelos informacionais. Esta nova variável aleatória será chamada de autovariável aleatória, pois seus valores são resultantes de uma função não-linear das próprias probabilidades. De forma tabular, tem-se:

Evento ω_i	*Probabilidade* $\Pr(\omega_i) = p_i$	*Valor=Informacional (bit)* $x_i = s(p_i) = -\log_2 p_i$
ω_1	$\Pr(\omega_1) = p_1$	$x_1 = s(p_1) = -\log_2 p_1$
ω_2	$\Pr(\omega_2) = p_2$	$x_2 = s(p_2) = -\log_2 p_2$
$\vdots$	$\vdots$	$\vdots$
ω_n	$\Pr(\omega_n) = p_n$	$x_n = s(p_n) = -\log_2 p_n$

Graficamente:

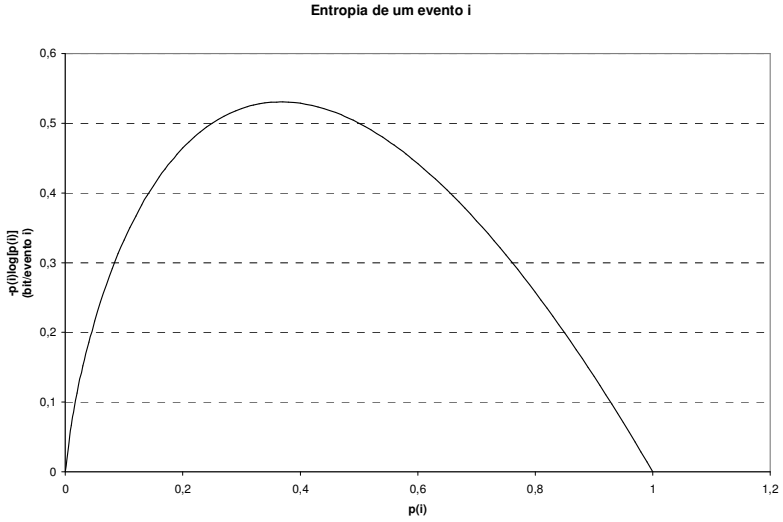


Desta forma, pode-se calcular o informacional esperado (médio) de uma variável aleatória. Matematicamente, tem-se o valor esperado do informacional na medida de probabilidade $\mathbb{P}$ é:

$$\begin{aligned} E_{\mathbb{P}}(X_{\mathbb{P}}) &= \sum_{i=1}^{n} p_i \left(-\log_2 p_i\right) \\ &= -\sum_{i=1}^{n} p_i \log_2 p_i \text{ bit/evento} \end{aligned}$$

em que $0 \log_2 0 = 0$, por definição.

$E_{\mathbb{P}}(\bullet)$ é o operador linear, chamado **esperança matemática**, que executa o produto escalar entre as probabilidades e os valores da variável aleatória.



O valor esperado do informacional na medida de probabilidade $\mathbb{P}$ é também chamado de **entropia de Shannon** e é simbolizada por $h_{\mathbb{P}}(\bullet)$. Portanto:

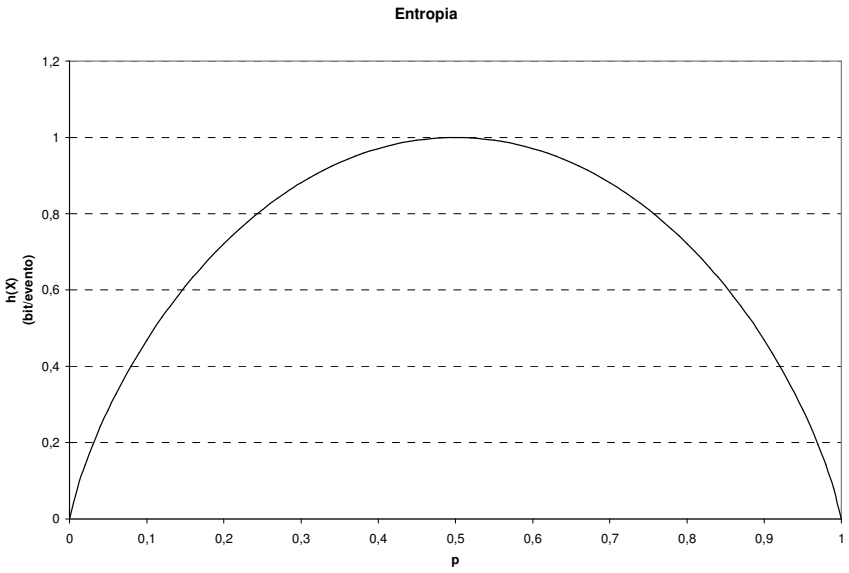
$$h_{\mathbb{P}}(X_{\mathbb{P}}) = E_{\mathbb{P}}(X_{\mathbb{P}}) .$$

$h_{\mathbb{P}}(X_{\mathbb{P}})$ representa a média das quantidades de surpresa ponderadas pelas respectivas probabilidades que se recebe ao saber da ocorrência de um evento qualquer do espaço amostral. A entropia pode também ser interpretada como a quantidade média de informação a ser recebida quando ocorre um evento qualquer do espaço amostral.

Para o caso de um experimento de Bernoulli, a entropia da variável aleatória com probabilidade p de sucesso $\big(X \sim \text{Bernoulli}(p) \big)$ é:

$$h_{\mathbb{P}}(X_{\mathbb{P}}) = -p \log_2 p - (1-p) \log_2 (1-p) \,.$$

Pelo gráfico a seguir, observa-se que a entropia da variável aleatória de Bernoulli apresenta o valor máximo de entropia 1 para a probabilidade de sucesso igual a ½.



O exemplo a seguir será útil para estabelecer as unidades de medida do informacional e da entropia.

2.3.3.5 Exemplo paradigmático do jogo de adivinhar o número pensado por outra pessoa

Edwards (1971, p. 44) propõe a seguinte pergunta para definir a unidade do informacional: "Que grau de dificuldade existe para a identificação de um

elemento particular de certo conjunto?" Portanto, é necessária a definição do conjunto, alfabeto ou espaço amostral.

Lauand (1986, p. 111-7) relata um problema proposto na Idade Média. O problema consiste em adivinhar o número pensado por outra pessoa. Edwards (1971, p. 44) propõe um jogo disputado entre duas pessoas no qual cada uma escolhe uma letra do alfabeto, e ganha o jogo quem descobrir primeiro a letra escolhida pelo outro. O novo jogo a ser proposto é uma combinação dos dois anteriores. Imagine-se um jogo no qual cada um dos jogadores (honestos) seleciona um número do conjunto de números naturais de 1 a 8. O jogo será jogado várias vezes. Os jogadores revezam-se nas perguntas cujas respostas são necessariamente **sim** ou **não**. O jogo termina quando um dos jogadores acerta o número escolhido pelo outro jogador.

O estado informacional inicial de cada um deles, em relação ao número escolhido pelo outro jogador, é o de máxima ignorância, isto é, todos os números possuem a mesma chance de terem sido escolhidos (1/8).

O jogador A armou a estratégia de perguntar se o outro jogador escolheu o número 1. Caso não tenha sido este o número escolhido, a próxima pergunta será se o número escolhido foi o 2 e assim por diante. Evidentemente, se o outro jogador escolheu o número 1, o jogo termina na primeira pergunta deste jogador. No entanto, se o número escolhido foi o 7 ou o 8, sete perguntas têm que ser feitas, uma vez que o número médio de perguntas/respostas binárias a fazer para identificar o número selecionado pelo oponente é 4. Isto é, se o conjunto de números naturais tiver n elementos distintos e eqüiprováveis, haverá $n/2$ respostas binárias, em média, para identificar o número selecionado pelo adversário. No caso de o conjunto conter os números de 1 a 1.024, o número médio de respostas binárias será 512. Esta estratégia equivale a perguntar aleatoriamente, com a agravante de o jogador A ter que manter o controle dos números perguntados para não correr o risco de repeti-los.

O jogador B, tendo feito recentemente um curso de Teoria da Informação, planejou perguntar inicialmente se o número selecionado pelo seu contendedor é menor ou igual a 4. Caso a resposta seja afirmativa, ele irá perguntar se o número

é menor ou igual a 2 e assim por diante. Como pode ser percebido, a quantidade de perguntas/respostas binárias necessárias para que o número selecionado pelo rival seja identificado é sempre 3, isto é, $\log_2 8 = \log_2 2^3 = 3$. Portanto, o número médio de respostas binárias também será 3. Isto é, se o conjunto de números naturais tiver n elementos distintos e equiprováveis, $\log_2 n$ respostas binárias têm que ser feitas, em média (neste caso exatamente), para identificar o número selecionado pelo competidor. No caso de o conjunto conter os números de 1 a 1.024, o número de respostas binárias será 10. Note que esta é a melhor estratégia possível [Edward (1971, p. 46)].

Parece razoável estabelecer a unidade do grau de dificuldade para identificar um evento do espaço amostral, selecionado pela natureza, com base no número de respostas **sim-não** que se tornem necessárias quando o processo ótimo é empregado. Esta quantidade $\log_2 n = -\log_2 \frac{1}{n}$ respostas **sim-não** é a unidade do informacional. Portanto, diz-se, no jogo com oito números, que a informação "o número que o jogador selecionou foi o x", contém 3 unidades de informação ou de incerteza. Note-se que a quantidade de incerteza é o número mínimo de perguntas necessárias e suficientes para sair de um estado de completa incerteza para um estado de completa certeza.

Uma resposta **sim-não** será chamada de bit (*binary digit*). A palavra **bit** foi criada por J. W. Tukey. Portanto, um evento equiprovável de um espaço amostral de tamanho n possui $\log_2 n = -\log_2 \frac{1}{n}$ bit de informação que será simbolizado por s.

Como a quantidade de informação relaciona-se diretamente com o tamanho do conjunto, pode-se aplicar a medida a ele mesmo. No caso do jogo com oito números equiprováveis, o conjunto possui 3 bit/número. Esta quantidade de informação é chamada de **entropia** e será simbolizada por h[24].

---

[24] Conforme Cagliari (1998, p. 24-5), H é a letra maiúscula **eta** do alfabeto grego. η é a letra minúscula **eta** do alfabeto grego. Esta última corresponde à letra **h** do alfabeto romano. A letra **h** vem da letra semítica **heth**. O grego usou essa letra

$$
\begin{aligned}
h &= \sum_{i=1}^{n} \frac{1}{n} \log_2 n \\
&= -\sum_{i=1}^{n} \frac{1}{n} \log_2 \frac{1}{n} \\
&= -\log_2 \frac{1}{n} \\
&= \log_2 n \text{ bit/numero}
\end{aligned}
$$

Continuando o jogo, imagine-se que o jogador B (que parece ser mais inteligente que o A) descobre que, a partir deste momento, o seu rival selecionará números com as seguintes probabilidades:

$$
\begin{aligned}
&\mathrm{Pr}(1) = \frac{1}{2} \\
&\mathrm{Pr}(2) = \frac{1}{4} \\
&\mathrm{Pr}(3) = \mathrm{Pr}(4) = \frac{1}{8} \\
&\mathrm{Pr}(5) = \mathrm{Pr}(6) = \mathrm{Pr}(7) = \mathrm{Pr}(8) = 0
\end{aligned}
$$

Este conjunto pode ser visto da seguinte maneira: 11112234.

Portanto, os informacionais dos números do novo conjunto são:

Número	Informacional (bit)
1	$-\log_2 \frac{1}{2} = 1$
2	$-\log_2 \frac{1}{4} = 2$
3 e 4	$-\log_2 \frac{1}{8} = 3$

A entropia desse conjunto é 1,75 bit/número:

$$
\begin{aligned}
h &= \frac{1}{2} * 1 + \frac{1}{4} * 2 + \frac{1}{8} * 3 + \frac{1}{8} * 3 \\
&= \frac{1}{2} + \frac{1}{2} + \frac{3}{8} + \frac{3}{8} \\
&= 1{,}75 \text{ bit/numero}
\end{aligned}
.
$$

---

para representar a vogal longa **eta**. A origem egípcia dessa letra é o hieróglifo que significa 'cerca' ou 'corda trançada'.

Isto é, o número médio mínimo de respostas binárias necessárias para identificar o número escolhido pelo contendor é 1,75 por número.

Por exemplo, imagine-se que o número escolhido pelo jogador A tenha sido o 1. Como a primeira pergunta do jogador B é "O número é menor que 2?" ou, equivalentemente, "O número é o 1?", ele pode acertar o número selecionado com apenas um pergunta com a probabilidade de ½, isto é, em metade das partidas disputadas.♦

2.3.3.6 Entropia de algumas variáveis aleatórias discretas

Ebrahimi *et al.* (1999) apresentam uma tabela com a entropia de algumas das principais variáveis aleatórias discretas: binomial, geométrica, Poisson e uniforme.

Quadro 2: Variância e entropia de algumas distribuições discretas

Função de massa de probabilidade $\mathbb{P}$	*Variância* $Var_{\mathbb{P}}(X)$	*Entropia (nit/evento)* $h_{\mathbb{P}}(X_{\mathbb{P}})$	*Relação*
$X \sim$ uniforme(n) $\Pr(x \mid n) = \dfrac{1}{n}$ $x = 1,2,\ldots,n$	$\dfrac{n^2-1}{12}$	$\ln n$	$\uparrow n[\uparrow Var_{\mathbb{P}}(X);\uparrow h_{\mathbb{P}}(X_{\mathbb{P}})]$
$X \sim$ geometrica(q) $\Pr(x \mid q) = q(1-q)^x$ $x = 0,1,\ldots$	$\dfrac{1-q}{q^2}$	$-\ln q - \dfrac{(1-q)\ln(1-q)}{q}$	$\uparrow q[\downarrow Var_{\mathbb{P}}(X);\downarrow h_{\mathbb{P}}(X_{\mathbb{P}})]$
$X \sim$ binomial$(n;q)$ $X \sim$ Bernoulli$(1;q)$ $\Pr(x \mid n;q) = \binom{n}{x} q^x (1-q)^{n-x}$ $x = 0,1,\ldots,n$	$nq(1-q)$	$-nq\ln q - n(1-n)\ln(1-q)$ $-\ln\Gamma(n+1)$ $+\sum_{k=0}^{n}[\ln\Gamma(k+1)+\ln\Gamma(n-k+1)]\,p(k \mid n;q)$	$\uparrow q > \frac{1}{2}[\downarrow Var_{\mathbb{P}}(X);\downarrow h_{\mathbb{P}}(X_{\mathbb{P}})]$ $\uparrow q < \frac{1}{2}[\uparrow Var_{\mathbb{P}}(X);\uparrow h_{\mathbb{P}}(X_{\mathbb{P}})]$
$X \sim$ Poisson(λ) $\Pr(x \mid \lambda) = \dfrac{\exp(-\lambda)\lambda^x}{x!}$ $x = 0,1,\ldots$	λ	$\lambda - \lambda\ln\lambda$ $+\sum_{k=0}^{\infty}[\ln\Gamma(k+1)]\,p(k \mid \lambda)$	$\uparrow \lambda[\uparrow Var_{\mathbb{P}}(X);\uparrow h_{\mathbb{P}}(X_{\mathbb{P}})]$

Observe-se que a entropia acompanha o sentido de aumento/diminuição da variância quando os parâmetros das distribuições variam. Tanto a variância

quanto a entropia dependem apenas dos parâmetros da distribuição. Conforme Ebrahimi *et al.* (1999, p. 318), a entropia é também uma medida de dispersão, incerteza, avaliação de ajuste etc. Portanto, ela concorre com a variância. Ambas as medidas medem a idéia de concentração, mas, enquanto a variância mede concentração em torno da medida de posição central chamada **média** (valor esperado da distribuição), a entropia mede dispersabilidade da distribuição independentemente de uma medida de posição [Ebrahimi *et al.* (1999, p. 320)].

No caso discreto a variância é estritamente positiva enquanto a entropia é não negativa.

No caso discreto a entropia é invariante com relação a uma transformação um a um dos valores do domínio. A variância não é invariante com relação a este tipo de transformação. Por exemplo, imagine-se uma variável aleatória X que tenha distribuição de Bernoulli com parâmetro q (probabilidade de sucesso).

Evento ω_i	*Probabilidade* $\Pr(\omega_i) = p_i$	*Valor* x_i
ω_1 =sucesso	$\Pr(\omega_1) = q$	1
ω_2 =fracasso	$\Pr(\omega_2) = 1 - q$	0

Observa-se que a esperança é q e a variância é $q(1\text{-}q)$.

A entropia vale $-q\log_2 q - (1-q)\log_2(1-q)$.

Imagine-se agora que façamos a seguinte transformação $2X$.

Evento ω_i	*Probabilidade* $\Pr(\omega_i) = p_i$	*Valor* x_i
ω_1 =sucesso	$\Pr(\omega_1) = q$	2
ω_2 =fracasso	$\Pr(\omega_2) = 1 - q$	0

A entropia certamente não muda, pois ela só depende da medida de probabilidade.

No entanto, a esperança muda para $2q$, e a variância, para $4q(1-q)$.

2.3.4 Divergência de Kullback-Leibler

"O executivo de sucesso é antes de tudo um previsor; comprar, produzir, vender, fixar preços e organizar vêm depois."

Bernstein (1997, p. 21)

Uma outra grandeza definida pela TI é a **divergência** de Kullback-Leibler. Ela também é conhecida como **entropia relativa** ou **cross-entropy**. O exemplo paradigmático a seguir contribui para o exato entendimento e formalização matemática deste importante conceito.

2.3.4.1 Mensuração do ganho informacional da previsão e da eficiência informacional relativa da previsão usando o conceito de divergência: uma extensão do análise bayesiana

O Teorema de Bayes gera a distribuição de probabilidade posterior a partir das probabilidades prévias e das características do previsor (sensibilidade e especificidade). No entanto, a teoria Estatística Bayesiana da decisão não permite responder às questões de quanto de informação é ganha com o serviço do previsor, e qual e quão melhor é o previsor. Será mostrado como a Teoria da Informação usa os resultados da Teoria da Decisão para responder a estas perguntas: "Qual o ganho informacional proporcionado por uma previsão para o decisor responsável?", "Quão útil é um previsor para um decisor?" e "Quão melhor um previsor é em relação a outro?".

2.3.4.1.1 Exemplo paradigmático da compra de um carro usado

Imagine-se que uma pessoa esteja pensando em adquirir um carro do tipo C da ED's Honesto. A fim de tomar uma decisão informada, observam-se os dados disponíveis referentes aos carros C numa revista especializada de automóveis e encontra-se que, infelizmente, 30% têm defeito de transmissão. Para ter mais informação sobre o carro C da ED's, contrata-se um mecânico que possa fazer um julgamento técnico depois de uma rápida volta de quarteirão na direção do carro. Evidentemente, o mecânico embora não diagnostique sempre corretamente, possui uma excelente folha de serviço: de todos os carros defeituosos que ele examinou no passado, diagnosticou corretamente 90% das vezes (afirmou que o

carro do tipo C era "defeituoso" quando de fato ele era defeituoso). Em outras palavras, ele diagnosticou incorretamente somente 10% (afirmou que o carro do tipo C era "ok" quando de fato era defeituoso). Ele tem também uma boa folha de serviço de carros em bom estado (sem defeito de transmissão): pronunciou 80% de "ok" quando o carro era de fato ok.

A expressão "defeituoso", entre aspas, descreve a previsão (opinião) do mecânico e **defeituoso**, sem aspas, descreve o estado real do carro C, analogamente para "ok" e **ok**.

2.3.4.2 Estados informacionais do decisor

Há quatro estados informacionais do comprador do carro C:

- Estado informacional 1: após o conhecimento sobre o defeito de transmissão no carro C e antes da contratação do mecânico;
- Estado informacional 2: após a contratação do mecânico e antes de seu diagnóstico;
- Estado informacional 3: após o diagnóstico do mecânico e antes da revelação do estado real do carro C;
- Estado informacional 4: revelação do estado real do carro C.

Os três primeiros estados informacionais são estados de incerteza ou de informação incompleta sobre o estado do carro C. O estado informacional 4 é o estado informacional de interesse do comprador do carro C, i.e., é o estado de certeza. Quanto menor a incerteza sobre o real estado do carro C, melhor para o decisor (comprador do carro C) pelo princípio da aversão à incerteza.

1) Estado informacional 1: após o conhecimento sobre o defeito de transmissão no carro C e antes da contratação do mecânico

Qual a chance de que o carro C, que se está querendo comprar de ED's, tenha defeito de transmissão?

A resposta a esta pergunta é a seguinte: do universo de carros C, a chance de selecionar ao acaso um carro defeituoso é de 0,30. O carro C da ED's é um carro

selecionado ao acaso dentro do universo de carros C, portanto há 30% de chance desse carro ser defeituoso. Essa é a única informação que se tem sobre o estado do carro C.

Logo, pode-se definir um espaço amostral com dois estados da natureza: **ok** e **defeituoso**, com as seguintes probabilidades:

Evento	Probabilidade
Ok	$P(\text{ok}) = 0,70$
Defeituoso	$P(\text{defeituoso}) = 0,30$

2) Estado informacional 2: após a contratação do mecânico e antes de seu diagnóstico

Neste estado informacional, o comprador (decisor responsável) conta com as informações sobre as características do mecânico (sensibilidade=0,90 e especificidade=0,80 do previsor imperfeito), além das informações sobre o carro C do estado informacional 1 (probabilidades prévias).

Neste novo estado informacional serão determinadas as probabilidades posteriores utilizando o teorema de Bayes.

Neste estado informacional, deseja-se saber qual a chance de que o carro C que se está querendo comprar tenha defeito de transmissão se o mecânico pronunciar: (1) "defeituoso" ou (2) "ok".

Para responder a estas perguntas e a outras, deve-se empregar o Teorema de Bayes.

Kinney (1997, p. 22-6) e Wonnacott & Wonnacott (1990, cap. 19) representam graficamente, de forma visual, o Teorema de Bayes. Fica, desta forma, evidente que a fórmula de inversão da probabilidade condicional é a razão das áreas de dois retângulos cujos lados são menores ou iguais a um. Em primeiro lugar, já é difícil estimar a área de um retângulo com lados, em geral, menores que um. Em segundo lugar, é também difícil estimar a razão de duas áreas originadas pelo cálculo anterior. A interação das dificuldades dos dois cálculos anteriores torna a

estimação da probabilidade posterior um exercício de difícil solução mental. Entra em cena de forma determinante o teorema desenvolvido pelo pastor inglês Thomas Bayes (1701-1761) no *Ensaio sobre a solução de um problema da doutrina das chances* [Bernstein (1997, p. 128)].

O Teorema de Bayes pode ser entendido como uma forma de obter a árvore de probabilidades do interessado, invertendo a árvore de probabilidade da natureza [Bekman & Costa Neto (1980, 15-6)].

Para isso constrói-se a árvore de probabilidades da natureza:

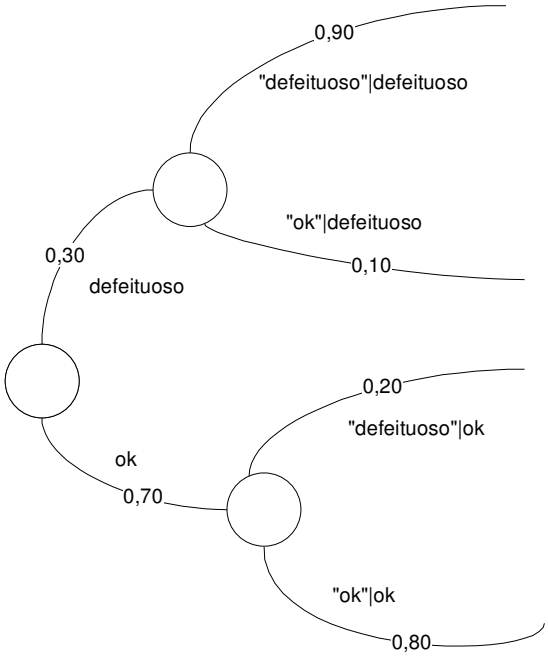


Figura 1: árvore de probabilidades da natureza

Pelo Teorema de Bayes, uma vez que a probabilidade posterior ou revisada do carro C ser defeituoso, dado que o mecânico diagnosticou que ele é "defeituoso", é:

$$\begin{aligned}
\mathrm{P}(\text{defeituoso} \mid \text{"defeituoso"}) &= \frac{\mathrm{P}(\text{"defeituoso"} \mid \text{defeituoso}) * \mathrm{P}(\text{defeituoso})}{\mathrm{P}(\text{"defeituoso"})} \\
&= \frac{\mathrm{P}(\text{"defeituoso"} \mid \text{defeituoso}) * \mathrm{P}(\text{defeituoso})}{\mathrm{P}(\text{"defeituoso"} \mid \text{defeituoso}) * \mathrm{P}(\text{defeituoso}) + \mathrm{P}(\text{"defeituoso"} \mid \text{ok}) * \mathrm{P}(\text{ok})} \\
&= 0,66
\end{aligned}$$

Logo, uma vez o carro tenha sido diagnosticado pelo mecânico como "defeituoso", a chance de que ele seja realmente defeituoso aumenta dos originais 30% para 66%.

Pode-se concluir também que:

$$\begin{aligned} \mathrm{P}(\text{ok} \mid \text{"defeituoso"}) &= 1 - \mathrm{P}(\text{defeituoso} \mid \text{"defeituoso"}) \\ &= 1 - 0,66 \\ &= 0,34 \end{aligned}$$

Caso o mecânico diagnostique como “ok”, as seguintes probabilidades posteriores são calculadas:

$$\begin{aligned} \mathrm{P}(\text{defeituoso} \mid \text{"ok"}) &= 0,05 \\ \mathrm{P}(\text{ok} \mid \text{"ok"}) &= 1 - \mathrm{P}(\text{defeituoso} \mid \text{"ok"}) \\ &= 1 - 0,05 \\ &= 0,95 \end{aligned}$$

Os resultados da figura 1 podem ser sumariados em uma outra árvore, a árvore reversa ou do interessado, que é a árvore das probabilidades posteriores. Em outras palavras, trata-se da árvore que exprime o ponto de vista do interessado sobre a compra do carro C, dado o julgamento do mecânico. Isso não envolve novos cálculos, mas apenas trata-se uma nova forma de dispor os cálculos já feitos.

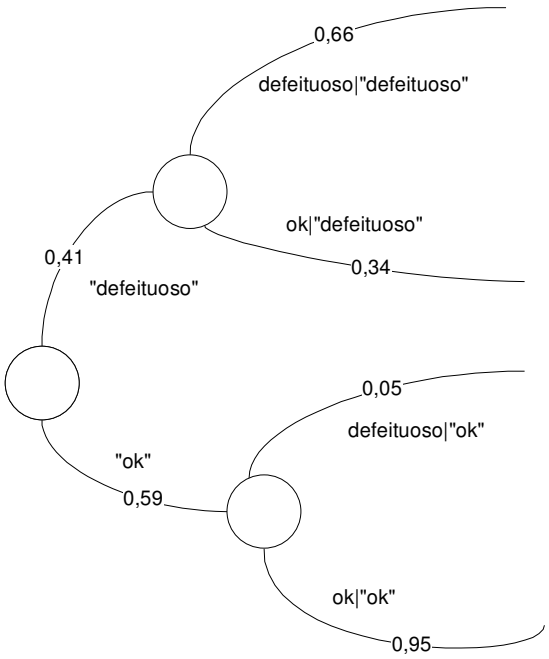


Figura 2: árvore de probabilidades do interessado

Por exemplo, a árvore do interessado dá a resposta à pergunta 1, isto é, P(defeituoso|"defeituoso")=0,66. Uma vez que o carro C de ED's tenha sido julgado pelo mecânico como "defeituoso", a chance de que ele seja verdadeiramente defeituoso torna-se 66%.

No terceiro ramo a partir do topo temos a resposta à pergunta 2, isto é, P(defeituoso|"ok")=0,05. Uma vez o carro C seja julgado "ok", a chance que ele seja defeituoso é de apenas 5%.

As probabilidades prévias, a característica do previsor imperfeito e as probabilidades posteriores relacionam-se pela seguinte expressão:

$$\frac{P(\text{defeituoso} \mid \text{"defeituoso"})}{P(\text{ok} \mid \text{"defeituoso"})} = \frac{P(\text{"defeituoso"} \mid \text{defeituoso})}{P(\text{"defeituoso"} \mid \text{ok})} * \frac{P(\text{defeituoso})}{P(\text{ok})}$$

A razão das probabilidades prévias P(defeituoso)/P(ok) é chamada de **vantagem prévia em favor de defeituoso**.

A razão das probabilidades posteriores P(defeituoso|"defeituoso")/P(ok|"defeituoso") é chamada de vantagem posterior a favor de defeituoso.

Finalmente, a razão entre as probabilidades P("defeituoso"|defeituoso)/P("defeituoso"|ok) é chamada de **fator de Bayes ou verossimilhança**.

Outra forma de representar a reversão das probabilidades (apresentação das probabilidades posteriores) é a de inicialmente apresentar em forma tabular as probabilidades conjuntas, como se segue:

	Estado real do carro C		Probabilidade marginal
Diagnóstico do mecânico	defeituoso	ok	
"defeituoso"	p_{11}	p_{12}	$p_{1.}$
"ok"	p_{21}	p_{22}	$p_{2.}$
Probabilidade marginal	$p_{.1}$	$p_{.2}$	1

No exemplo dado temos:

	Estado real do carro C		Probabilidade marginal
Diagnóstico do mecânico	defeituoso	ok	
"defeituoso"	0,27	0,14	0,41
"ok"	0,03	0,56	0,59
Probabilidade marginal	0,30	0,70	1

Agora é possível apresentar a tabela das probabilidades posteriores:

	Estado real do carro C		
Diagnóstico do mecânico	defeituoso	ok	Probabilidades marginais
"defeituoso" ($\mathbb{T}^{\text{"def"}}$)	0,27/0,41=0,66	0,14/0,41=0,34	1
"ok" ($\mathbb{T}^{\text{"ok"}}$)	0,03/0,59=0,05	0,56/0,59=0,95	1

A razão g^{+}=P(defeituoso|"defeituoso")/P(defeituoso)=0,66/0,30=2,20 indica que, quando o julgamento do mecânico é "defeituoso", a probabilidade de o carro ser defeituoso aumenta 2,20 vezes em relação à probabilidade prévia.

Analogamente, a razão g^{-}=P(defeituoso|"ok")/P(defeituoso)=0,05/0,30=0,17 indica que, quando o julgamento do mecanico é "ok", a probabilidade de o carro ser defeituoso diminui para 17% da probabilidade prévia.

2.3.4.3 Análise informacional da decisão em condição de incerteza

Para analisar um problema qualquer, a TI se vale de três princípios, a saber:

1. Princípio do uso da máxima informação disponível: um decisor racional deve usar adequadamente toda a informação disponível;
2. Princípio da comparação de esperanças ou entropias: duas esperanças ou entropias são comparáveis se elas estão definidas na mesma medida de probabilidade;
3. Princípio de aversão à incerteza.

Por outro lado, é importante considerar o que se entende por "estado do conhecimento ou estado informacional de uma pessoa interessada ou decisor na solução de um problema em condição de incerteza". Trata-se do conhecimento que o decisor responsável possui do conjunto de todas as distribuições de probabilidades disponíveis num dado instante do tempo.

Considera-se o espaço amostral chamado "estado do carro C", que apresenta dois eventos: carro C com defeito de transmissão, associado à probabilidade P(defeituoso)=0,30 e carro C ok (sem defeito de transmissão), associado à probabilidade P(ok)=0,70.

Nesse espaço amostral, define-se uma variável aleatória associada aos dois eventos, que adquire os seguintes valores:

$$\begin{aligned} x_{\text{defeituoso}} &= s(\text{P(defeituoso)}) \\ &= -\log_2[\text{P(defeituoso)}] \\ x_{\text{ok}} &= s(\text{P(ok)}) \\ &= -\log_2[\text{P(ok)}] \end{aligned}$$

Chama-se $s(p)$ o informacional do evento associado à probabilidade p. Sua unidade é o bit.

O vetor de probabilidades $\mathbb{P}$={P(defeituoso); P(ok)} está associado ao domínio constituído por informacionais ξ = {s[P(ok)]; s[P(defeituoso)]}.

Indica-se a variável aleatória acima mencionada por $X_{\mathbb{P}} : \mathbb{P} \mapsto \xi$.

O valor médio ou esperado da variável aleatória $X_{\mathbb{P}}$ é o que chamamos de **entropia**, isto é:

$$\begin{aligned} E_{\mathbb{P}}(X_{\mathbb{P}}) &= h_{\mathbb{P}}(X_{\mathbb{P}}) \\ &= -\mathrm{P(defeituoso)} * \log_2 \mathrm{P(defeituoso)} - \mathrm{P(ok)} * \log_2 \mathrm{P(ok)} \end{aligned}$$

O índice $\mathbb{P}$ indica que o valor esperado é calculado nesta distribuição de probabilidade. A unidade da entropia é bit/estado do carro C.

A entropia mede a quantidade de incerteza ou de surpresa esperada que o decisor teria ao vir a conhecer o verdadeiro estado do carro C, isto é, ela mede a quantidade de surpresa entre um determinado estado de conhecimento com incerteza e o de certeza.

No exemplo considerado, este é o primeiro estado de conhecimento por parte do decisor na compra do carro C, pois as únicas probabilidade disponíveis são as prévias, fornecidas pela revista especializada. Tem-se, assim:

$$\begin{aligned} h_{\mathbb{P}}(X_{\mathbb{P}}) &= -0,3 * \log_2 0,3 - 0,70 * \log_2 0,70 \\ &= 0,88 \text{ bit/estado do carro C} \end{aligned}$$

3) Estado informacional 3: após o diagnóstico do mecânico e antes da revelação do estado real do carro C

Veja-se agora o terceiro estado de conhecimento quando se dispõe das informações da revista (probabilidades prévias) e da disponibilidade das qualidades do mecânico (probabilidades posteriores), e o resultado apresentado pelo teste do mecânico.

Como se viu anteriormente, conhecidas as probabilidades prévias e as qualificações do mecânico, com auxílio do Teorema de Bayes é possível também conhecer as probabilidades posteriores. Este é o estado de máximo conhecimento do comprador do carro C.

Para este estado de conhecimento quais probabilidades e informacionais devem ser utilizados para cálculo da entropia?

Evidentemente, aquelas de máximo conhecimento, isto é, as probabilidades posteriores, de acordo com o princípio 1.

Deve-se, aqui, considerar as duas possíveis respostas do mecânico após o seu julgamento, a saber, "defeituoso" ou "ok".

Se a opinião do mecânico for "defeituoso", a entropia será:

$$\begin{aligned} h_{\mathbb{T}^{\text{"def"}}}\left(X_{\mathbb{T}^{\text{"def"}}}\right) &= -\mathrm{P}(\text{def} \mid \text{"def"}) * \log_2 \mathrm{P}(\text{def} \mid \text{"def"}) - \mathrm{P}(\text{ok} \mid \text{"def"}) * \log_2 \mathrm{P}(\text{ok} \mid \text{"def"}) \\ &= 0{,}92 \text{ bit/estado do carro C} \end{aligned}$$

A entropia $h_{\mathbb{T}^{\text{"def"}}}\left(X_{\mathbb{T}^{\text{"def"}}}\right)$ mede a quantidade média de surpresa ou de incerteza entre o estado de conhecimento 3, quando o mecânico diagnostica o carro C como “defeituoso”, e o estado 4 de certeza.

Se a opinião do mecânico for "ok", a entropia será:

$$\begin{aligned} h_{\mathbb{T}^{\text{"ok"}}}\left(X_{\mathbb{T}^{\text{"ok"}}}\right) &= -\mathrm{P}(\text{def} \mid \text{"ok"}) * \log_2 \mathrm{P}(\text{def} \mid \text{"ok"}) - \mathrm{P}(\text{ok} \mid \text{"ok"}) * \log_2 \mathrm{P}(\text{ok} \mid \text{"ok"}) \\ &= 0{,}29 \text{ bit/estado do carro C} \end{aligned}$$

A entropia $h_{\mathbb{T}^{\text{"ok"}}}\left(X_{\mathbb{T}^{\text{"ok"}}}\right)$ mede a quantidade média de surpresa ou de incerteza entre o estado de conhecimento 3, quando o mecânico diagnostica o carro C como “ok”, e o estado 4 de certeza.

É importante observar que não se podem comparar estas duas últimas entropias com a do primeiro estado de conhecimento, posto que elas são distribuições de probabilidades diferentes (de acordo com o princípio 2).

Considere-se, finalmente, um estado de conhecimento intermediário 2, aquele em que se conhecem as probabilidades prévias fornecidas pela revista, as qualificações das opiniões do mecânico, porém não se conhece o resultado do seu diagnóstico.

Numa situação como esta, podem-se usar as probabilidades posteriores das qualificações do mecânico, mas não se podem usar seus informacionais, pois estes não estão disponíveis ainda.

Quais informacionais podem ser usadas? Aqueles da pessoa interessada, que ainda são os do estado informacional 1 (inicial).

Também neste estado do conhecimento 2 devem-se considerar as probabilidades relativas às duas possíveis previsões do mecânico.

Tem-se, portanto:

$$\begin{aligned} h_{\mathbb{T}^{\text{"def"}}}\left(X_{\mathbb{P}}\right) &= -\mathrm{P}(\text{def} \mid \text{"def"}) * \log_2 \mathrm{P}(\text{def}) - \mathrm{P}(\text{ok} \mid \text{"def"}) * \log_2 \mathrm{P}(\text{ok}) \\ &= 1,32 \text{ bit/estado do carro C} \end{aligned}$$

e

$$\begin{aligned} h_{\mathbb{T}^{\text{"ok"}}}\left(X_{\mathbb{P}}\right) &= -\mathrm{P}(\text{def} \mid \text{"ok"}) * \log_2 \mathrm{P}(\text{def}) - \mathrm{P}(\text{ok} \mid \text{"ok"}) * \log_2 \mathrm{P}(\text{ok}) \\ &= 0,58 \text{ bit/estado do carro C} \end{aligned}$$

É importante observar que se podem comparar $h_{\mathbb{T}^{\text{"def"}}}\left(X_{\mathbb{T}^{\text{"def"}}}\right)$ e $h_{\mathbb{T}^{\text{"def"}}}\left(X_{\mathbb{P}}\right)$, pois as duas entropias estão definidas na mesma distribuição de probabilidade (princípio 2). O mesmo ocorre com as duas outras entropias.

Verifica-se também que:

$$h_{\mathbb{T}^{\text{"def"}}}\left(X_{\mathbb{P}}\right) > h_{\mathbb{T}^{\text{"def"}}}\left(X_{\mathbb{T}^{\text{"def"}}}\right)$$

e

$$h_{\mathbb{T}^{\text{"ok"}}}\left(X_{\mathbb{T}^{\text{"ok"}}}\right) > h_{\mathbb{T}^{\text{"ok"}}}\left(X_{\mathbb{T}^{\text{"ok"}}}\right).$$

Por que essas desigualdades ocorrem?

Basta notar que, ao tomar conhecimento da opinião do mecânico (informação adicional) elimina-se uma certa quantidade de surpresa ou de incerteza, ou melhor, ganha-se mais informação sobre o verdadeiro estado do carro C.

O que faz as inequações anteriores transformarem-se em equações são as divergências (valores positivos) que são subtraídas do lado esquerdo das desigualdades.

Assim as divergências são:

$$\begin{aligned} h_{\mathbb{T}^{\text{"def"}}}\left(X_{\mathbb{P}}\right) - h_{\mathbb{T}^{\text{"def"}}}\left(X_{\mathbb{T}^{\text{"def"}}}\right) &= 1,32 - 0,92 \\ &= 0,40 \text{ bit/estado do carro C} \end{aligned}$$

e

$$\begin{aligned} h_{\mathbb{T}^{\text{"ok"}}}\left(X_{\mathbb{P}}\right) - h_{\mathbb{T}^{\text{"ok"}}}\left(X_{\mathbb{T}^{\text{"ok"}}}\right) &= 0,58 - 0,29 \\ &= 0,29 \text{ bit/estado do carro C} \end{aligned}$$

Essa diferença é chamada de **entropia relativa** ou **divergência de Kullback-Leibler**. A divergência também pode ser formalizada matematicamente, tanto para o resultado "defeituoso" quanto para o resultado "ok", como sendo a esperança de $X_{\mathbb{T}^{\text{"}\bullet\text{"}}/\mathbb{P}} = \log_2 \frac{\mathbb{T}^{\text{"}\bullet\text{"}}}{\mathbb{P}}$ na distribuição de probabilidade $\mathbb{T}^{\text{"}\bullet\text{"}}$.

$$
\begin{aligned}
d\left(X_{\mathbb{T}^{\text{"def"}}} \,||\, X_{\mathbb{P}}\right) &= E_{\mathbb{T}^{\text{"def"}}}\left(X_{\mathbb{T}^{\text{"def"}}/\mathbb{P}}\right) \\
&= E_{\mathbb{T}^{\text{"def"}}}\left(X_{\mathbb{P}}\right) - E_{\mathbb{T}^{\text{"def"}}}\left(X_{\mathbb{T}^{\text{"def"}}}\right) \\
&= h_{\mathbb{T}^{\text{"def"}}}\left(X_{\mathbb{P}}\right) - h_{\mathbb{T}^{\text{"def"}}}\left(X_{\mathbb{T}^{\text{"def"}}}\right) \\
&= \mathrm{P}(\mathrm{def} \mid \text{"def"}) \log_2 \left[\frac{\mathrm{P}(\mathrm{def} \mid \text{"def"})}{\mathrm{P}(\mathrm{def})}\right] + \mathrm{P}(\mathrm{ok} \mid \text{"def"}) \log_2 \left[\frac{\mathrm{P}(\mathrm{ok} \mid \text{"def"})}{\mathrm{P}(\mathrm{ok})}\right] \\
&= 0{,}40 \text{ bit/estado do carro C}
\end{aligned}
$$

$$
\begin{aligned}
d\left(X_{\mathbb{T}^{\text{"ok"}}} \,||\, X_{\mathbb{P}}\right) &= E_{\mathbb{T}^{\text{"ok"}}}\left(X_{\mathbb{T}^{\text{"ok"}}/\mathbb{P}}\right) \\
&= E_{\mathbb{T}^{\text{"ok"}}}\left(X_{\mathbb{P}}\right) - E_{\mathbb{T}^{\text{"ok"}}}\left(X_{\mathbb{T}^{\text{"ok"}}}\right) \\
&= h_{\mathbb{T}^{\text{"ok"}}}\left(X_{\mathbb{P}}\right) - h_{\mathbb{T}^{\text{"ok"}}}\left(X_{\mathbb{T}^{\text{"ok"}}}\right) \\
&= 0{,}29 \text{ bit/estado do carro C}
\end{aligned}
$$

Observe-se que a divergência para o resultado "defeituoso" é maior que a do "ok".

Isto significa que o resultado do teste defeituoso é mais informativo (reduz mais a incerteza) que o do "ok" aproximadamente (0,40/0,29)=1,38 vezes.

A divergência é a diferença média de incerteza entre os estados informacionais 2 e 3.

No presente exemplo, o resultado "defeituoso" proporciona uma economia de aproximadamente 0,4 bit/estado do carro C em média para determinar a verdadeira condição do carro C. Já o resultado "ok" economiza 0,29 bit/estado do carro C em média.

Observe-se ainda que as entropias do estado informacional 3 não são comparáveis com a entropia do estado de conhecimento 1. Um simples cálculo mostra esta condição. É de se esperar que o estado de conhecimento 3 contenha mais informação (menos incerteza) sobre o estado do carro C. Portanto, a entropia do estado 1 deveria ser maior que a do 3. No entanto, sendo

$h_{\mathbb{P}}\left(X_{\mathbb{P}}\right) \approx 0{,}88$ bit/estado do carro C e $h_{\mathbb{T}^{\text{"def"}}}\left(X_{\mathbb{T}^{\text{"def"}}}\right) = 0{,}92$ bit/estado do carro C, a diferença é negativa! Este resultado configura uma situação absurda.

Portanto, os estados informacionais 1 e 3 não são comparáveis:

$$\begin{aligned} h_{\mathbb{P}}\left(X_{\mathbb{P}}\right) - h_{\mathbb{T}^{\text{"def"}}}\left(X_{\mathbb{T}^{\text{"def"}}}\right) &= 0{,}88 - 0{,}92 \\ &= -0{,}04 \text{ bit/estado do carro C} \end{aligned}$$

Este resultado ocorre porque apenas os estados de conhecimento 2 e 3 são comparáveis. Note-se que, nestes estados de conhecimento (2 e 3) as entropias (esperanças) estão definidas na mesma medida de probabilidade $\mathbb{T}^{\text{"def"}}$ e $\mathbb{T}^{\text{"ok"}}$.

Pode-se agora definir o ganho informacional total da previsão γ como sendo a divergência total da seguinte forma:

$$\begin{aligned} \gamma &= d\left(X_{\mathbb{T}^{\text{"def"}}} \,||\, X_{\mathbb{P}}\right) + d\left(X_{\mathbb{T}^{\text{"ok"}}} \,||\, X_{\mathbb{P}}\right) \\ &= 0{,}40 + 0{,}29 \\ &= 0{,}69 \text{ bit/estado do carro C} \end{aligned}$$

Duas outras grandezas úteis são os ganhos relativos de informação ou divergências relativas $\gamma^{\text{"def"}}$ e $\gamma^{\text{"ok"}}$, que serão definidas como:

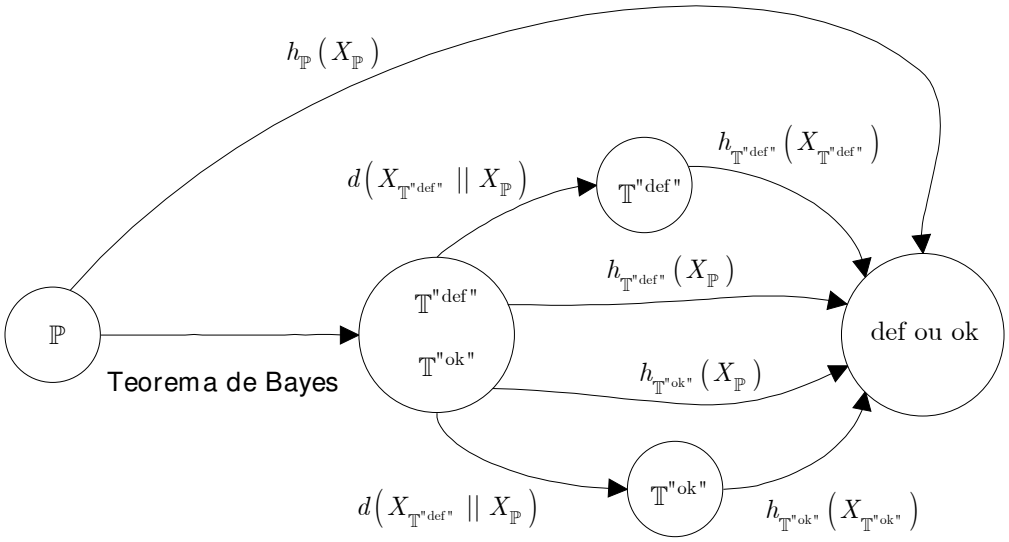
$$\begin{aligned} \gamma^{\text{"def"}} &= \frac{d\left(X_{\mathbb{T}^{\text{"def"}}} \,||\, X_{\mathbb{P}}\right)}{d\left(X_{\mathbb{T}^{\text{"def"}}} \,||\, X_{\mathbb{P}}\right) + d\left(X_{\mathbb{T}^{\text{"ok"}}} \,||\, X_{\mathbb{P}}\right)} \\ &= \frac{d\left(X_{\mathbb{T}^{\text{"def"}}} \,||\, X_{\mathbb{P}}\right)}{\gamma} \\ &= \frac{0{,}40}{0{,}69} \\ &= 0{,}58 \\ \gamma^{\text{"ok"}} &= 1 - \gamma^{\text{"def"}} \\ &= 1 - 0{,}58 \\ &= 0{,}42 \end{aligned}$$

Como mostram as equações anteriores, estas grandezas medem as proporções dos ganhos informacionais da previsão em relação ao ganho total de informação absoluto do teste.

O ganho informacional relativo da previsão do carro C ser "defeituoso" é maior que ganho informacional relativo se a previsão é "ok".

A entropia mede a quantidade média de bit/estado do carro C necessária para realizar o diagnóstico correto. Isto é, ela mede a quantidade média de surpresa (informacional) entre um estado informacional com incerteza (1, 2 e 3) e o estado de conhecimento certo (estado informacional 4). Portanto, é necessário ter em mente que o foco do cálculo da entropia é o evento selecionado pela natureza - que no presente exemplo é ser ou não defeituoso - tendo como ponto de partida um determinado estado de conhecimento com incerteza. A divergência é a economia média em bit/estado do carro C para descobrir a verdadeira condição do carro C de ED's. Esquematizando:

Figura 7: Relações entre as entropias e divergências dos estados informacionais



Considerando agora duas autovariáveis aleatórias discretas binárias $X_{\mathbb{P}}$ e $X_{\mathbb{T}}$, então a divergência é:

$$d(X_{\mathbb{T}} \,||\, X_{\mathbb{P}}) = t \log_2 \frac{t}{p} + (1-t) \log_2 \frac{1-t}{1-p}\,.$$

A divergência é positiva e é mínima (zero) quando $p = t$.

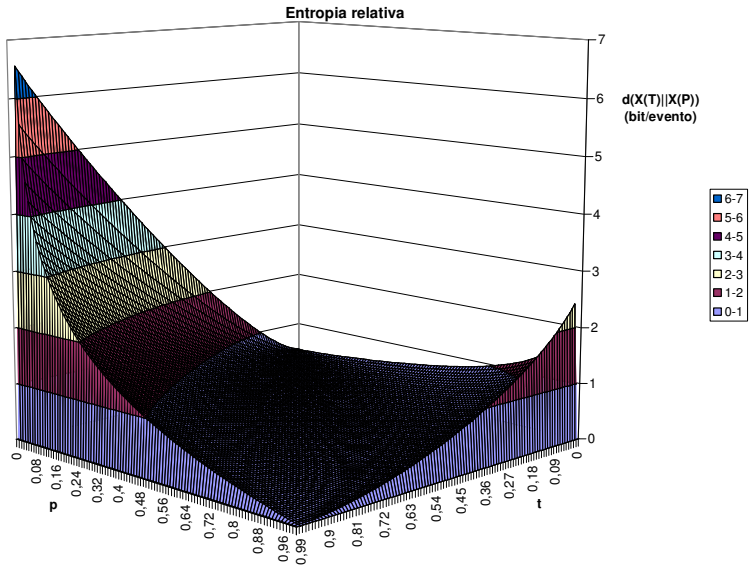
Note-se que, se $p = \frac{1}{2}$, então:

$$d(X_{\mathbb{T}} \,||\, X_{\mathbb{P}}) = 1 - h_{\mathbb{T}}(X_{\mathbb{T}})$$
$$d(X_{\mathbb{T}} \,||\, X_{\mathbb{P}}) + h_{\mathbb{T}}(X_{\mathbb{T}}) = 1$$

Quando $X_{\mathbb{P}}$ possui distribuição uniforme discreta, a medida de probabilidade $\mathbb{T}$, inferida pela maximização da entropia sujeita a restrições, equivale à minimização da divergência.

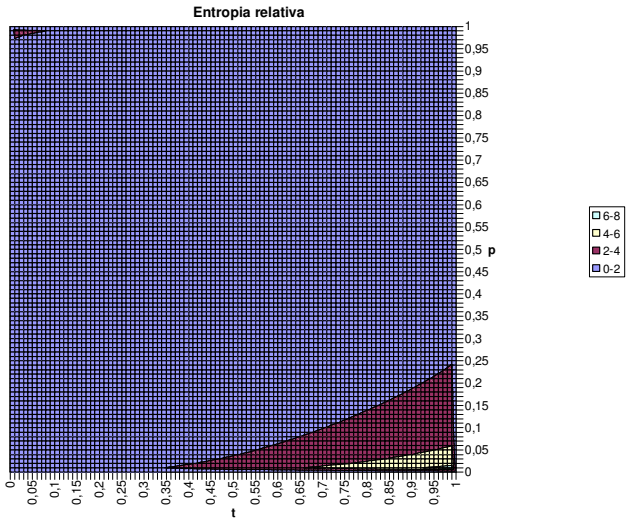
O gráfico de superfície da divergência é:

Figura 8: Gráfico da superfície da divergência



O gráfico das curvas de nível da superfície da divergência é:

Figura 9: Gráfico de curvas de nível da divergência



Devido à assimetria apresentada nos gráficos anteriores, a divergência não é uma distância euclidiana, isto é, $d(X_{\mathbb{T}} \,||\, X_{\mathbb{P}})$ não é necessariamente igual a $d(X_{\mathbb{P}} \,||\, X_{\mathbb{T}})$.

Kullback (1997, p. 6-7) denomina $d(X_{\mathbb{T}} \,||\, X_{\mathbb{P}})$ e $d(X_{\mathbb{P}} \,||\, X_{\mathbb{T}})$ de **divergências direcionadas**. A soma delas origina o que ele denomina propriamente de **divergência**. A divergência resultante desta soma possui todas as propriedades de distância (ou métrica) como definida pela topologia, com exceção da desigualdade triangular. Desta forma, ela não pode ser chamada de distância.

Em Wolf (1996, node96.html), há dois importantes teoremas decorrentes do conceito de divergência em conjunção com o teorema de Bayes:

- Teorema 1 (a informação aumenta em média): embora um conjunto particular de novos dados possa diminuir a informação sobre os parâmetros, em média a informação sobre os parâmetros aumenta.
- Teorema 2 (o aumento médio de informação é a divergência): o aumento médio da informação sobre os parâmetros é a média da divergência entre a distribuição dos parâmetros condicionada aos dados após e antes da nova amostra ter sido observada.

2.3.4.4 Eficiência informacional relativa da previsão

Outra medida importante que resta determinar é a de eficiência informacional da previsão. Para isto é necessário determinar o limite superior da divergência total para este problema. Dessa forma são calculadas as divergências, considerando um previsor perfeito (clarividente).

O clarividente possui as seguintes características: P("ok"|ok)=1 e P("def"|def)=1. Isto implica que P(ok)=P("ok"), P(def)=P("def"), P(ok|"ok")=1 e P(def|"def")=1.

$$
\begin{aligned}
\gamma_{\max}^{\text{"def"}} &= -\log_2 \mathrm{P}(\text{def}) \\
&= -\log_2 0,30 \\
&= 1,74 \text{ bit/estado do carro C} \\
\gamma_{\max}^{\text{"ok"}} &= -\log_2 \mathrm{P}(\text{ok}) \\
&= -\log_2 0,70 \\
&= 0,51 \text{ bit/estado do carro C} \\
\gamma_{\max} &= \gamma_{\max}^{\text{"def"}} + \gamma_{\max}^{\text{"ok"}} \\
&= -\log_2 \mathrm{P}(\text{def})\mathrm{P}(\text{ok}) \\
&= 2,25 \text{ bit/estado do carro C}
\end{aligned}
$$

Como o limite inferior da divergência total é zero, a razão entre a divergência observada e a total fornece um índice de eficiência informacional relativa da previsão:

$$\begin{aligned}\eta &= \frac{\gamma}{\gamma_{\max}} \\ &= \frac{0,69}{2,25} \\ &= 0,31\end{aligned}$$

Portanto, este previsor (mecânico) - para este decisor responsável (comprador do carro C) - atinge 31% da divergência total em relação a um mecânico clarividente.

O valor mínimo da divergência total máxima é 2 bit/estado do carro C e ocorre quando P(ok)=1/2=P(def).

A divergência total atinge o valor zero quando a sensibilidade e a especificidade são iguais a 1/2. Desta forma, a eficiência relativa do previsor é nula, qualquer que seja a probabilidade prévia.

Esta medida de eficiência relativa pode ser utilizada para comparar a eficiência de previsores.

Quadro 3: Eficiência informacional relativa da previsão

	Característica do carro C e do previsor												
	a	b	c	d	e	f	g	h	i	j	k	l	m
P(+)	0,50	0,30	0,30	0,40	0,40	0,30	0,30	0,40	0,40	0,001	0,001	0,001	0,999
Sensibilidade = P("+"|+)	0,50	0,90	0,80	0,80	0,90	0,80	0,90	0,80	0,90	0,999	0,99	0,9	0,9
Especificidade = P("-"|-)	0,50	0,80	0,90	0,90	0,80	0,80	0,90	0,80	0,90	0,999	0,998	0,9	0,9
Divergência("+") (bit/estado do carro C)	0,00	0,39	0,69	0,60	0,36	0,34	0,75	0,32	0,65	3,98	2,39	0,02	0,00
Divergência("-") (bit/estado do carro C)	0,00	0,29	0,19	0,26	0,39	0,17	0,30	0,23	0,42	0,00	0,00	0,00	0,02
Divergência total (bit/estado do carro C)	0,00	0,68	0,88	0,86	0,76	0,51	1,06	0,55	1,06	3,99	2,39	0,02	0,02
Divergência total máxima (bit/estado do carro C)	2,00	2,25	2,25	2,06	2,06	2,25	2,25	2,06	2,06	9,97	9,97	9,97	9,97
Eficiência informacional relativa da previsão (%)	0	30	39	42	37	23	47	27	52	40	24	0,2	0,0018

+ significa ter defeito de transmissão no carro C ou sucesso.

Sendo p a probabilidade de sucesso, α a sensibilidade e β a especificidade, então as divergências são:

$$d(X_{\mathbb{T}^{"+"}} \,||\, X_{\mathbb{P}}) = \frac{p\alpha}{p\alpha + (1-p)(1-\beta)} \log_2 \frac{\alpha}{p\alpha + (1-p)(1-\beta)} + \frac{(1-p)(1-\beta)}{p\alpha + (1-p)(1-\beta)} \log_2 \frac{1-\beta}{p\alpha + (1-p)(1-\beta)}$$

$$d(X_{\mathbb{T}^{"-"}} \,||\, X_{\mathbb{P}}) = \frac{p(1-\alpha)}{p(1-\alpha) + (1-p)\beta} \log_2 \frac{1-\alpha}{p(1-\alpha) + (1-p)\beta} + \frac{(1-p)\beta}{p(1-\alpha) + (1-p)\beta} \log_2 \frac{\beta}{p(1-\alpha) + (1-p)\beta}$$

Fazendo $\alpha = \beta$ tem-se:

$$d(X_{\mathbb{T}^{"+"}} \,||\, X_{\mathbb{P}}) = \frac{p\alpha}{p\alpha + (1-p)(1-\alpha)} \log_2 \frac{\alpha}{p\alpha + (1-p)(1-\alpha)} + \frac{(1-p)(1-\alpha)}{p\alpha + (1-p)(1-\alpha)} \log_2 \frac{1-\alpha}{p\alpha + (1-p)(1-\alpha)}$$

$$d(X_{\mathbb{T}^{"-"}} \,||\, X_{\mathbb{P}}) = \frac{p(1-\alpha)}{p(1-\alpha) + (1-p)\alpha} \log_2 \frac{1-\alpha}{p(1-\alpha) + (1-p)\alpha} + \frac{(1-p)\alpha}{p(1-\alpha) + (1-p)\alpha} \log_2 \frac{\alpha}{p(1-\alpha) + (1-p)\alpha}$$

Figura 10: Gráfico de superfície da divergência("+") quando sensibilidade=especificidade

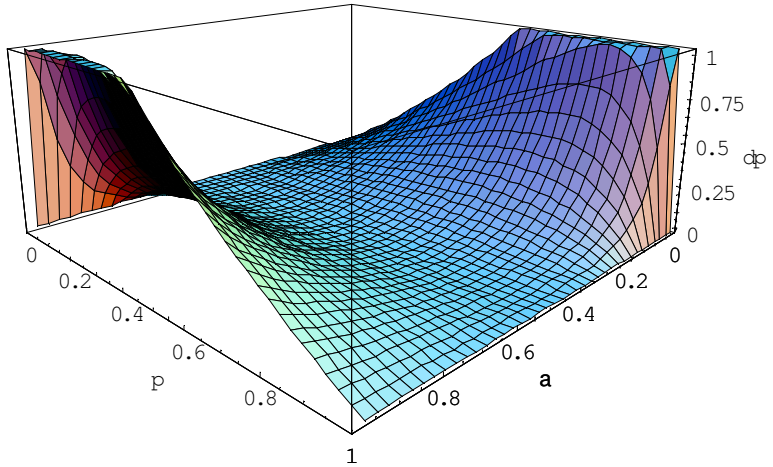
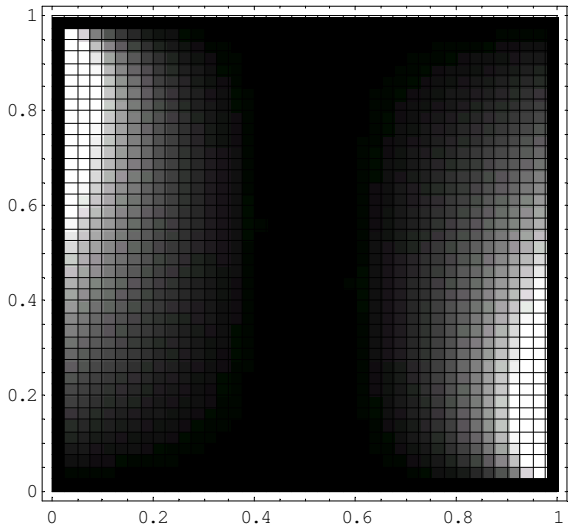


Figura 11: Gráfico de densidade da divergência("+") quando sensibilidade=especificidade



As divergências parciais máximas que ocorrem na situação de clarividência, isto é, com especificidade e sensibilidade iguais a um $(\alpha = \beta = 1)$ e definindo que $0 \log_2 0 = 0$, são:

$$d_{\max}(X_{\mathbb{T}^{"+"}} \,||\, X_{\mathbb{P}}) = -\log_2 p$$
$$d_{\max}(X_{\mathbb{T}^{"-"}} \,||\, X_{\mathbb{P}}) = -\log_2 (1-p)$$

A divergência total máxima é:

$$\begin{aligned}\gamma_{\max} &= d_{\max}(X_{\mathbb{T}^{"+"}} \mid\mid X_{\mathbb{P}}) + d_{\max}(X_{\mathbb{T}^{"-"}} \mid\mid X_{\mathbb{P}}) \\ &= -\log_2 p - \log_2(1-p) \\ &= -\log_2 p(1-p)\end{aligned}$$

Figura 12: Gráfico da divergência total máxima quando sensibilidade=especificidade=1

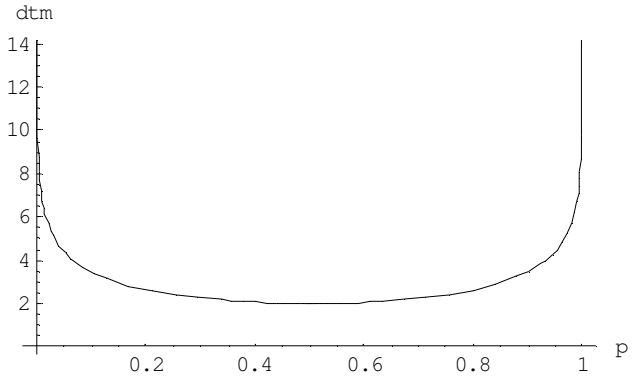


Figura 13: Gráfico de superfície da divergência total quando sensibilidade=especificidade

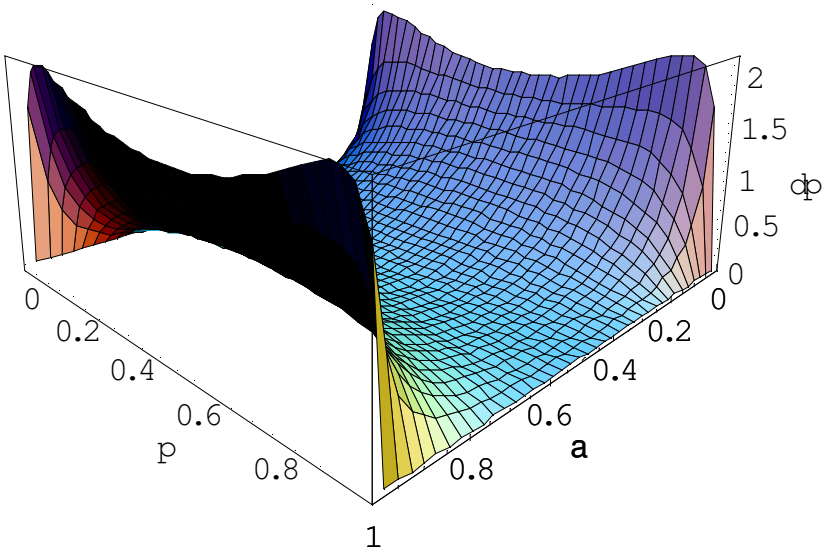
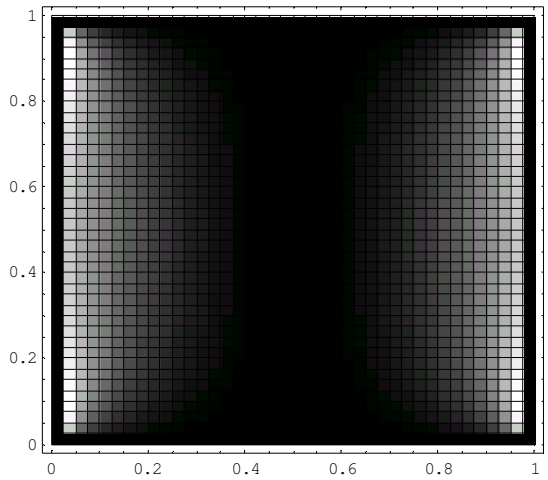


Figura 14: Gráfico de densidade da divergência total quando sensibilidade=especificidade



Pode-se ainda definir a eficiência informacional relativa da previsão da seguinte forma:

$$\eta = \frac{\gamma}{\gamma_{\max}} .$$

Quão útil é um previsor específico ou uma determinada previsão? A resposta é: depende das probabilidades prévias.

Note-se que numa situação na qual a probabilidade prévia de sucesso é 0,5, qualquer previsor com sensibilidade e especificidade diferentes de 0,5 proporciona ganho de informação. Esta situação com probabilidade de sucesso igual a 0,5 é a de máxima entropia (estado de conhecimento 1 de máxima ignorância ou incerteza), isto é, há muita quantidade de incerteza para ser eliminada por um previsor.

Já quando a probabilidade de sucesso encontra-se perto dos extremos (0 ou 1), o previsor necessita ser praticamente um clarividente para proporcionar ganho substancial de informação. Esta situação é a de baixa entropia, isto é, há pouca quantidade de incerteza para ser eliminada por um previsor.

Quadro 4: Eficiência informacional relativa da previsão - II

P(+)	0,50	0,50	0,50	0,001	0,001	0,001	0,001	0,001	0,001
Sensibilidade = P("+"|+)	0,50	0,60	0,40	0,80	0,90	0,99	0,999	0,9999	0,99999
Especificidade = P("-"|-)	0,50	0,50	0,50	0,80	0,90	0,99	0,999	0,9999	0,99999
Divergência("+") (bit/estado do carro C)	0,00	0,01	0,01	0,00	0,02	0,46	3,98	8,62	9,79
Divergência("-") (bit/estado do carro C)	0,00	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00
Divergência total (bit/estado do carro C)	0,00	0,01	0,01	0,00	0,02	0,46	3,99	8,62	9,79
Divergência total máxima (bit/estado do carro C)	2,00	2,00	2,00	9,97	9,97	9,97	9,97	9,97	9,97
Eficiência informacional relativa da previsão (%)	0,0	0,7	0,7	0,0	0,2	4,7	40,0	86,5	98,2

+ significa ter defeito de transmissão no carro C ou sucesso.

2.3.4.5 Medida entrópica de ganho de informação aplicada em Economia

Theil (1967, cap. 1) foi o precursor da idéia de mensurar o ganho informacional da previsão, utilizando a Teoria da Informação em Economia. No entanto, ele a definiu de maneira imprecisa. Definiu o ganho de informação da previsão de aumento de preço como uma diferença entre informacionais de uma maneira inadequada.

O exemplo a seguir refere-se ao novo enfoque de ganho de informação proporcionado por uma nova evidência, pela Teoria da Informação, com base no conceito de divergência anteriormente abordado. Agora será ele aplicado a uma variável econômica de variação de preço de um bem.

Para isso considera-se um exemplo com dados reais, exposto em Theil (1967, cap. 1). Este autor foi um dos pioneiros da aplicação da TI a problemas de previsões do sentido de variação de variáveis econômicas. Assim, de acordo com esse autor, suponhamos que sejam obtidas as previsões e correspondentes realizações de variação do preço de um bem num mercado, pela aplicação de questionários enviados por um instituto de pesquisas econômicas a um certo número de

empresas. O mercado é o interessado no resultado da pesquisa. O instituto de pesquisas econômicas prevê a cada mês ou trimestre a direção da variação de preço do bem. Aos participantes é inquirido predizer (manifestar a expectativa sobre) a variação do preço de um bem no mês seguinte (resultado do teste). Os tipos de variação (eventos deste espaço amostral) são: “1” = "aumento do preço do bem no próximo mês", “2” = "permanência do preço do bem no próximo mês" e “3” = "diminuição do preço do bem no próximo mês".

Aos participantes também é inquirido relatar a variação do preço de um bem percebida como ocorrida no último mês. Os tipos de variação de preço são: 1 = aumento do preço no último mês, 2 = permanência do preço no último mês, e 3 = diminuição do preço no último mês.

Conforme Theil (1967, p. 11) os dados são reais.

Representa-se $\hat{p}_{ij}$ como a freqüência relativa (estimativa da probabilidade) de todos os casos nos quais algum participante prediz uma variação $i = \text{"1"},\text{"2"},\text{"3"}$ e reporta uma variação $j = 1,2,3$ ocorrida. Assim por exemplo, $\hat{p}_{12}$ é a freqüência relativa de todos os casos nos quais o aumento (“1”) foi previsto e no qual foi relatado que o preço se manteve (2) no último mês.

$\hat{p}_{ij}$ é a estimativa da probabilidade conjunta de ocorrer i e j, isto é:

$$\hat{p}_{ij} = \widehat{\Pr}(i;j).$$

Com estes dados observados, pode-se formar a seguinte matriz de estimativas das probabilidades conjuntas:

		Realização			Probabilidade marginal
		1	2	3	
	“1”	$\hat{p}_{11}$	$\hat{p}_{12}$	$\hat{p}_{13}$	$\hat{p}_{1.}$
Previsão	“2”	$\hat{p}_{21}$	$\hat{p}_{22}$	$\hat{p}_{23}$	$\hat{p}_{2.}$
	“3”	$\hat{p}_{31}$	$\hat{p}_{32}$	$\hat{p}_{33}$	$\hat{p}_{3.}$

Probabi-lidade marginal		$\widehat{p}_{\cdot 1}$	$\widehat{p}_{\cdot 2}$	$\widehat{p}_{\cdot 3}$	1

As freqüências relativas marginais são aqui:

$$\widehat{p}_{i\cdot} = \sum_{j=1}^{3} \widehat{p}_{ij}, \ i = \text{"1","2","3"}$$
$$\widehat{p}_{\cdot j} = \sum_{i=1}^{3} \widehat{p}_{ij}, \ j = 1,2,3$$

Elas representam a freqüência relativa total das previsões e das realizações, respectivamente.

Considere-se que os valores das estimativas das probabilidades conjuntas são os da tabela a seguir.

		Realização			
		1	2	3	Probabilidade marginal
	"1"	0,111	0,043	0,003	0,157
Previsão	"2"	0,039	0,523	0,089	0,651
	"3"	0,004	0,069	0,119	0,192
	Marginal ($\mathbb{P}$)	0,154	0,635	0,211	1,000

Pode-se verificar que as probabilidades prévias são dadas pelas probabilidades marginais de realização. Isto é, a probabilidade prévia de o preço do bem no mercado permanecer o mesmo, de um mês para o outro, é 0,635. A probabilidade de a expectativa de mercado ser de permanência do preço é 0,651. A soma da diagonal principal é 753. Este valor é a probabilidade de permanência da direção da variação do preço no mercado.

A pesquisa permite estimar tanto a medida de probabilidade prévia quanto as posteriores.

As probabilidades posteriores estão calculadas na tabela a seguir da seguinte maneira:

$$\begin{aligned}\widehat{\Pr}(j \mid i) &= \frac{\widehat{\Pr}(j;i)}{\widehat{\Pr}(i)} \\ &= \frac{\hat{p}_{ij}}{\hat{p}_{i\cdot}}\end{aligned}$$

		Realização				
		1	2	3	Probabilida de marginal	Medida de probabilidade posterior $\mathbb{T}^i$
	"1"	0,707	0,263	0,019	1	$\mathbb{T}^{"1"}$
Previsão	"2"	0,060	0,803	0,137	1	$\mathbb{T}^{"2"}$
	"3"	0,021	0,359	0,620	1	$\mathbb{T}^{"3"}$

A probabilidade de o preço do bem aumentar no próximo mês, dado que o mercado tem a expectativa de que o preço irá aumentar no próximo mês, é 0,707. A probabilidade de o preço do bem não variar no próximo mês, dado que o mercado tem a expectativa de que o preço irá aumentar no próximo mês, é 0,263.

Theil (1967, cap. 1) havia definido o ganho de informação da previsão de variação de preço como:

$$\begin{aligned} \gamma^i_{\text{Theil}} &= s\left(\hat{p}_{\cdot j}\right) - s\left(\frac{\hat{p}_{ij}}{\hat{p}_{i\cdot}}\right) \\ &= \log_2\left(\frac{\hat{p}_{ij}}{\hat{p}_{i\cdot}\hat{p}_{\cdot j}}\right) \text{ bit} \\ \text{para } i &= \texttt{"1"},\texttt{"2"},\texttt{"3"} \\ \text{e } i &= j \end{aligned}$$

Portanto, a medida de ganho de informação de Theil possui várias deficiências:

- Não exige que o instituto tenha que anunciar ao mercado sua previsão da direção da variação do preço. Sendo assim, não fica caracterizado o estado informacional 3.
- Não aproveita toda a informação das medidas de probabilidade disponíveis, isto, o conhecimento disponível.

- É um informacional, portanto, sua unidade é bit. Assim, a definição de Theil nem chegou a utilizar o conceito de entropia (bit/evento).
- Não respeita o princípio de utilização máxima do conhecimento disponível.
- Não respeita o princípio de comparação entre estados de conhecimento.

Para sanar todos os problemas anteriores, propõe-se aqui a divergência como medida de ganho informacional da previsão. Pode-se calcular os ganhos informacionais relativos da previsão do instituto, isto é, as divergências para cada expectativa de mercado sobre a variação do preço do bem.

A fórmula para o estimador das divergências é:

$$\begin{aligned}\hat{d}\left(X_{\mathbb{T}^i} \,||\, X_{\mathbb{P}}\right) &= \widehat{E}_{\mathbb{T}^i}\left(X_{\mathbb{T}^i/\mathbb{P}}\right)\\ &= \widehat{E}_{\mathbb{T}^i}\left(X_{\mathbb{P}}\right) - \widehat{E}_{\mathbb{T}^i}\left(X_{\mathbb{T}^i}\right)\\ &= \hat{h}_{\mathbb{T}^i}\left(X_{\mathbb{P}}\right) - \hat{h}_{\mathbb{T}^i}\left(X_{\mathbb{T}^i}\right)\\ &= \sum_{j=1}^{3} \frac{\hat{p}_{ij}}{\hat{p}_{i\cdot}} \log_2\left(\frac{\hat{p}_{ij}}{\hat{p}_{i\cdot}\hat{p}_{\cdot j}}\right)\end{aligned}$$

para $i =$ "1","2","3"

Previsão	Divergência (bit/tipo de variação)
“1”	1,154
“2”	0,105
“3”	0,608
Total	1,867

A fórmula para o estimador das divergências relativas é:

$$\hat{\gamma}^i = \frac{\hat{d}(X_{\mathbb{T}^i} \,||\, X_{\mathbb{P}})}{\sum_{i=1}^{3} \hat{d}(X_{\mathbb{T}^i} \,||\, X_{\mathbb{P}})}$$

para $i =$ "1","2","3"

Previsão	Divergência relativa
“1”	0,618
“2”	0,056
“3”	0,326
Total	1,000

Convém lembrar que a divergência é a redução da quantidade média de bit por tipo de variação de preço do bem necessária, para descobrir a variação de preço do bem (acertar o prognóstico econômico) que ocorrerá no próximo mês, proporcionada pela informação da expectativa do tipo de variação pelo mercado.

O resultado da pesquisa de opinião (expectativa de mercado anunciada pelo instituto para o próximo mês) reduz em média aproximadamente 1,9 bit/tipo de variação de preço do bem para o mercado identificar a próxima variação de preço.

O resultado da pesquisa, anunciado pelo instituto, de que o mercado tem a expectativa de aumento do preço para o próximo mês (“1”), proporciona uma redução média de aproximadamente 1,2 bit/tipo de variação, que corresponde a 62% do ganho de informação total.

A divergência relativa da previsão de permanência do preço pelo mercado indica para o mercado que este tipo de expectativa de variação é muito pouco informativa em relação aos outros dois resultados apontados pela pesquisa.

As divergências relativas do exemplo indicam que a informação sobre permanência do preço no próximo mês reduz relativamente muito pouco a incerteza, isto é, o ganho relativo de informação é de aproximadamente 5% sobre a real variação do preço do próximo mês. No entanto, a informação sobre a diminuição do preço do bem no próximo mês representa uma considerável redução relativa da incerteza, isto é, o ganho relativo de informação é de aproximadamente 33%, sobre a real variação do preço do bem no próximo mês. Já a expectativa de mercado de aumento do preço do bem no próximo mês representa uma expressiva redução relativa da incerteza, isto é, o ganho relativo

de informação é de aproximadamente 62%, sobre a determinação da real variação do preço do bem no próximo mês.

Desta forma, pode-se ordenar informacionalmente estes resultados em termos de sua divergência relativa. Quanto mais alta a divergência relativa de uma expectativa do mercado sobre a variação do preço, mais relevante é a expectativa para determinar a real variação do preço do bem no próximo mês.

Esta medida é útil também para avaliar o estado informacional do mercado com relação à próxima variação de preço de um bem. Isto é, pode-se avaliar se o mercado é informacionalmente eficiente ou se está em equilíbrio entrópico [Gulko (1998)].

Suponha-se que as divergências relativas fossem:

Previsão	Divergência relativa (bit/tipo de variação)
"1"	0,5
"2"	0,0
"3"	0,5
Total	1,0

As divergências relativas hipotéticas indicam que o mercado considera irrelevante relativamente a informação sobre a expectativa do mercado de permanência do preço no próximo mês, como redutora da incerteza sobre a real variação do preço do próximo mês. No entanto, as informações sobre as expectativas do mercado sobre a variação (aumento ou diminuição) do preço do bem no próximo mês são igualmente relevantes relativamente, como redutoras da incerteza sobre a real variação do preço do bem no próximo mês. Este resultado corrobora o fato de que, na definição de mercado informacionalmente eficiente, apenas a expectativa de aumento e de diminuição do preço no próximo instante são consideradas e que

estas expectativas estão em equilíbrio informacional a cada instante (Gulko (1998)).

2.3.5 Entropia diferencial

Quando o número de células de um sistema fechado tende para o infinito, isto é, $m \to +\infty$, tem-se o surgimento da entropia diferencial de Boltzmann. A entropia diferencial é a versão contínua da entropia discreta de Shannon (Golan *et al.* (1996, p. 40)).

A entropia diferencial é a entropia de uma variável aleatória contínua (Cover & Thomas (1991, p. 224)).

A princípio seria razoável supor que a entropia de uma variável aleatória contínua fosse infinita positivamente [Haykin (1999, p. 485)]. No entanto será definida uma medida derivada da entropia chamada **entropia diferencial**.

Seja X uma variável aleatória contínua com distribuição de probabilidade acumulada:

$$\begin{aligned} F(x) &= \Pr(X \leq x) \\ &= \int_{-\infty}^{x} f(z)dz \end{aligned}$$

$f(x)$ é a função de densidade de probabilidade da variável aleatória contínua X e está relacionada com a função de distribuição da seguinte forma:

$$\begin{aligned} \frac{d}{dx}[F(x)] &= \frac{dF(x)}{dx} \\ &= f(x) \\ dF(x) &= f(x)dx \end{aligned}$$

$\int_{-\infty}^{+\infty} f(x)dx = 1$ é a condição análoga à soma das probabilidades igual a 1.

O conjunto de valores do domínio para os quais $f(x) > 0$ é chamado de **conjunto de apoio** da variável aleatória X e é representado por ξ.

Uma definição mais geral da entropia é:

$$
\begin{aligned}
h_{\mathbb{P}}(X_{\mathbb{P}}) &= -\int_{\xi} \log_2 f(x) dF(x) \\
&= -\int_{\xi} \log_2 [f(x)] f(x) dx \\
&= -\int_{\xi} [f(x) dx] \log_2 [f(x)] \\
&= -\int_{\xi} f(x) \log_2 [f(x)] dx
\end{aligned}
$$

Note-se que o termo $f(x)dx$ (função densidade de probabilidade multiplicada pelo diferencial de x) é o correspondente da $\Pr(x)$ (massa de probabilidade de x).

2.3.5.1 Entropia de uma variável aleatória uniforme contínua

Imagine-se $X \sim \frac{d\mathbb{P}}{dx} = \text{uniforme}(0;a),\ a > 0$. Logo, a função de densidade de probabilidade é $f(x) = \frac{1}{a}$. O conjunto de apoio é o intervalo 0-a, isto é, $\xi = [0;a]$.

A entropia diferencial de X é:

$$
\begin{aligned}
h_{\mathbb{P}}(X_{\mathbb{P}}) &= -\int_{0}^{a} \frac{1}{a} \log_2 \frac{1}{a} dx \\
&= \log_2 a
\end{aligned}
$$

Note-se que se $a < 1 \Rightarrow \log_2 a < 0$.

Portanto, a entropia diferencial pode ser negativa.

O volume do conjunto de apoio é, neste caso unidimensional, o comprimento do conjunto apoio (segmento de reta), isto é, a.

Há, no entanto, uma maneira analítica de calcular o volume do conjunto de apoio:

$$
\begin{aligned}
vol(\xi) &= 2^{h_{\mathbb{P}}(X_{\mathbb{P}})} \\
&= 2^{\log_2 a} \\
&= a
\end{aligned}
$$

Note-se que a descrição exata dos valores do domínio passa a ser uma informação necessária para o cálculo da entropia diferencial. O mesmo não ocorre com a entropia discreta para variável aleatória discreta.

2.3.5.2 *Entropia de uma variável aleatória normal contínua*

Imagine-se que $X \sim \frac{d\mathbb{P}}{dx} = \text{normal}(\mu;\sigma^2)$. Então:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$
$$x \in \mathbb{R}$$

Note-se que o conjunto de apoio tem volume infinito (reta real). No entanto, como será mostrado a seguir, a entropia diferencial é finita. Logo, o volume do conjunto de apoio também é finito.

A entropia diferencial de uma variável aleatória normal contínua é:

$$h_{\mathbb{P}}(X_{\mathbb{P}}) = \frac{1}{2}\log_2(2\pi e\sigma^2) \text{ bit/evento}$$

Portanto, o volume do conjunto de apoio é:

$$\begin{aligned} vol(\xi) &= 2^{h_{\mathbb{P}}(X_{\mathbb{P}})} \\ &= 2^{\frac{1}{2}\log_2(2\pi e\sigma^2)} \\ &= 2^{\log_2(2\pi e\sigma^2)^{\frac{1}{2}}} \\ &= \sqrt{2\pi e\sigma^2} \end{aligned}$$

No caso específico de $X \sim \mathbb{P} = \text{normal}(0;1) = \text{ normal padronizada}$ tem-se:

$$\begin{aligned} h_{\mathbb{P}}(X_{\mathbb{P}}) &\approx 2{,}065 \text{ bit/evento} \\ vol(\xi) &\approx 4{,}13 \end{aligned}$$

O volume do conjunto de apoio significa que o intervalo de –2,065 a 2,065 centrado na média (zero) contém a maior parte da probabilidade [Cover & Thomas (1991, p. 227)]. De fato, este intervalo contém aproximadamente 0,96 de probabilidade.

Como o conjunto de apoio de uma variável aleatória discreta é enumerável, então seu volume é nulo. A fórmula para o cálculo de volume de um conjunto de apoio reforça esta idéia:

$$\begin{aligned} vol(\xi) &= 2^{h_{\mathbb{P}}(X_{\mathbb{P}})} \\ &= 2^{-\infty} \\ &= 0 \end{aligned}$$

Desta forma, a entropia diferencial de uma variável aleatória discreta é negativamente infinita [Cover & Thomas (1991, p. 229)].

2.3.6 Entropia diferencial como limitante da entropia discreta

Lembre-se que a entropia discreta de uma variável aleatória discreta com domínio finito é máxima quando a distribuição discreta é uniforme. O mesmo não ocorre com a entropia diferencial.

A entropia discreta não depende dos descritores da variável aleatória (medidas de posição, variabilidade etc.), pois ela depende apenas das probabilidades, e não dos valores. O mesmo não ocorre com a entropia diferencial que, no caso da normal, depende da variância e da uniforme dos limites do intervalo, portanto, em certa medida, ela depende do conjunto de apoio.

A distribuição normal é a que possui a máxima entropia de todas as distribuições contínuas com a mesma variância [Cover & Thomas (1991, p. 234-6)]. Então, além da informação do conjunto de apoio, é necessária a informação da variância para a determinação da distribuição de máxima entropia.

A entropia de uma variável aleatória discreta X com medida de probabilidade $\mathbb{P}$ possui um limite superior que é dado pela seguinte expressão:

$$h_{\mathbb{P}}(X_{\mathbb{P}}) \leq \frac{1}{2}\log_2\left[2\pi e\left(\mathrm{Var}_{\mathbb{P}}(Y) + \frac{1}{12}\right)\right],$$

sendo Y uma variável aleatória com medida de probabilidade $\mathbb{P}$ no domínio $\xi = [1,2,..]$.

2.3.7 Entropia mútua ou informação mútua

Imagine-se que dois espaços amostrais $\Omega_1 = \{\omega_{11};\omega_{12}\}$ e $\Omega_2 = \{\omega_{21};\omega_{22}\}$ tenham as seguintes medidas de probabilidade associadas $\mathbb{P}_1 = \{\Pr(\omega_{11});\Pr(\omega_{12})\}$ e $\mathbb{P}_2 = \{\Pr(\omega_{21});\Pr(\omega_{22})\}$, respectivamente.

Produto cartesiano dos espaços amostrais $\Omega = \Omega_1 \otimes \Omega_2$	ω_{11}	ω_{12}
ω_{21}	$(\omega_{11};\omega_{21})$	$(\omega_{12};\omega_{21})$
ω_{22}	$(\omega_{11};\omega_{22})$	$(\omega_{12};\omega_{22})$

A medida de probabilidade associada ao espaço amostral resultante do produto cartesiano dos espaços amostrais $\Omega = \Omega_1 \otimes \Omega_2$ é:

$$\mathbb{P} = \{\Pr(\omega_{11};\omega_{21});\Pr(\omega_{12};\omega_{21});\Pr(\omega_{11};\omega_{22});\Pr(\omega_{12};\omega_{22})\}.$$

Evento do espaço amostral Ω	Probabilidade conjunta da medida de probabilidade conjunta $\mathbb{P}$
$(\omega_{11};\omega_{21})$	$\Pr(\omega_{11};\omega_{21})$
$(\omega_{12};\omega_{21})$	$\Pr(\omega_{12};\omega_{21})$
$(\omega_{11};\omega_{22})$	$\Pr(\omega_{11};\omega_{22})$
$(\omega_{12};\omega_{22})$	$\Pr(\omega_{12};\omega_{22})$

As medidas $\mathbb{P}_1$ e $\mathbb{P}_2$ são conhecidas como **medidas de probabilidade marginais** de $\mathbb{P}$. Esta última, por sua vez, é conhecida por **medida de probabilidade conjunta**.

Por exemplo, considerando que a probabilidade marginal é a soma de duas probabilidades conjuntas:

$$\Pr(\omega_{11}) = \Pr(\omega_{11};\omega_{21}) + \Pr(\omega_{11};\omega_{22}).$$

Se os espaços amostrais são independentes, então:

$$\begin{aligned}\mathbb{P} &= \mathbb{P}_1 * \mathbb{P}_2 \\ &= \{\Pr(\omega_{11}) * \Pr(\omega_{21});\Pr(\omega_{12}) * \Pr(\omega_{21});\Pr(\omega_{11}) * \Pr(\omega_{22});\Pr(\omega_{12}) * \Pr(\omega_{22})\}\end{aligned}.$$

No caso mais geral, quando a independência entre os espaços amostrais ocorre, tem-se que as probabilidades conjuntas podem ser escritas como probabilidades condicionais:

$$\begin{aligned}\mathbb{P} &= (\mathbb{P}_1 \mid \mathbb{P}_2) * \mathbb{P}_2 \\ &= \{\Pr(\omega_{11} \mid \omega_{21}) * \Pr(\omega_{21});\Pr(\omega_{12} \mid \omega_{21}) * \Pr(\omega_{21});\Pr(\omega_{11} \mid \omega_{22}) * \Pr(\omega_{22});\Pr(\omega_{12} \mid \omega_{22}) * \Pr(\omega_{22})\} \\ &= (\mathbb{P}_2 \mid \mathbb{P}_1) * \mathbb{P}_1\end{aligned}$$

Como a quantidade média de surpresa da medida de probabilidade conjunta $\mathbb{P}$ pode ser medida? Isto é, qual a entropia desta medida em função das medidas marginais? A resposta é a expressão a seguir:

$$
\begin{aligned}
E_{\mathbb{P}}\left(X_{\mathbb{P}}\right) &= h_{\mathbb{P}}\left(X_{\mathbb{P}}\right) \\
&= h_{(\mathbb{P}_1|\mathbb{P}_2)*\mathbb{P}_2}\left(X_{(\mathbb{P}_1|\mathbb{P}_2)*\mathbb{P}_2}\right) \\
&= h_{\mathbb{P}_1|\mathbb{P}_2}\left(X_{\mathbb{P}_1|\mathbb{P}_2}\right) + h_{\mathbb{P}_2}\left(X_{\mathbb{P}_2}\right)
\end{aligned}
$$

Portanto, a entropia conjunta é a soma da entropia condicional com a entropia marginal.

Como a expressão pode ser interpretada?

A quantidade média de surpresa por evento do espaço amostral conjunto é a soma das surpresas médias por evento do espaço 1 condicionado ao 2 e do espaço 2. Isto é, quando um evento do espaço amostral 2 é conhecido, ele elimina uma parte da surpresa do espaço amostral 1. O que sobra de surpresa esperada do espaço amostral 1 após a ocorrência de um evento do espaço amostral 2 é somado à surpresa esperada do espaço amostral 2. Note-se que a intersecção entre as duas surpresas esperadas é contada uma vez apenas. Esta intersecção é chamada de **entropia mútua ou informação mútua**. Conforme MacKay (1995, p. 1), a entropia condicional mede a incerteza (surpresa) média por evento do espaço amostra 1, que permanece quando o evento do espaço amostral 2 é conhecido.

Portanto, conforme MacKay (1995, p. 1), dados são úteis, pois eles não aumentam a surpresa ou incerteza em média.

Se os espaços amostrais são independentes, então:

$$
\begin{aligned}
E_{\mathbb{P}}\left(X_{\mathbb{P}}\right) &= h_{\mathbb{P}}\left(X_{\mathbb{P}}\right) \\
&= h_{\mathbb{P}_1*\mathbb{P}_2}\left(X_{\mathbb{P}_1*\mathbb{P}_2}\right) \\
&= h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) + h_{\mathbb{P}_2}\left(X_{\mathbb{P}_2}\right)
\end{aligned}
$$

O resultado anterior indica que a entropia conjunta é a soma das entropias das variáveis aleatórias informacionais. Note-se que a entropia mútua é nula. Isto é, o conhecimento da ocorrência de um evento de um espaço amostral não diminui a surpresa esperada do outro.

A partir deste momento, será criada uma nova notação para entropia conjunta:

$$
h_{\mathbb{P}_1;\mathbb{P}_2}\left(X_{\mathbb{P}_1;\mathbb{P}_2}\right) = h_{\mathbb{P}_1|\mathbb{P}_2}\left(X_{\mathbb{P}_1|\mathbb{P}_2}\right) + h_{\mathbb{P}_2}\left(X_{\mathbb{P}_2}\right).
$$

A expressão anterior é conhecida na Teoria da Informação como **regra da cadeia**.

A entropia conjunta é igual a soma de uma entropia condicional e de uma entropia.

Se $\mathbb{P}_1 = \mathbb{P}_1$, então significa que a entropia condicional é nula, pois o conhecimento do evento do espaço amostral elimina completamente a incerteza sobre o próprio espaço amostral. Desta forma, a entropia conjunta é a entropia. Note-se que a entropia mútua é igual à entropia também. Matematicamente:

$$\begin{aligned} h_{\mathbb{P}_1;\mathbb{P}_1}\left(X_{\mathbb{P}_1;\mathbb{P}_1}\right) &= h_{\mathbb{P}_1|\mathbb{P}_1}\left(X_{\mathbb{P}_1|\mathbb{P}_1}\right) + h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) \\ &= 0 + h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) \\ &= h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) \end{aligned}$$

É possível definir agora definir mais formalmente a informação mútua da variável aleatória informacional conjunta $X_{\mathbb{P}}$:

$$\begin{aligned} h_{\mathbb{P}_1,\mathbb{P}_2}\left(X_{\mathbb{P}_1,\mathbb{P}_2}\right) &= h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) - h_{\mathbb{P}_1|\mathbb{P}_2}\left(X_{\mathbb{P}_1|\mathbb{P}_2}\right) \\ &= h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) + h_{\mathbb{P}_2}\left(X_{\mathbb{P}_2}\right) - h_{\mathbb{P}_1;\mathbb{P}_2}\left(X_{\mathbb{P}_1;\mathbb{P}_2}\right) \end{aligned}$$

ou, equivalentemente:

$$h_{\mathbb{P}_2,\mathbb{P}_1}\left(X_{\mathbb{P}_2,\mathbb{P}_1}\right) = h_{\mathbb{P}_2}\left(X_{\mathbb{P}_2}\right) - h_{\mathbb{P}_2|\mathbb{P}_1}\left(X_{\mathbb{P}_2|\mathbb{P}_1}\right).$$

$h_{\mathbb{P}_1,\mathbb{P}_2}\left(X_{\mathbb{P}_1,\mathbb{P}_2}\right)$ mede a redução média de surpresa por evento do espaço amostral 1 provocada pelo conhecimento de um evento do espaço amostral 2.

Observe-se que, se os espaços amostrais são independentes, o conhecimento de evento de um espaço amostral não ensina nada sobre o outro espaço amostral. Desta forma a entropia mútua é nula, isto é:

$$\begin{aligned} h_{\mathbb{P}_1,\mathbb{P}_2}\left(X_{\mathbb{P}_1,\mathbb{P}_2}\right) &= h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) - h_{\mathbb{P}_1|\mathbb{P}_2}\left(X_{\mathbb{P}_1|\mathbb{P}_2}\right) \\ 0 &= h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) - h_{\mathbb{P}_1|\mathbb{P}_2}\left(X_{\mathbb{P}_1|\mathbb{P}_2}\right) \\ h_{\mathbb{P}_1|\mathbb{P}_2}\left(X_{\mathbb{P}_1|\mathbb{P}_2}\right) &= h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) \end{aligned}$$

O resultado anterior era o esperado.

Já se os espaços amostrais são idênticos (totalmente dependentes), o conhecimento sobre um evento de um espaço amostral ensina tudo sobre o outro

espaço amostral. Desta forma, a entropia mútua é a própria entropia de um espaço amostral. Matematicamente:

$$
\begin{aligned}
h_{\mathbb{P}_1,\mathbb{P}_2}\left(X_{\mathbb{P}_1,\mathbb{P}_2}\right) &= h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) - h_{\mathbb{P}_1|\mathbb{P}_2}\left(X_{\mathbb{P}_1|\mathbb{P}_2}\right) \\
h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) &= h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) - h_{\mathbb{P}_1|\mathbb{P}_2}\left(X_{\mathbb{P}_1|\mathbb{P}_2}\right) \\
h_{\mathbb{P}_1|\mathbb{P}_2}\left(X_{\mathbb{P}_1|\mathbb{P}_2}\right) &= 0
\end{aligned}
$$

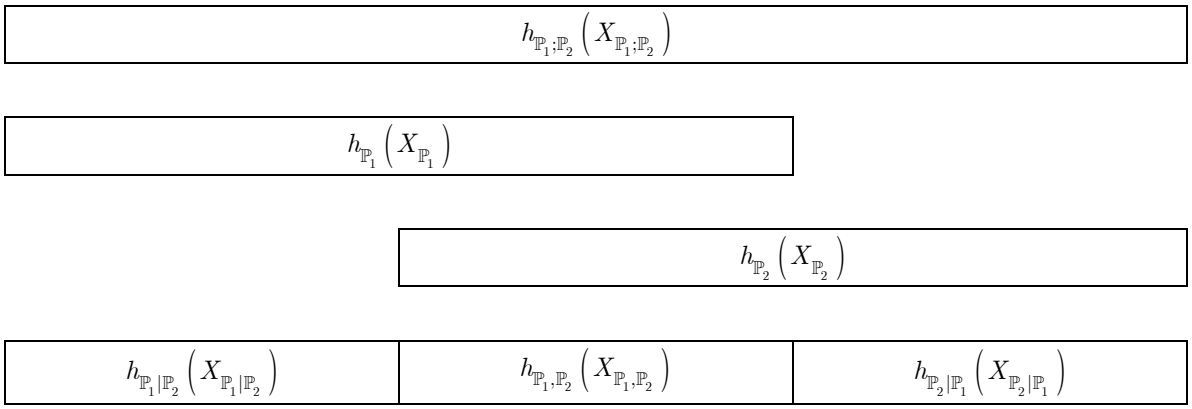
O resultado anterior também era o esperado.

Reza (1961, cap. 3) faz a analogia destas medidas da Teoria da Informação com a teoria dos conjuntos.

2.3.8 Interpretação gráfica das medidas de entropia

MacKay (1999, p. 53) fornece uma interpretação mais precisa sobre as relações entre estas medidas.

Figura 15: Representação gráfica em forma de barras das entropias entre duas variáveis aleatórias dependentes



2.3.9 Versões contínuas das medidas de entropia

As versões contínuas destas medidas são:

- Entropia marginal diferencial: $h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) = -\int f_{\mathbb{P}_1}(x)\log_2 f_{\mathbb{P}_1}(x)dx$
- Entropia conjunta diferencial: $h_{\mathbb{P}_1;\mathbb{P}_2}\left(X_{\mathbb{P}_1;\mathbb{P}_2}\right) = -\iint f_{\mathbb{P}}(x;y)\log_2 f_{\mathbb{P}}(x;y)dxdy$
- Entropia condicional diferencial: $h_{\mathbb{P}_1|\mathbb{P}_2}\left(X_{\mathbb{P}_1|\mathbb{P}_2}\right) = -\iint f_{\mathbb{P}}(x;y)\log_2\left(\frac{f_{\mathbb{P}}(x;y)}{f_{\mathbb{P}_2}(y)}\right)dxdy$

- Entropia mútua: $h_{\mathbb{P}_1,\mathbb{P}_2}\left(X_{\mathbb{P}_1,\mathbb{P}_2}\right) = -\iint f_{\mathbb{P}}(x;y)\log_2\left(\frac{f_{\mathbb{P}}(x;y)}{f_{\mathbb{P}_1}(x)f_{\mathbb{P}_2}(y)}\right)dxdy$

2.3.10 Relação entre entropia mútua e divergência

A entropia mútua é a seguinte divergência de Kullback-Leibler:

$$\begin{aligned} h_{\mathbb{P}_1,\mathbb{P}_2}\left(X_{\mathbb{P}_1,\mathbb{P}_2}\right) &= d\left(X_{\mathbb{P}} \mid\mid X_{\mathbb{P}_1*\mathbb{P}_2}\right) \\ &= d\left(X_{(\mathbb{P}_1|\mathbb{P}_2)*\mathbb{P}_2} \mid\mid X_{\mathbb{P}_1*\mathbb{P}_2}\right) \end{aligned}$$

Ao contrário da divergência, a entropia mútua goza da seguinte propriedade:

$$h_{\mathbb{P}_1,\mathbb{P}_2}\left(X_{\mathbb{P}_1,\mathbb{P}_2}\right) = h_{\mathbb{P}_2,\mathbb{P}_1}\left(X_{\mathbb{P}_2,\mathbb{P}_1}\right).$$

Como a entropia mútua é positiva e simétrica, ela possui propriedades que a tornam candidata a ser uma medida de distância.

A entropia mútua goza de importantes propriedades conforme Reza (1961, p. 275-8). Ela é a única medida que independe do fato de a variável aleatória ser contínua ou discreta em termos de interpretação.

2.3.11 Entropia mútua da normal bivariada e sua relação com o coeficiente de correlação e a independência Estatística

Conforme Haykin (1999, 507-8) e Reza (1961, 282-3), no caso de duas variáveis aleatórias terem uma distribuição normal conjunta, a entropia mútua adquire um novo e importante significado.

$$\begin{aligned} &(X_1;X_2) \sim \frac{d\mathbb{P}}{dx_1dx_2} = \text{normal}\left(\mu = \begin{bmatrix}0 & 0\end{bmatrix}; \sigma^2 = \begin{bmatrix}\sigma^2_{X_1} & \sigma_{X_1;X_2} \\ \sigma_{X_2;X_1} & \sigma^2_{X_2}\end{bmatrix}\right) \\ &\sigma_{X_1;X_2} = \sigma_{X_2;X_1} = E_{\mathbb{P}}\left[\left(X_1 - E_{\mathbb{P}_1}[X_1]\right)\left(X_2 - E_{\mathbb{P}_2}[X_2]\right)\right] \\ &\sigma^2_{X_1} = \sigma_{X_1;X_1} = E_{\mathbb{P}}\left[\left(X_1 - E_{\mathbb{P}_1}[X_1]\right)^2\right] \\ &\sigma^2_{X_2} = \sigma_{X_2;X_2} = E_{\mathbb{P}}\left[\left(X_2 - E_{\mathbb{P}_2}[X_2]\right)^2\right] \end{aligned}$$

O coeficiente de correlação é definido como:

$$\begin{aligned}\rho_{X_1;X_2} &= \frac{\sigma_{X_1;X_2}}{\sqrt{\sigma^2_{X_1}}\sqrt{\sigma^2_{X_2}}} \\ &= \frac{\sigma_{X_1;X_2}}{\sigma_{X_1}\sigma_{X_2}}\end{aligned}$$

Portanto, conforme Reza (1961, p. 283), a entropia conjunta é:

$$h_{\mathbb{P}_1;\mathbb{P}_2}\left(X_{\mathbb{P}_1;\mathbb{P}_2}\right) = \log_2 2\pi\sigma_{X_1}\sigma_{X_2}\sqrt{1-\rho^2_{X_1;X_2}}\ .$$

A entropia mútua é:

$$h_{\mathbb{P}_1,\mathbb{P}_2}\left(X_{\mathbb{P}_1,\mathbb{P}_2}\right) = -\frac{1}{2}\log_2\left(1-\rho^2_{X_1;X_2}\right).$$

Note-se que a entropia mútua depende apenas do coeficiente de correlação.

Um teorema da Estatística estabelece que se a distribuição conjunta é normal e o coeficiente de correlação é nulo, então as duas variáveis aleatórias são estatisticamente independentes. Por outro lado, quando as duas variáveis aleatórias são estatisticamente independentes, então o coeficiente de correlação é nulo.

Desta forma, quando duas variáveis aleatórias são estatisticamente independentes, a entropia mútua é nula.

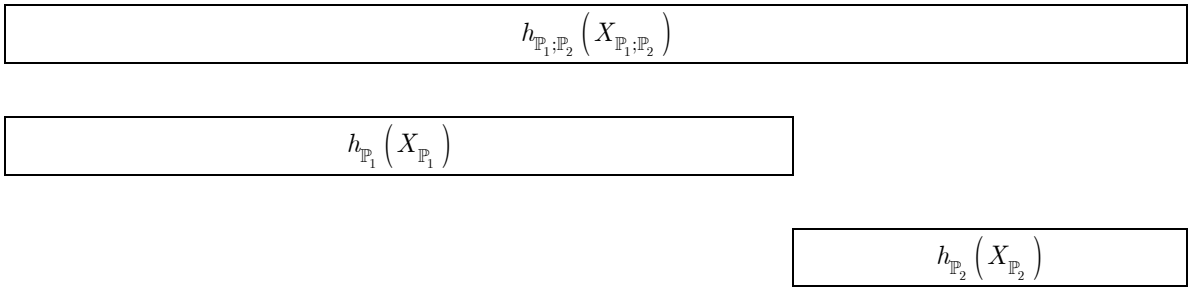
Conforme Haykin (1999, p. 514), quando a entropia mútua é nula, as duas variáveis aleatórias são estatisticamente independentes.

Black & Weigend (1998), *apud* Haykin (1999, p. 513), aplicaram a medida de entropia mútua na análise de dados do mercado financeiro com o intuito de extrair dos dados de ações o conjunto de componentes latentes independentes. Trata-se de uma generalização da análise de componentes principais, que extrai componentes não correlacionadas.

2.3.12 Entropia de um processo estocástico markoviano ou taxa de entropia

Quando duas variáveis aleatórias são independentes, graficamente temos as seguintes relações:

Figura 16: Representação gráfica em forma de barras das entropias entre duas variáveis aleatórias independentes



A primeira medida de probabilidade pode ser imagina como associada a um dado canônico e a Segunda, a uma moeda canônica.

Portanto, a entropia conjunta é:

$$\begin{aligned} h_{\mathbb{P}_1;\mathbb{P}_2}\left(X_{\mathbb{P}_1;\mathbb{P}_2}\right) &= h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) + h_{\mathbb{P}_2}\left(X_{\mathbb{P}_2}\right) \\ &= \log_2 6 + \log_2 2 \end{aligned}$$

Note-se que a entropia mútua é nula.

Caso elas sejam independentes e não identicamente distribuídas, então:

$$h_{\mathbb{P}_1;\mathbb{P}_2;\ldots;\mathbb{P}_n}\left(X_{\mathbb{P}_1;\mathbb{P}_2;\ldots;\mathbb{P}_n}\right) = \sum_{i=1}^{n} h_{\mathbb{P}_i}\left(X_{\mathbb{P}_i}\right).$$

Portanto, a entropia de n variáveis aleatórias independentes e identicamente distribuídas é:

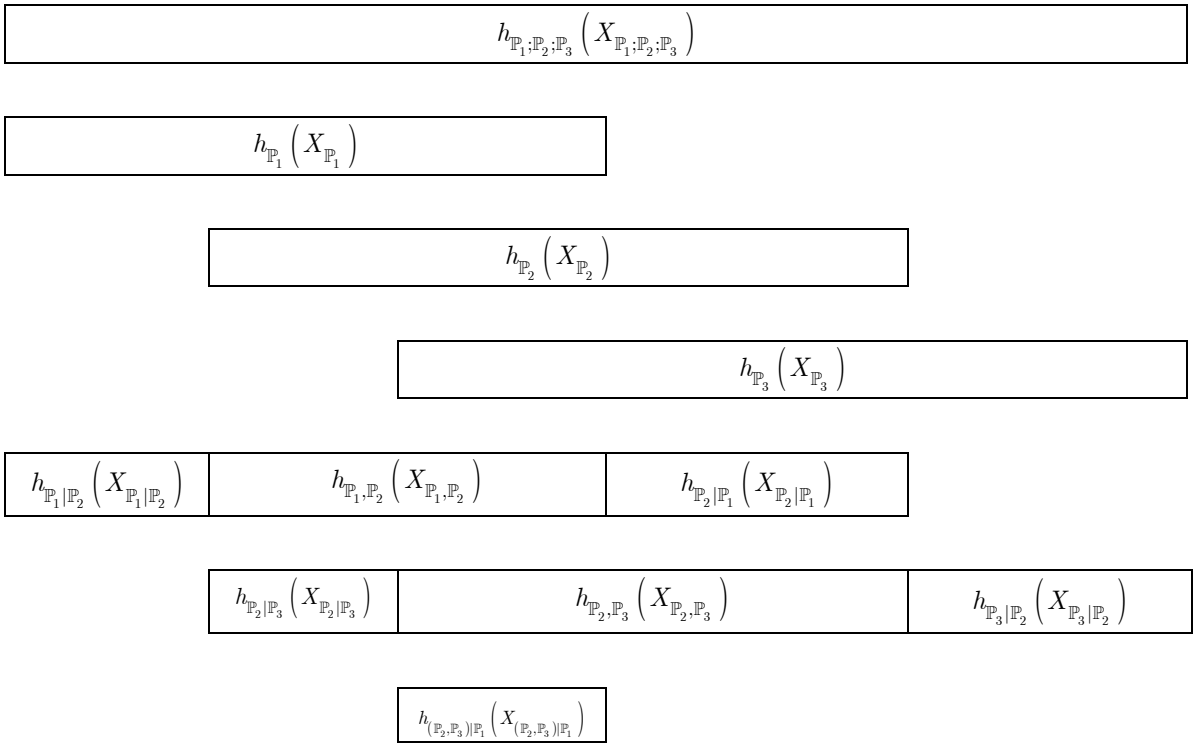
$$h_{\mathbb{P};\mathbb{P};\ldots;\mathbb{P}}\left(X_{\mathbb{P};\mathbb{P};\ldots;\mathbb{P}}\right) = nh_{\mathbb{P}}\left(X_{\mathbb{P}}\right).$$

Qual o valor da entropia de um conjunto de variáveis aleatórias dependentes?
Inicialmente será estudado o processo estocástico mais simples, isto é, a cadeia de Markov. Considere-se um processo estocástico $\{X_n; n = 0,1,2,\ldots\}$. Se $X_n = i$, então o processo é dito estar no estado i no estágio ou instante de tempo n. Seja $\Pr\left(X_{n+1} = j \mid X_n = i\right) = p_{ij} \geq 0$ a mesma probabilidade de transição do estado i para o j qualquer que seja n. Isto é, suponha-se que:

$$\Pr\left(X_{n+1} = j \mid X_n = i; X_{n-1} = i_{n-1}; \ldots; X_1 = i_1, X_0 = i_0\right) = p_{ij}$$

para todos os estados $i_0, i_1, \ldots, i_{n-1}, i, j$ e todo $n \geq 0$.
A cadeia de Markov pode ser representada da seguinte maneira, pela Teoria da Informação:

Figura 17: Representação gráfica em forma de barras das entropias de um processo estocástico markoviano



Note-se que $h_{\mathbb{P}_3|\mathbb{P}_2}\left(X_{\mathbb{P}_3|\mathbb{P}_2}\right) = h_{\mathbb{P}_2|\mathbb{P}_1}\left(X_{\mathbb{P}_2|\mathbb{P}_1}\right)$.

A entropia conjunta pode ser expressa em termos das condicionais de forma análoga à cadeia de Markov:

$$\begin{aligned} h_{\mathbb{P}_1;\mathbb{P}_2;\mathbb{P}_3}\left(X_{\mathbb{P}_1;\mathbb{P}_2;\mathbb{P}_3}\right) &= h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) + h_{\mathbb{P}_2|\mathbb{P}_1}\left(X_{\mathbb{P}_2|\mathbb{P}_1}\right) + h_{\mathbb{P}_3|\mathbb{P}_2}\left(X_{\mathbb{P}_3|\mathbb{P}_2}\right) \\ &= h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) + 2h_{\mathbb{P}_2|\mathbb{P}_1}\left(X_{\mathbb{P}_2|\mathbb{P}_1}\right) \end{aligned}.$$

Generalizando:

$$h_{\mathbb{P}_1;\mathbb{P}_2;\ldots;\mathbb{P}_n}\left(X_{\mathbb{P}_1;\mathbb{P}_2;\ldots;\mathbb{P}_n}\right) = h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) + nh_{\mathbb{P}_2|\mathbb{P}_1}\left(X_{\mathbb{P}_2|\mathbb{P}_1}\right)$$

A entropia do terceiro estágio da cadeia de Markov dados os dois primeiros é:

$$h_{\mathbb{P}_1;\mathbb{P}_2;\ldots;\mathbb{P}_n}\left(X_{\mathbb{P}_1;\mathbb{P}_2;\ldots;\mathbb{P}_n}\right) = h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) + nh_{\mathbb{P}_2|\mathbb{P}_1}\left(X_{\mathbb{P}_2|\mathbb{P}_1}\right)$$

A entropia do terceiro estágio condicionada aos dois primeiros é a entropia do segundo estágio condicionado ao primeiro. Este resultado vale para qualquer número de estágios consecutivos.

Note-se que este resultado pode ser generalizado, conforme Cover & Thomas (1991, p. 66), da seguinte maneira:

$$h_{\mathbb{P}_n|(\mathbb{P}_1;\mathbb{P}_2;\ldots;\mathbb{P}_{n-1})}\left(X_{\mathbb{P}_n|(\mathbb{P}_1;\mathbb{P}_2;\ldots;\mathbb{P}_{n-1})}\right) = h_{\mathbb{P}_2|\mathbb{P}_1}\left(X_{\mathbb{P}_2|\mathbb{P}_1}\right).$$

Conforme Cover & Thomas (1991, p. 66), $\{X_i; i = 1,2,\ldots,n,\ldots\}$ é uma cadeia de Markov estacionária com distribuição estacionária $\mathbb{P}_*^{(\infty)}$ e matriz de transição $\mathbb{P}^{(1)}$. Então a taxa de entropia ou entropia do processo estocástico markoviano é dada pela equação seguinte:

$$\begin{aligned}
E_{\mathbb{P}_*^{(\infty)}}\left(X_{\mathbb{P}^{(1)}}\right) &= \frac{1}{n} h_{\mathbb{P}_1;\mathbb{P}_2;\ldots;\mathbb{P}_n}\left(X_{\mathbb{P}_1;\mathbb{P}_2;\ldots;\mathbb{P}_n}\right) \\
&= \frac{1}{n}\left[h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) + n h_{\mathbb{P}_2|\mathbb{P}_1}\left(X_{\mathbb{P}_2|\mathbb{P}_1}\right)\right] \\
&= \frac{1}{n} h_{\mathbb{P}_1}\left(X_{\mathbb{P}_1}\right) + h_{\mathbb{P}_2|\mathbb{P}_1}\left(X_{\mathbb{P}_2|\mathbb{P}_1}\right) \\
&\to h_{\mathbb{P}_2|\mathbb{P}_1}\left(X_{\mathbb{P}_2|\mathbb{P}_1}\right) \text{ quando } n \to \infty
\end{aligned}$$

onde

$$X_{\mathbb{P}^{(1)}} : \left\{\Omega, \left\{\mathbb{P}_i^{(1)}; i = 1,2,\ldots,m\right\}\right\} \mapsto \xi = \left\{-\mathbb{P}_i^{(1)} \log_2 \mathbb{P}_i^{(1)} = h_{\mathbb{P}_i^{(1)}}\left(X_{\mathbb{P}_i^{(1)}}\right)\right\}$$

m = quantidade de estados

A unidade desta medida é bit/estado/estágio e representa a quantidade média de surpresa por evento por estágio do processo markoviano estacionário.

Para o caso particular de uma cadeia de Markov de dois estados, as matrizes de transição e a distribuição estacionária são, respectivamente:

$$\begin{aligned}
\mathbb{P}^{(1)} &= \begin{bmatrix} \mathbb{P}_1^{(1)} \\ \mathbb{P}_2^{(1)} \end{bmatrix} \\
&= \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix} \\
\mathbb{P}_*^{(\infty)} &= \begin{bmatrix} \dfrac{\beta}{\alpha+\beta} & \dfrac{\alpha}{\alpha+\beta} \end{bmatrix}
\end{aligned}$$

A entropia da variável aleatória $X_{n\to\infty}$ é:

$$h_{\mathbb{P}_\infty}\left(X_{\mathbb{P}_\infty}\right) = -\frac{\beta}{\alpha+\beta}\log_2\left(\frac{\beta}{\alpha+\beta}\right) - \frac{\alpha}{\alpha+\beta}\log_2\left(\frac{\alpha}{\alpha+\beta}\right).$$

A taxa de entropia do processo markoviano estacionário é:

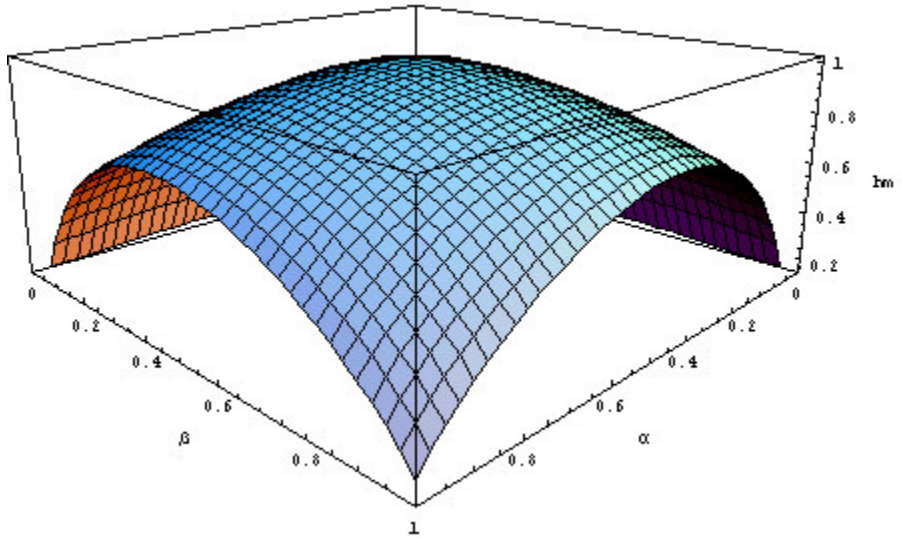
$$h_{\mathbb{P}_2|\mathbb{P}_1}\left(X_{\mathbb{P}_2|\mathbb{P}_1}\right) = \left(\frac{\beta}{\alpha+\beta}\right) h_{\mathbb{P}_1^{(1)}}\left(X_{\mathbb{P}_1^{(1)}}\right) + \left(\frac{\alpha}{\alpha+\beta}\right) h_{\mathbb{P}_2^{(1)}}\left(X_{\mathbb{P}_2^{(1)}}\right)$$

$$h_{\mathbb{P}_1^{(1)}}\left(X_{\mathbb{P}_1^{(1)}}\right) = -(1-\alpha)\log_2(1-\alpha) - \alpha\log_2\alpha$$

$$h_{\mathbb{P}_2^{(1)}}\left(X_{\mathbb{P}_2^{(1)}}\right) = -(1-\beta)\log_2(1-\beta) - \beta\log_2\beta$$

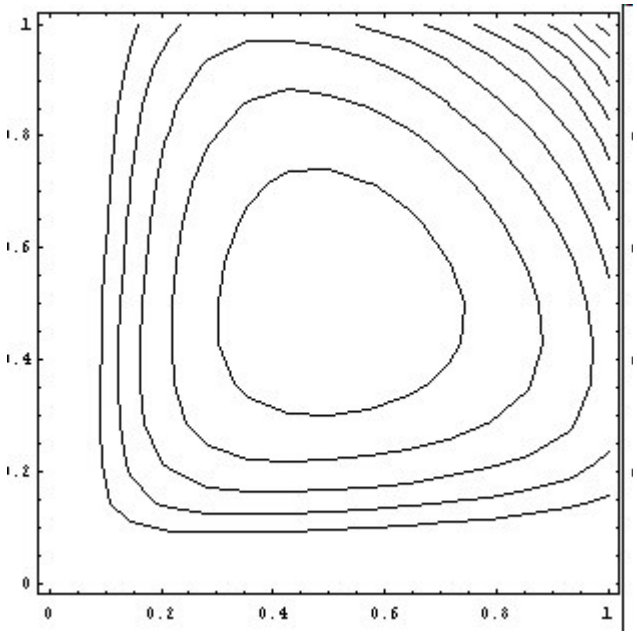
A superfície da taxa de entropia deste processo markoviano está representada no gráfico a seguir:

Figura 18: Gráfico de superfície da taxa de entropia de um processo markoviano estacionário de dois estados



As curvas de nível da taxa de entropia deste processo markoviano estão representadas no gráfico a seguir:

Figura 19: Gráfico de curvas de nível da taxa de entropia de um processo markoviano estacionário de dois estados



2.3.13 Entropia mútua total

A entropia mútua dos dois primeiros estágios da cadeia de Markov é:

$$\begin{aligned} h_{(\mathbb{P}_2,\mathbb{P}_3)|\mathbb{P}_1}\left(X_{(\mathbb{P}_2,\mathbb{P}_3)|\mathbb{P}_1}\right) &= h_{\mathbb{P}_1,\mathbb{P}_2}\left(X_{\mathbb{P}_1,\mathbb{P}_2}\right) - h_{\mathbb{P}_2|\mathbb{P}_3}\left(X_{\mathbb{P}_2|\mathbb{P}_3}\right) \\ &= h_{\mathbb{P}_2,\mathbb{P}_3}\left(X_{\mathbb{P}_2,\mathbb{P}_3}\right) - h_{\mathbb{P}_2|\mathbb{P}_1}\left(X_{\mathbb{P}_2|\mathbb{P}_1}\right) \end{aligned}$$

Poder-se-ia considerar esta última entropia como uma entropia mútua conjunta ou total dos três primeiros estágios consecutivos de uma cadeia de Markov, isto é:

$$h_{(\mathbb{P}_2,\mathbb{P}_3)|\mathbb{P}_1}\left(X_{(\mathbb{P}_2,\mathbb{P}_3)|\mathbb{P}_1}\right) = h_{\mathbb{P}_1,\mathbb{P}_2,\mathbb{P}_3}\left(X_{\mathbb{P}_1,\mathbb{P}_2,\mathbb{P}_3}\right).$$

Desta forma, temos que a entropia mútua conjunta (dependência) dos n primeiros estágios da cadeia de Markov decresce, tendendo para zero, isto é:

$$h_{\mathbb{P}_1,\mathbb{P}_2,\ldots,\mathbb{P}_n}\left(X_{\mathbb{P}_1,\mathbb{P}_2,\ldots,\mathbb{P}_n}\right) \to 0 \text{ quando } n \to \infty\,.$$

Observe-se que, na expressão a seguir, a medida de probabilidade conjunta entre dois estágios consecutivos da cadeia de Markov aproxima-se cada vez mais do produto das medidas de probabilidade, pois a dependência a cada passo diminui.

Desta forma, a divergência tende para zero também. Isto significa que o ganho de informação sucessivo é decrescente.

$$\begin{aligned} h_{\mathbb{P}_n,\mathbb{P}_{n+1}}\left(X_{\mathbb{P}_n,\mathbb{P}_{n+1}}\right) &= d\left(X_{\mathbb{P}} \mid\mid X_{\mathbb{P}_n * \mathbb{P}_{n+1}}\right) \\ &= d\left(X_{\left(\mathbb{P}_n \mid \mathbb{P}_{n+1}\right) * \mathbb{P}_{n+1}} \mid\mid X_{\mathbb{P}_n * \mathbb{P}_{n+1}}\right) \\ &\approx d\left(X_{\mathbb{P}_n * \mathbb{P}_{n+1}} \mid\mid X_{\mathbb{P}_n * \mathbb{P}_{n+1}}\right) \\ &= 0 \end{aligned}$$

2.3.14 Relação entre a entropia de Shannon e a de Boltzmann

"*A modern definition of entropy is a quantitative measure of the uncertainty associated with probability distribution, or the amount of missing information.*"

Yang (1997, p. xi)

Tendo por base Cover & Thomas (1991, p. 33-6) pode-se afirma que o conceito de entropia de Shannon contém o de Boltzmann, pois daquele pode-se derivar este.

O conceito de entropia termodinâmica de Boltzmann baseia-se na hipótese de que o comportamento das partículas é aleatório. Desta forma, o próximo estado do sistema fechado apenas depende do estado atual. Este sistema pode ser modelado pela cadeia de Markov.

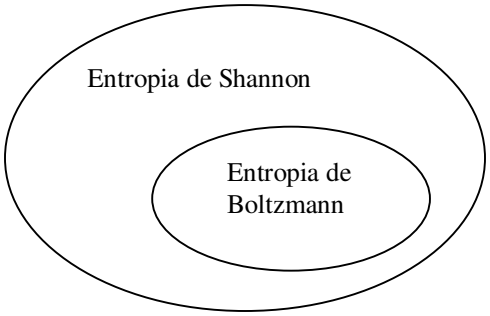
O macroestado mais equilibrado de todos em termos de quantidade de pontos de fase por célula é aquele em que os pontos de fase distribuem-se uniformemente entre as células do sistema fechado. Este macroestado especial pode ser chamado, desta forma, de **estado de equilíbrio** ou de **máxima probabilidade termodinâmica**. Se o sistema não está em equilíbrio, mudanças ocorrerão no sistema até que o estado de equilíbrio tenha sido atingido. Portanto, a mecânica Estatística interpreta o aumento da entropia num sistema fechado como uma conseqüência natural de um sistema tender de um estado menos provável para um mais provável. Quanto maior a desordem de um sistema, maior a sua probabilidade termodinâmica e maior a entropia. O mais alto grau de ordem de um sistema é ter todos os pontos de fase concentrados numa única célula. Neste estado a probabilidade termodinâmica relativa (w) é 1 e a entropia é nula. O estado de máxima entropia ou de equilíbrio não é um estado estático, pois os pontos de fase estão em permanente movimento no sistema. Logo, ocasionalmente o sistema

diferirá do estado de equilíbrio e, consequentemente, a entropia deixará de ser máxima. Pequenas mudanças são mais prováveis do que as grandes, mas estas não são impossíveis, apenas altamente improváveis. Este fenômeno é estudado na Teoria das Flutuações [Sears (1955, cap. 15)].

Do ponto de vista markoviano, quando o sistema entra em regime estacionário (equilíbrio) significa que a medida de probabilidade associada ao número de microestados do sistema fechado torna-se fixa.

No sentido termodinâmico esta medida de probabilidade estacionária deve ser uniforme para que a entropia do sistema fechado aumente.

No entanto, a medida de probabilidade estacionária uniforme ocorre se e somente se a matriz de transição da cadeia de Markov for duplamente estocástica, isto é, se a soma de cada linha e de cada coluna das probabilidades de transição somar um. Evidentemente, no caso mais geral da cadeia de Markov, apenas a soma de cada linha das probabilidades de transição (condicionais) deve somar um. Trata-se, portanto, de uma cadeia de Markov muito particular.



2.3.15 Entropia de ordem α e a distinção entre processos caóticos e aleatórios

Schützenberger (1954), Rényi (1961) e Pompe (1994), *apud* Golan *et al.* (1996, p. 36-7), apresentam o conceito de entropia generalizada. A expressão da entropia discreta generalizada é:

$$h_{\mathbb{P}}\left(X_{\mathbb{P}};\alpha\right) = \frac{1}{1-\alpha}\log_2 \sum_{i=1}^{n} p_k^{\alpha}$$
$$\alpha \in \mathbb{R}$$

A entropia discreta de Shannon é um caso particular da medida anterior quando $\alpha \to 1$.

Conforme Pompe (1994), *apud* Golan *et al.* (1996, p. 37), o parâmetro alfa é utilizado para distinguir processos caóticos dos aleatórios.

2.3.16 Divergência de Amari e a função utilidade HARA

Samperi (1998, p. 144) relata que a divergência de Kullback-Leibler generalizada de Rényi é:

$$d_r\left[X_{\mathbb{P}};X_{\mathbb{T}};\alpha\right]=\frac{1}{1-\alpha}\log_2\int f_{\mathbb{P}}^{\alpha}(x)f_{\mathbb{T}}^{1-\alpha}(x)dx\,.$$

Quando $\alpha\to 1$, há a divergência de Kullback-Leibler:

$$\begin{aligned}d_r\left[X_{\mathbb{P}};X_{\mathbb{T}};\alpha\to 1\right]&=\int f_{\mathbb{T}}(x)\log_2\left(\frac{f_{\mathbb{T}}(x)}{f_{\mathbb{P}}(x)}\right)dx\\&=d\left(X_{\mathbb{T}}\,||\,X_{\mathbb{P}}\right)\end{aligned}$$

Amari (1996), *apud* Samperi (1998, p. 144), define outra medida de divergência generalizada:

$$\begin{aligned}&d_a\left[X_{\mathbb{P}};X_{\mathbb{T}};\alpha\right]=\int f_{\mathbb{P}}(x)g_\alpha\left(\frac{f_{\mathbb{T}}(x)}{f_{\mathbb{P}}(x)}\right)dx\\&y=\frac{f_{\mathbb{T}}(x)}{f_{\mathbb{P}}(x)}\\&g_\alpha(y)=\begin{cases}\dfrac{4}{1-\alpha^2}\left(1-y^{\left(\frac{1+\alpha}{2}\right)}\right),\ \alpha\neq\pm1\\ y\log_2 y,\ \alpha=1\\ -\log_2 y,\ \alpha=-1\end{cases}\end{aligned}$$

Para alfa igual a 1, -1 e 0 têm-se as duas divergências de Kullback-Leibler e a distância de Hellinger [Samperi (1998, p. 145)].

$$\begin{aligned}&d_a\left[X_{\mathbb{P}};X_{\mathbb{T}};\alpha=1\right]=d\left(X_{\mathbb{T}}\,||\,X_{\mathbb{P}}\right)\\&d_a\left[X_{\mathbb{P}};X_{\mathbb{T}};\alpha=-1\right]=d\left(X_{\mathbb{P}}\,||\,X_{\mathbb{T}}\right)\\&d_a\left[X_{\mathbb{P}};X_{\mathbb{T}};\alpha=0\right]=2\int\left(\sqrt{f_{\mathbb{P}}(x)}-\sqrt{f_{\mathbb{T}}(x)}\right)^2dx\end{aligned}$$

Samperi (1998, p. 147) apresenta uma tabela reproduzida a seguir, na qual relaciona a função utilidade *HARA* (*hyperbolic absolute risk aversion*) [Pratt *et al.* (1996, apêndice 3) e Luenberger (1998, p. 256)] com a α-divergência de Amari.

A função utilidade HARA é:

$$u(x) = \frac{1-\gamma}{\gamma}\left(\frac{ax}{1-\gamma} + b\right)^{\gamma}, \; b > 0 \, .$$

O coeficiente de aversão à incerteza de Arrow-Pratt é:

$$\varsigma = \frac{1}{cx + d} \, .$$

α	γ	*Função utilidade HARA*
$-\infty$	$\to 1^{-}$	Linear (neutro à incerteza)
-3	1/2	Quadrática
-1	$\to 0^{\pm}$	Logarítmica
0	-1	Inversa (competitiva)
1	$\pm\infty$	Exponencial
3	2	Quadrática
$+\infty$	$\to 1^{+}$	Linear (neutro à incerteza)

Como pode ser observado, há uma estreita conexão entre função utilidade e divergência. Neste momento, tem-se definido o caminho da integração entre Teoria da Informação e Teoria Estatística Bayesiana da Decisão. A tese de Samperi (1998) abre este possível unificação de abordagens.

2.3.17 Processamento ótimo de informação incompleta ou estimação robusta com dados limitados utilizando o princípio da máxima entropia ou da mínima divergência

Conforme Jaynes (1996, p. ix) e Samperi (1998, p. 4; 215-8), a resolução do problema de formação de preço de opção pode ser entendido como um problema de processamento ótimo de informação incompleta. Este processamento utiliza o princípio de máxima entropia (=mínimo risco) ou de mínima divergência (=máxima utilidade). Samperi (1998, p. 118) afirma que a maximização da entropia é claramente equivalente à minimização da divergência com relação à distribuição prévia uniforme [Benish (1999) e Lence & Miller (1998)]. No entanto, a entropia relativa é invariante à mudança de coordenada, enquanto a entropia

não goza desta propriedade (Haykin (1999, p. 496)). Desta forma, é preferível utilizar o princípio de mínima entropia relativa. Em Samperi (1998, p.125) há uma interpretação geométrica deste princípio numa situação discreta. Samperi (1998, p. 221) afirma que a entropia relativa pode ser entendida como um automomento.

A aplicação destes princípios permite obter a maioria das distribuições de probabilidade [Kapur & Kesavan (1992), *apud* Samperi (1998, p. 11)]. Além disso, a entropia relativa está intimamente relacionada com o conceito de suficiência da Estatística e energia livre de um sistema.

O princípio de mínima entropia relativa (*cross-entropy*) generalizada é uma expansão do princípio de máxima entropia, pois considera também a distribuição prévia [Lence & Miller (1998, p. 854-5)].

Samperi (1998, p. 32; 187-9) demonstra que o problema de obter a medida de probabilidade neutralizadora do preço da incerteza pelo princípio de máxima entropia é o dual do problema de maximizar a utilidade esperada da riqueza terminal cuja função utilidade é exponencial (critério de Kelly Jr.). Este último problema é tratado por Luenberger (1998, cap. 9 e 15) e Cover & Thomas (1991).

Samperi (1998, cap. 4) estuda a relação da classe de funções utilidade chamada *HARA* (*hyperbolic absolute risk aversion*) [Pratt *et al.* (1996, apêndice 3)] com divergência.

Jaynes (1998, p. 104) afirma que a analogia com teorias físicas é muito mais profunda que uma mera analogia de método [Haykin (1999, p. 586-92) e Batschelet (1978, p. 361-5)]. Na p. xii, afirma que a maximização da entropia cria um modelo a partir dos dados disponíveis. Na p. 3-29, afirma que, enquanto a Estatística Clássica e Bayesiana são especulativas, a de máxima entropia é não especulativa, isto é, (i) o modelo criado é o mais simples para capturar toda informação das restrições e (ii) é o único modelo para o qual as restrições são Estatísticas suficientes. Portanto, evita-se a consideração de informações que não se têm e aproveita-se ao máximo as que se têm. Na p. xiii, afirma que o princípio de máxima entropia é a ferramenta apropriada (a mais segura) quando se tem

muito pouco conhecimento além dos dados brutos; ele protege o estudioso contra tirar conclusões não garantidas pelos dados disponíveis. Entretanto, quando a informação é extremamente vaga, pode ser difícil definir qualquer espaço amostral e poder-se-ia causar surpresa se ainda um princípio mais primitivo que este se encontrasse. Há espaço para muito mais pensamento criativo aqui. Na p. 2-18, afirma que qualquer distribuição de probabilidade é necessariamente subjetiva no sentido de descrever apenas um estado de conhecimento, e não pode de modo algum ser mensurada por um experimento físico. Ela pode ser objetiva se independer da personalidade do usuário (sentimentos como esperança, medos, valores de julgamento etc.). Na p. 3-29, afirma que, praticamente em todos os problemas reais de inferência científica, os dados são conhecidos, mas a hipótese correta não o é. Então, o problema enfrentado pelo cientista é de natureza inversa ao imposto pela regra dominante da pedagogia convencional (teoria da amostragem). Fornecidos os dados, qual é a probabilidade de alguma hipótese especificada ser verdadeira? De fato, a motivação do cientista para coletar dados é geralmente possibilitar que ele aprenda algo sobre o fenômeno. O objetivo do cientista é estabelecer uma conexão lógica entre os dados e as hipóteses.

2.3.17.1 Exemplo paradigmático do princípio de máxima entropia aplicado a um problema com indeterminação

Desta forma, Jaynes estabelece que, mesmo os problemas com indeterminação (*ill-posed*), podem e devem ser resolvidos.

Jaynes, *apud* Golan *et al.* (1996, p. 12-4), apresenta um problema sobre um experimento com o lançamento de um dado de seis faces. Seja X a variável aleatória associada ao lançamento do dado.

Evento	*Probabilidade*	*Valor*
w_1=face 1	p_1	$x_1 = 1$
w_2=face 2	p_2	$x_2 = 2$
w_3=face 3	p_3	$x_3 = 3$
w_4=face 4	p_4	$x_4 = 4$
w_5=face 5	p_5	$x_5 = 5$
w_6=face 6	p_6	$x_6 = 6$

Observe-se que:

$$\sum_{i=1}^{6} p_i = 1$$
$$p_i \geq 0$$

O problema consiste em estimar as probabilidades (incógnitas não observáveis) dos eventos do experimento.

Imagine-se que é sabido que o resultado médio de um grande número de lançamentos independentes deste dado é μ.

Mesmo com esta informação, não é possível utilizar o estimador de máxima verossimilhança de uma distribuição multinomial.

Este problema é claramente indeterminado (*ill-posed*) porque há seis incógnitas (probabilidades) e apenas duas fontes de informação, isto é, a soma das probabilidades é um e a média é μ.

A formulação a seguir é a de um problema de programação não-linear.

$$\max\ h = -\sum_{i=1}^{6} p_i \ln p_i$$
$$\text{restricoes:}$$
$$\sum_{i=1}^{6} p_i = 1$$
$$\sum_{i=1}^{6} p_i x_i = \mu$$

A medida de probabilidade de máxima entropia segue abaixo para alguns valores da média.

μ	$\hat{p}_1$	$\hat{p}_2$	$\hat{p}_3$	$\hat{p}_4$	$\hat{p}_5$	$\hat{p}_6$	max h
2	0,478	0,255	0,136	0,072	0,038	0,021	1,367
3	0,247	0,207	0,174	0,146	0,123	0,103	1,748
3,5	0,167	0,167	0,167	0,167	0,167	0,167	1,792
4	0,103	0,123	0,146	0,174	0,207	0,247	1,748
5	0,021	0,038	0,072	0,136	0,255	0,478	1,367

Note-se que, como era de esperar, quando $\mu = 3{,}5$, a distribuição de máxima entropia é a uniforme discreta.

Como os dados econômicos (de modo particular) e das ciências sociais costumam ser de natureza não experimental, isto é, observacionais, muitos dos modelos estatístico-econômicos que constituem a base da estimação e inferência são indeterminados [Golan *et al.* (1996, p. 1)].♦

Conforme Yang (1997) o princípio da máxima entropia (PME) foi estabelecido primeiramente por E. T. Jaynes, em 1957. O PME tem sido amplamente aplicado em termodinâmica e mecânica Estatística, análise de séries temporais, reconstrução de imagens, teoria das filas, teoria dos jogos e teoria da decisão, Economia, análise de carteira etc. O PME pode ser descrito como: quando é feita a inferência baseada em informação incompleta, deve-se fazê-la a partir da distribuição de probabilidade que possui a máxima entropia permitida pela informação disponível. Portanto, o PME fornece uma regra de processamento ótimo de informação. A entropia mede a quantidade de informação faltante, e a distribuição de entropia máxima é a distribuição menos subjetiva compatível com as restrições dadas, ou melhor, é a menos comprometida com as informações faltantes. Estatisticamente, isto implica que não há razão para preferir qualquer outra distribuição. O princípio geral da máxima entropia é aplicável a qualquer problema com espaço de hipóteses bem definido e informação incompleta.

Conforme Reza (1961, p. 278), a máxima entropia no caso discreto, com domínio finito, ocorre quando os eventos são igualmente prováveis. Já no caso contínuo ou com domínio infinito este resultado não se aplica. No caso contínuo a entropia diferencial pode ser negativa ou infinita também.

No entanto, estas situações ocorrem porque, na prática, faltam restrições que tornem a entropia relativa bem-comportada, isto é, limitada e positiva.

Cabe lembrar que, no caso discreto, a entropia tinha as restrições de não-negatividade das probabilidades e da soma destas ser igual a 1. Os valores da variável aleatória (se o conjunto de apoio ou domínio for finito), porém não interferem no cálculo da entropia discreta.

Já no caso contínuo (conjunto de apoio não enumerável) ou discreto infinito (conjunto de apoio enumerável), as restrições sobre o conjunto de apoio afetam a entropia diferencial e a entropia, respectivamente. Lembre-se que, no caso da medida de probabilidade normal, a entropia diferencial é função da variância dos valores da variável aleatória. Esta, por sua vez, é função do conjunto de apoio!

Desta forma, pode-se inverter o problema. Isto é, pode-se perguntar qual a medida de probabilidade que maximiza a entropia discreta ou diferencial sujeita às restrições plausíveis sobre a variável aleatória de interesse.

O problema invertido pode ser resolvido por métodos de otimização não-linear, como o dos multiplicadores de Lagrange.

Os três casos seguintes contínuos encontram-se em Reza (1961, p. 278-82) e Golan *et al.* (1996, p. 34-6).

2.3.17.1.1 Qual a função densidade de probabilidade que maximiza a entropia quando o domínio da variável aleatória é finito?

A distribuição uniforme contínua no intervalo dado é a resposta a esta pergunta.

$$\underset{f_{\mathbb{P}}(x)}{\text{Max}} \; -\int_{\xi} f_{\mathbb{P}}(x) \ln\left[f_{\mathbb{P}}(x)\right] dx$$

restricoes:

$$(0) \int_{\xi} f_{\mathbb{P}}(x) dx = 1$$

$$(1) \; f_{\mathbb{P}}(x) > 0, \; x \in \xi = [a;b]$$

As duas restrições garantem que a função procurada seja uma densidade de probabilidade.

Conforme Golan *et al.* (1996, cap. 3), Gulko (1998, cap. 1) e Reza (1967, cap. 8), pela equação de Euler tem-se que a função de densidade de probabilidade que maximiza a entropia diferencial é a uniforme contínua no conjunto de apoio dado:

$$\begin{aligned}
\ln f^*_{\mathbb{P}}(x) + 1 + \lambda^*_0 &= 0 \\
f^*_{\mathbb{P}}(x) &= \exp\left(-1 - \lambda^*_0\right) \\
&= \frac{1}{\exp\left(1 + \lambda^*_0\right)} \\
\lambda^*_0 &= \ln(b-a) - 1 \\
f^*_{\mathbb{P}}(x) &= \frac{1}{b-a}
\end{aligned}$$

Note-se que este resultado confere com a intuição de que a distribuição mais honesta (a mais objetiva) da variável aleatória, quando apenas o domínio finito é fornecido, é a uniforme contínua. O λ^*_0 é o multiplicador de Lagrange associado à restrição zero.

Se o domínio for discreto, a função densidade de máxima entropia será a uniforme discreta.♦

2.3.17.1.2 Qual a função densidade de probabilidade que maximiza a entropia quando o domínio da variável aleatória é positivo e sua esperança é conhecida?

A distribuição exponencial, tendo como parâmetro a esperança dada, é a resposta a esta pergunta.

$$\underset{f_{\mathbb{P}}(x)}{\text{Max}} \; -\int_{\xi} f_{\mathbb{P}}(x) \ln\left[f_{\mathbb{P}}(x)\right] dx$$

restricoes:

$$(0) \int_{\xi} f_{\mathbb{P}}(x) dx = 1$$

$$(1) \int_{\xi} x f_{\mathbb{P}}(x) dx = \theta_1$$

$$(2) \; f_{\mathbb{P}}(x) > 0, \; x \in \xi = [0; +\infty)$$

As restrições (0) e (2) garantem que a função procurada seja uma densidade de probabilidade.

A restrição (1) exige que a esperança (primeiro momento) seja igual a uma certa constante (média).

Conforme Golan *et al.* (1996, cap. 3), Gulko (1998, cap. 1) e Reza (1967, cap. 8), pela equação de Euler tem-se que a função de densidade de probabilidade que maximiza a entropia diferencial é a exponencial no conjunto de apoio dado:

$$\begin{aligned}
\ln f^*_{\mathbb{P}}(x) + 1 + \lambda^*_0 + x\lambda^*_1 &= 0 \\
f^*_{\mathbb{P}}(x) &= \exp\left(-1 - \lambda^*_0 - x\lambda^*_1\right) \\
&= \frac{\exp\left(-x\lambda^*_1\right)}{\exp\left(1 + \lambda^*_0\right)} \\
\lambda^*_0 &= -1 + \ln \theta_1 \\
\lambda^*_1 &= \frac{1}{\theta_1} \\
f^*_{\mathbb{P}}(x) &= \frac{1}{\theta_1} \exp\left(-\frac{1}{\theta_1} x\right)
\end{aligned}$$

Note-se que este resultado confere com a intuição de que a distribuição exponencial é a mais honesta (a mais objetiva) quando o domínio na forma de uma semi-reta e a média da variável aleatória são fornecidos. Os λ^*_0 e λ^*_1 são os multiplicadores de Lagrange, associados às restrições zero e um, respectivamente.

Se o domínio for discreto, a função densidade de máxima entropia será a geométrica [Johnson *et al.* (1992, cap. 5) e Kemp (1997)].♦

2.3.17.1.3 Qual a função densidade de probabilidade que maximiza a entropia quando o domínio da variável aleatória é infinito e sua variância é conhecida?

A distribuição normal com esperança nula e a variância dada é a resposta a esta pergunta.

$$\underset{f_{\mathbb{P}}(x)}{\text{Max}} \; -\int_{\xi} f_{\mathbb{P}}(x) \ln\left[f_{\mathbb{P}}(x)\right] dx$$

restricoes:

$$(0) \int_{\xi} f_{\mathbb{P}}(x) dx = 1$$

$$(1) \int_{\xi} x f_{\mathbb{P}}(x) dx = \theta_1 = 0 = \mu$$

$$(2) \int_{\xi} \left(x - \theta_1\right)^2 f_{\mathbb{P}}(x) dx = \int_{\xi} x^2 f_{\mathbb{P}}(x) dx = \theta_2 = \sigma^2$$

$$(3) \; f_{\mathbb{P}}(x) > 0, \; x \in \xi = \mathbb{R}$$

As restrições (0) e (3) garantem que a função procurada seja uma densidade de probabilidade.

A restrição (1) exige que a esperança (primeiro momento) seja igual a uma certa constante (média).

A restrição (2) exige que a variância (segundo momento centrado na média) seja igual a uma certa constante (variância).

Conforme Golan *et al.* (1996, cap. 3), Gulko (1998, cap. 1) e Reza (1967, cap. 8), pela equação de Euler tem-se que a função de densidade de probabilidade que maximiza a entropia diferencial é a normal no conjunto de apoio dado:

$$\begin{aligned}
\ln f^*_{\mathbb{P}}(x) + 1 + \lambda^*_0 + x\lambda^*_1 + x^2\lambda^*_2 &= 0 \\
f^*_{\mathbb{P}}(x) &= \exp\left(-1 - \lambda^*_0 - x\lambda^*_1 - x^2\lambda^*_2\right) \\
&= \frac{\exp\left(-x\lambda^*_1 - x^2\lambda^*_2\right)}{\exp\left(1 + \lambda^*_0\right)} \\
\lambda^*_0 &= \frac{1}{2}\ln\left(2\pi\sigma^2\right) - 1 \\
\lambda^*_1 &= 0 \\
\lambda^*_2 &= \frac{1}{2}\sigma^2 \\
f^*_{\mathbb{P}}(x) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right]
\end{aligned}$$

Note-se que este resultado confere com a intuição de que a distribuição normal é a mais honesta (a mais objetiva) quando o domínio na forma de uma reta, a média e a variância da variável aleatória são fornecidos. Os λ_0^*, λ_1^* e λ_2^* são os multiplicadores de Lagrange, associados às restrições zero, um e dois, respectivamente.

Se o domínio for discreto, a função densidade de máxima entropia será a normal discreta [Kemp (1997)].♦

As funções de densidade de probabilidade encontradas anteriormente são chamadas de **funções de densidade de probabilidade de máxima entropia**.

Se outros momentos forem considerados na determinação da função densidade de probabilidade, uma expressão generalizada encontra-se em Golan *et al.* (1996, p. 36).

2.3.17.2 Princípio de mínima divergência

Conforme foi mostrado anteriormente e em Samperi (1998, p. 118) maximizar a entropia é claramente equivalente a minimizar a divergência com relação à medida de probabilidade prévia uniforme. No entanto, como a divergência é invariante com relação à mudança de coordenada, ela é preferível.

Portanto, a minimização da divergência é uma generalização da maximização da entropia.

Samperi (1998, p. 119) sugere trabalhar com a seguinte medida, chamada de **divergência simétrica**, que possui as características de distância:

$$d\left(X_{\mathbb{P}};X_{\mathbb{T}}\right) = d\left(X_{\mathbb{T}} \mid\mid X_{\mathbb{P}}\right) + d\left(X_{\mathbb{P}} \mid\mid X_{\mathbb{T}}\right).$$

Esta divergência foi proposta por Kullback (1997, p. 6) e, segundo este autor, possui todas as propriedades de uma distância (métrica), como definida na topologia, com exceção da propriedade da desigualdade triangular, e, por isso, ela não pode ser chamada de distância.

O princípio de mínima divergência tem a seguinte formulação:

$$\underset{f_{\mathbb{P}}(x)}{\text{Max}} \; -\int_{\xi} f_{\mathbb{P}}(x) \ln\left[\frac{f_{\mathbb{P}}(x)}{f_{\mathbb{S}}(x)}\right] dx$$

restricoes:

$$(0) \int_{\xi} f_{\mathbb{P}}(x)dx = 1$$

$$(k) \int_{\xi} g_k(x) f_{\mathbb{P}}(x)dx = \theta_k, \; k = 1,2,...,n$$

$$(n+1) \; f_{\mathbb{P}}(x) > 0, \; x \in \xi$$

Note-se que, se as informações disponíveis sobre a variável aleatória são os momentos, então $g_k(x) = x^k$.

A função de densidade de probabilidade de mínima divergência é dada pela seguinte expressão:

$$f_{\mathbb{P}}^*(x) = \frac{f_{\mathbb{S}}(x) \exp\left(-\sum_{k=1}^{n} g_k(x)\lambda_k^*\right)}{\int_{\xi} f_{\mathbb{S}}(x) \exp\left(-\sum_{k=1}^{n} g_k(x)\lambda_k^*\right) dx}$$

Note-se que, quando $f_{\mathbb{S}}(x) = c$ conveniente, isto é, a função de densidade de probabilidade é uniforme, obtém-se a função de densidade de probabilidade de máxima entropia.

$$f_{\mathbb{P}}^*(x) = \frac{\exp\left(-\sum_{k=1}^{n} g_k(x)\lambda_k^*\right)}{\int_{\xi} \exp\left(-\sum_{k=1}^{n} g_k(x)\lambda_k^*\right) dx}$$

$$= \frac{\exp\left(-\sum_{k=1}^{n} g_k(x)\lambda_k^*\right)}{\exp\left(1 + \lambda_0^*\right)}$$

2.3.17.3 Axiomas de consistência

Shore & Johnson (1980), *apud* Haykin (1999, p. 535), provaram que o princípio de máxima entropia está correto no sentido de que, dado o conhecimento prévio na forma de restrições, há apenas uma distribuição que as satisfaz e que pode ser escolhida por um procedimento que satisfaz os axiomas de consistência. Esta distribuição é obtida pela maximização da entropia. Os axiomas de consistência são:

a) Unicidade: o resultado deve ser único.

b) Invariância: a escolha de coordenadas não deve afetar o resultado.

c) Independência de sistema: não deve importar se informação independente sobre sistemas independentes é explicada separadamente, em termos de diferentes densidades, ou juntas, em termos de uma densidade conjunta.

d) Independência de subconjunto: não deve importar se um subconjunto independente de estados do sistema é tratado em termos de uma densidade condicional separada ou em termos da densidade do sistema completo.

Eles também mostraram que a divergência de Kullback-Leibler também satisfaz estes axiomas.

3 Determinação entrópica do preço racional da opção européia simples ordinária sobre ação e *bond*

3.1 Introdução

Estimação e especificação da medida de probabilidade neutralizadora do preço da incerteza tornou-se um tema central em Finanças.

Conforme Yang (1997, p. 7), o princípio de não-arbitragem fornece uma condição suficiente para a existência de uma medida de probabilidade neutralizadora do preço da incerteza necessário para a formação de preço livre de arbitragem. Entretanto, ele não fornece um meio de determiná-la.

Ross (1976), *apud* Yang (1997, p. 5), considerou o conceito de arbitragem como o conceito unificador em Finanças.

Tornou-se comum caracterizar a medida de probabilidade neutralizadora do preço da incerteza, supondo um processo estocástico para o preço do valor mobiliário subjacente.

Conforme Harrison & Pliska (1981), *apud* Baxter & Rennie (1998, p. 197-8), (i) um mercado é livre de arbitragem se e somente se há pelo menos uma medida de probabilidade neutralizadora do preço da incerteza; (ii) e, neste caso, o mercado é completo se e somente se existe exatamente uma medida de probabilidade neutralizadora do preço da incerteza e mais nenhuma.

$$\underbrace{\text{hedge perfeito}}$$
$$\Updownarrow$$
$$\underbrace{\text{mercado completo}}$$
$$\Updownarrow$$
$$\text{existencia e unicidade da medida de probabilidade risco-neutro}$$
$$\Updownarrow$$
$$\underbrace{\text{preço unico e livre de arbitragem}}$$

Conforme Yang (1997, p. 12-3), o princípio de não-arbitragem estabelece que, se há uma carteira de valores mobiliários que replica o valor intrínseco de uma opção, então a opção e a carteira replicante devem ter o mesmo preço.

Ross (1976), *apud* Gulko (1998, p. 1-1), e Varian (1987) demonstram que não ocorrerá oportunidade de arbitragem se e somente se existir um vetor de preços estritamente positivos de ativos Arrow-Debreu na condição de mercado completo uniperíodo.

Conforme Ross (1976), *apud* Gulko (1998, p. 1-1), não ocorrerá oportunidade de arbitragem se e somente se existir um vetor de preços estritamente positivos de ativos Arrow-Debreu tal que:

$$\begin{aligned}
s(t) &= b(t)[\mathbb{Q}(\tau)]^T\, r(\tau) \\
&= b(t)[q_i(\tau)]^T \left[r_{ij}(\tau)\right] \\
&= b(t)E_{\mathbb{Q}(\tau)}\left\{\left[R_j(\tau)\right]\right\} \\
\text{em que } i &= \text{ estado da natureza} \\
j &= \text{ valor mobiliario} \\
b(t) &= \frac{b(\tau)}{1+r_f} \\
b(\tau) &= 1
\end{aligned}$$

Um mercado financeiro é dito completo uniperíodo se o posto da matriz de retornos $r(\tau)$ da data τ é igual ao número de estados da natureza m, i. e., se há m valores mobiliários linearmente independentes. A matriz $r(\tau)$ é uma matriz quadrada não-singular. Portanto, ela é inversível. Logo, o vetor de probabilidades risco-neutro é único:

$$[\mathbb{Q}(\tau)]^T = b^{-1}(t)s(t)[r(\tau)]^{-1}.$$

Se o posto de retornos for menor que m, então o mercado é dito ser incompleto uniperíodo e o vetor de probabilidades neutralizadora do preço da incerteza não é único. Portanto, o preço de um ativo não é único.

Conforme Cox & Rubinstein (1985) e Baxter & Rennie (1998), a condição de mercado completo multiperíodo pode ser obtida pela estratégia de carteira replicante, quando este mercado passa a se chamar **mercado dinamicamente completo** ou **completo multiperíodo**.

A abordagem econométrica e entrópica possibilita obter a medida única de probabilidade neutralizadora do preço da incerteza em condição de incerteza de mercado, sob as hipóteses de mercado perfeito, mercado em equilíbrio competitivo

e mundo com preço da incerteza neutralizado. Desta forma, com esta medida única de probabilidade, o preço racional da opção pode ser prontamente determinado. Além disso, este método entrópico não necessita das hipóteses normalmente utilizadas nos métodos tradicionais: mercado dinamicamente completo (hipótese esta equivalente à estratégia de *hedging* perfeito com a carteira replicante e autofinanciável dinâmica de valores mobiliários negociáveis) e processo estocástico markoviano específico associado ao movimento do preço do valor mobiliário subjacente (*e.g.* movimento browniano geométrico para o comportamento do preço da ação). Isto é, as hipóteses de mercado completo, eficiente e a medida de probabilidade real ou subjetiva (lognormal) não são necessárias. Conforme Baxter & Rennie (1998, p. 41), esta última condição, na verdade, sempre foi desnecessária, porém ela tem sido utilizada e a medida de probabilidade neutralizadora do preço da incerteza independe dela! Portanto, seu uso fere o princípio da parcimônia.

No espírito de Kelly Jr. (1956), quanto menos informações subjetivas (e.g. função de utilidade e probabilidade subjetiva) forem necessárias para a obtenção do preço racional, melhor (princípio da parcimônia).

O princípio da máxima entropia fornece um método prático, robusto e simples de determinar a medida de probabilidade neutralizadora do preço da incerteza, mesmo em condição de mercado incompleto e sem a especificação de um processo estocástico para o comportamento do preço do valor mobiliário subjacente.

3.2 O mundo informacionalmente eficiente é entrópico

A seguir será feita uma importante e nova conexão entre o mundo com preço da incerteza neutralizado de Yang (1997) e de Neftci (1996) e a hipótese de mercado informacionalmente eficiente de Gulko (1998). Será mostrado que esta última hipótese é um resultado deste mundo virtual. Logo, pelo princípio da parcimônia, ela é dispensável como hipótese para a determinação do preço da opção européia simples ordinária sobre ação e *bond.*

Lembrando que $\overline{S_i}(t) = \exp\left(-r_f t\right) S_i(t)$ é um martingale no mundo com preço da incerteza neutralizada e que martingales são variáveis aleatórias cujas variações

futuras são completamente imprevisíveis dada a história do preço. Isto é, a direção do movimento futuro do preço, num processo martingale, é impossível de ser prevista.

Então, o mundo com preço da incerteza neutralizada é informacionalmente eficiente, pois o investidor virtual é completamente ignorante a respeito da próxima mudança de preço. O preço corrente de valor mobiliário com incerteza reflete a expectativa de mercado (a média aritmética das expectativas subjetivas) sobre o desempenho do preço futuro. Um mercado competitivo equilibra multiformes expectativas subjetivas de tal modo que, a cada momento, metade dos investidores hipotéticos imagina que o preço de um valor mobiliário irá subir enquanto a outra metade acredita que o preço irá descer, isto é, os preços, neste mundo virtual, mantêm os investidores hipotéticos num estado de máxima incerteza com relação aos seus próximos movimentos de preço.

3.3 Expectativa de mercado

Conforme Gulko (1998) *apud* Yang (1997), expectativa de mercado é definida como a média aritmética de todas as expectativas dos investidores.

Note-se que estes investidores são os hipotéticos neutros à incerteza do mundo com preço da incerteza neutralizado.

Para construir o relacionamento entre o conceito de entropia e o de expectativas de mercado, associado com preços de ativos ou taxas de retorno em Finanças em condição de incerteza, assume-se que há n investidores no mercado.

O princípio de não-arbitragem assume que todos os investidores preferem mais a menos. Em acréscimo, admite-se que cada expectativa do investidor é igualmente provável se não se tem informação para atribuir um peso mais alto para algum particular ponto de vista do investidor. A suposição anterior é consistente com o princípio da razão insuficiente de Laplace que afirma que, a menos que haja informação em contrário, todos os resultados devem ser considerados como igualmente prováveis.

Suponha-se que há m estados associados com o preço futuro de um ativo S_τ. Há n investidores no mercado. O investidor k tem expectativa de que o estado j ocorra com probabilidade p_{kj}, sendo $k = 1,2,...,n$ e $j = 1,2,...,m$.

Considerando que as expectativas dos investidores têm o mesmo peso, então a expectativa de mercado para o estado j é simplesmente a média aritmética das expectativas individuais:

$$p_j = \frac{\sum_{k=1}^{n} p_{kj}}{n}.$$

Então a entropia da distribuição de probabilidade $\mathbb{P} = \{p_j; j = 1,2,...,m\}$ ou da expectativa de mercado é:

$$h_{\mathbb{P}}\left(X_{\mathbb{P}}\right) = -\sum_{j=1}^{m} p_j \ln p_j .$$

♦

Desta forma, o mundo informacionalmente eficiente é entrópico.

Portanto, a partir deste momento, este mundo sintético será denominado **mundo entrópico**.

Convém lembrar que o mundo entrópico originou-se de um mundo sem arbitragem, com linearidade de preços, com preço da incerteza neutralizado, martingale e, finalmente, de um mundo informacionalmente eficiente.

O exemplo a seguir mostrará uma maneira alternativa de obter a medida de probabilidade neutralizadora do preço da incerteza no mundo entrópico. Desta forma, evidencia-se o fato da não necessidade da hipótese de mercado completo e eficiente e da medida de probabilidade real ou subjetiva do valor mobiliário subjacente da data de exercício para a determinação do preço racional da opção.

3.4 Exemplo paradigmático da determinação do preço racional de uma opção racional pelo princípio da máxima entropia

No exemplo, conforme Gulko (1998, p. 1-4), as opções e os contratos futuros de eurodólar expiram em 16/09/1996= t_τ. Considere-se que hoje seja $t_0 = 0$ =01/07/1996 (data zero). Portanto, o prazo é $\tau = t_\tau - t_0 = t_\tau$. Estas opções serão tratadas como européias. Na data de exercício τ, as opções pagam $\max\{F_\tau - k_\tau;0\}$ em que k_τ é o preço de exercício (preço contratual futuro) e $\max\{F_\tau - k_\tau;0\}$ é o preço futuro estocástico na data τ. O preço futuro do eurodólar é determinado pela subtração da taxa de juros sobre os depósitos de eurodólar de três meses (Θ) de 100. Por exemplo, uma taxa de juros de 5,00% gera um preço futuro de 95. Em particular, na data de expiração os preços futuros convergem para 100 menos a taxa de juros *spot*. Considerar-se-ão cinco opções de compra de eurodólar que diferem apenas pelo preço de exercício[25]:

Opção	A	B	C	D	E
Preço de exercício	93,75	94,00	94,25	94,50	94,75
Preço da opção de compra européia	0,52	0,30	0,12	0,03	0,01

O preço de um ativo sem incerteza (título do Tesouro americano) que paga $1 em 16/09/1996 é $b = \dfrac{1}{1+r_f} = 0{,}98912$. Logo, a taxa certa neste período (de 01/07/96 até 16/09/96) é de 1,1%.

Um espaço de estados do preço do futuro na data de exercício, para que a matriz de retornos seja inversível, é:

$$F_\tau = \begin{bmatrix} 93,75 \\ 94,00 \\ 94,25 \\ 94,50 \\ 94,75 \\ 95,00 \end{bmatrix}$$

[25] *Wall Street Journal* (02/07/1996), *apud* Gulko (1998, p. 1-4).

Portanto, a matriz de valores intrínsecos r_τ, nos diferentes estados da natureza, é:

$$r_\tau = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0,25 & 0 & 0 & 0 & 0 \\ 1 & 0,50 & 0,25 & 0 & 0 & 0 \\ 1 & 0,75 & 0,50 & 0,25 & 0 & 0 \\ 1 & 1,00 & 0,75 & 0,50 & 0,25 & 0 \\ 1 & 1,25 & 1,00 & 0,75 & 0,50 & 0,25 \end{bmatrix}$$

O vetor de preços observados s das opções e do ativo sem incerteza é:

$$s = \begin{bmatrix} 0,98912 & 0,52 & 0,30 & 0,12 & 0,03 & 0,01 \end{bmatrix}$$

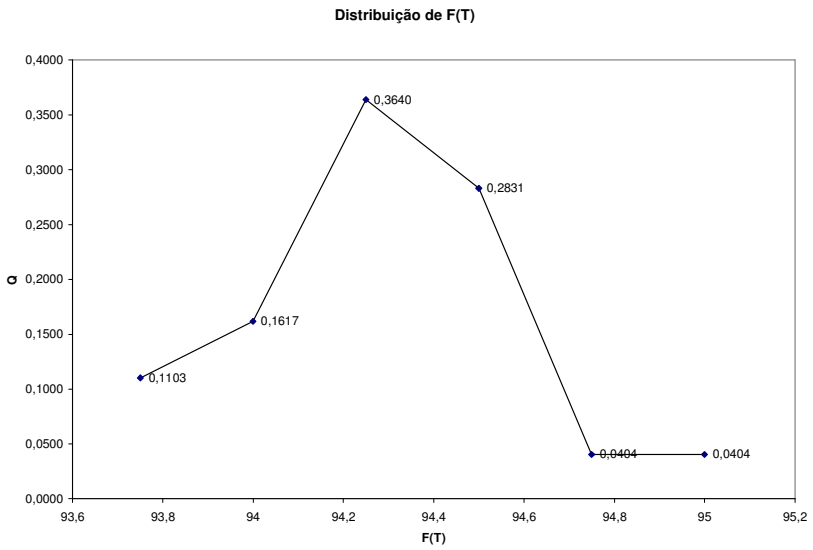
Como a matriz de valores intrínsecos possui posto seis, ela é posto completo. Isto implica que o mercado é completo. Logo, existe uma única medida de probabilidade neutralizadora do preço da incerteza $\mathbb{Q}$, dada pela expressão:

$$\begin{aligned} q &= b^{-1} s r_\tau^{-1} \\ &= \begin{bmatrix} 0,1103 & 0,1617 & 0,3640 & 0,2831 & 0,0404 & 0,0404 \end{bmatrix} \end{aligned}$$

3.4.1 Propriedades empíricas da medida de probabilidade neutralizadora do preço da incerteza

3.4.1.1 Propriedade empírica 1: a distribuição da medida de probabilidade neutralizadora do preço da incerteza é em forma de sino (unimodal)

A forma de sino da distribuição da medida de probabilidade neutralizadora do preço da incerteza sugere que esta medida reflete as expectativas de preços dos ativos-base dos investidores, apesar de $\mathbb{Q}$ ser uma medida de probabilidade abstrata sobre um espaço de estados da natureza ξ.



Qualquer conjunto de opções contemporâneas gera uma distribuição de probabilidade neutralizadora do preço da incerteza com forma de sino [Gulko (1998, p. 1-6)].

3.4.1.2 Propriedade empírica 2: o desvio-padrão na medida de probabilidade neutralizadora do preço da incerteza de F_τ é praticamente igual à volatilidade real de F_τ

O valor esperado e desvio-padrão de F_τ na medida $\mathbb{Q}$. são:

$$\begin{aligned}\mu_{F_\tau} &= E_{\mathbb{Q}}[F_\tau] \\ &= 94,2757 \\ \sigma_{F_\tau} &= \sqrt{Var_{\mathbb{Q}}[F_\tau]} \\ &= 0,2959\end{aligned}$$

Como $F_\tau = 100 - \Theta$, o desvio-padrão de F_τ é igual a volatilidade implícita da taxa de juros Θ de três meses, no período de dois meses e meio (de 01/07/96 a 16/09/96):

$$
\begin{aligned}
\mu_{\mathbb{Q}}(F_\tau) &= E_{\mathbb{Q}}[F_\tau] \\
&= E_{\mathbb{Q}}[100 - \Theta] \\
&= 100 - E_{\mathbb{Q}}[\Theta] \\
&= 94{,}2757 \\
E_{\mathbb{Q}}[\Theta] &= 100 - 94{,}2757 \\
&= 5{,}7243 \\
\sigma_{\mathbb{Q}}(F_\tau) &= \sqrt{Var_{\mathbb{Q}}[F_\tau]} \\
&= \sqrt{Var_{\mathbb{Q}}[100 - \Theta]} \\
&= \sqrt{Var_{\mathbb{Q}}[\Theta]} \\
&= \sigma_{\mathbb{Q}}(\Theta) \\
\sigma_{\mathbb{Q}}(\Theta) &= 0{,}2959
\end{aligned}
$$

É comum, para uma taxa de juros de curto prazo, que sua volatilidade varie entre 0,20 e 0,40 num período de dois meses e meio. O valor de 0,30 para a volatilidade é uma razoável estimativa para a taxa do eurodólar de três meses.

As volatilidades obtidas por meio dos preços das opções são sempre similares às reais do mercado (Gulko (1998, p. 1-6)).

Enquanto a teoria de arbitragem apresenta a medida de probabilidade neutralizadora do preço da incerteza como uma medida de probabilidade positiva sobre um espaço de estados da natureza ξ, as propriedades 1 e 2 sugerem que esta medida não é completamente abstrata. Além disso, a forma de sino desta medida sugere que a medida de probabilidade neutralizadora do preço da incerteza está sujeita a outras restrições que devem ser acrescidas à de positividade.

3.4.1.3 Propriedade empírica 3: nenhuma variação unilateral de um preço de opção por um tick *(0,01) aumenta a entropia acima de $h_{\mathbb{Q}}(F_\tau)$ dada a restrição de não poder aumentar o desvio-padrão acima de $\sigma_{\mathbb{Q}}(F_\tau)$.*

Para incluir novas restrições, serão examinados os efeitos das variações dos preços das opções na forma da medida $\mathbb{Q}$.

Duas medidas serão utilizadas para descrever a forma da distribuição: desvio-padrão e entropia.

A entropia de F_τ, na medida $\mathbb{Q}$, é dada por:

$$
\begin{aligned}
h_{\mathbb{Q}}(F_\tau) &= -\sum_{i=1}^{6} q_i \ln q_i \\
&= 1{,}522
\end{aligned}
$$

Quanto mais uniforme for a distribuição de F_τ, maior será o valor da entropia.

Quando o desvio padrão é zero, a entropia vale zero (valor mínimo). Por outro lado, quando a distribuição é uniforme, a entropia atinge o seu valor máximo.

Portanto, neste caso particular, a medida de probabilidade neutralizadora do preço da incerteza $\mathbb{Q}$ será caracterizada por $\left(\sigma_{\mathbb{Q}}(F_\tau); h_{\mathbb{Q}}(F_\tau)\right) = (0{,}2959; 1{,}522)$.

Suponha-se agora que os preços das opções variem continuamente para baixo e para cima o valor de 0,01 (um *tick*). Como as medidas de desvio-padrão e entropia variam?

Formalmente, para qualquer variação de preço unilateral da opção $s^* - s$ menor ou igual a um *tick* (0,01), isto é, $0 \le s^* \le s + 0{,}01$, verifica-se:

$$
\begin{aligned}
&\max\ h_{\mathbb{G}}(F_\tau) \\
&\text{sujeita as restriçoes:} \\
&g = b^{-1} s^* r_\tau^{-1} \\
&\sigma_{\mathbb{G}}(F_\tau) \le \sigma_{\mathbb{Q}}(F_\tau)
\end{aligned}
$$

Nestas condições, quando os preços das opções têm um aumento de até 0,01, a entropia pode chegar a aumentar, porém, ao final da variação do *tick*, ela sempre acaba menor que $h_{\mathbb{Q}}(F_\tau) = 1{,}522$. Portanto, a medida de probabilidade que maximiza a entropia de F_τ é $\mathbb{Q}$.

Figura 20: Entropia *versus* volatilidade [Gulko (1998, p. 1-8)]

A propriedade empírica 3 sugere a característica de máxima entropia dos preços das opções e da medida de probabilidade neutralizadora do preço da incerteza.

Seja I o conjunto das restrições denominado informação contida no conjunto de restrições. Então a propriedade de máxima entropia da medida de probabilidade neutralizadora do preço da incerteza $\mathbb{Q}^*$ pode ser formulada como:

$$\begin{aligned}\mathbb{Q}^* &= \arg\max_{\mathbb{G}} h_{\mathbb{G}}(F_\tau \mid I) \\ &\cong \mathbb{Q}\end{aligned}$$

A medida de probabilidade neutralizadora do preço da incerteza maximiza a entropia condicional de F_τ, isto é, a entropia condicional sobre o conjunto de informações I.

A medida de probabilidade neutralizadora do preço da incerteza $\mathbb{Q}^*$ é geralmente única em aplicações de interesse prático devido ao fato de $h_{\mathbb{G}}(F_\tau)$ ser uma função côncava sobre um conjunto convexo de medidas de probabilidade $\mathbb{G}$.

O conjunto I de informações determina completamente a medida $\mathbb{Q}^*$. No exemplo das opções de eurodólar, se os preços das cinco opções forem utilizados, basta inverter a matriz de resultados para determinar a única medida $\mathbb{Q}^*$. Se I inclui apenas quatro dos preços das opções e um limite superior para a variância, pode-se ainda determinar de maneira única a medida $\mathbb{Q}^*$ pela maximização da entropia de F_τ e, em seguida, formar o preço da opção omitida.

Qual é o menor conjunto informacional I necessário para construir a medida de probabilidade risco neutro? Dentro do espírito de Black-Scholes-Merton supõe-se que as medidas de probabilidade neutralizadora do preço da incerteza são completamente determinadas pelas informações fornecidas pelos ativos primários, tais como ações e *bonds*. Os ativos financeiros derivados, tais como opções, não oferecem informação adicional para a formação de preços.

Então o conjunto I contém apenas as informações dos ativos primários.

Por exemplo, pode-se reduzir este conjunto de informações ao preço atual do ativo-base e de sua volatilidade.

Para ilustrar os mecanismos e as implicações, considere-se o exemplo do mercado de opções de eurodólar. Primeiramente, será construída a medida $\mathbb{Q}^*$ sobre o

domínio $\xi = \{f_i; i = 1,2,...6\} = \{93,75; 94,00; 94,25; 94,50; 94,75; 95,00\}$, condicionada sobre o conjunto informacional $I = \{\xi; f; \sigma\}$ que contém o preço do futuro livre de arbitragem $f = \mu_{\mathbb{Q}}(F_\tau) = E_{\mathbb{Q}}(F_\tau) = 94,2757$ e a variância de F_τ na medida $\mathbb{Z}$ (observar Figura 20 em que $\mathbb{Z}$ é a medida com o mais alto valor de entropia), i.e., $\sigma^2_{\mathbb{Z}}(F_\tau) = Var_{\mathbb{Z}}(F_\tau) = 0,2925^2$ ao invés de $\sigma^2_{\mathbb{Q}}(F_\tau) = Var_{\mathbb{Q}}(F_\tau) = 0,2959^2$. Sendo assim, o valor do desvio-padrão de F_τ, quando a entropia é máxima, é a melhor estimativa de σ consistente com a hipótese de máxima entropia.

Em seguida, serão formados os preços das opções na medida $\mathbb{Q}^*$ e comparados com os valores cotados. Um mau ajuste rejeitaria a escolha de I e/ou a hipótese de mercado entrópico, enquanto um bom ajuste seria um sinal encorajador.

Dadas as estimativas de f e σ, a medida $\mathbb{Q}^*$ é dada pela resolução do seguinte programa de máxima entropia:

$$\begin{aligned}
&\max_{\mathbb{G}} \quad -\sum_{i=1}^{6} g_i \ln g_i \\
&\text{sujeita as restricoes:} \\
&E_{\mathbb{G}}(F_\tau) = f \\
&E_{\mathbb{G}}\left(F_\tau^2\right) = \sigma^2 + f^2 \\
&\sum_{i=1}^{6} g_i = 1 \\
&g_i > 0, \text{ para } i = 1,2,...,6
\end{aligned}$$

Mais detalhadamente:

$$\begin{aligned}
&\max_{\mathbb{G}} \quad -\sum_{i=1}^{6} g_i \ln g_i \\
&\text{sujeita as restricoes:} \\
&\sum_{i=1}^{6} g_i f_i = f \\
&\sum_{i=1}^{6} g_i f_i^2 = \sigma^2 + f^2 \\
&\sum_{i=1}^{6} g_i = 1 \\
&g_i > 0, \text{ para } i = 1,2,...,6
\end{aligned}$$

Numericamente:

$$\max_{\mathbb{G}} \quad -\sum_{i=1}^{6} g_i \ln g_i$$

sujeita as restricoes:

$$\sum_{i=1}^{6} g_i = 1$$

$$g_i 93,75 + g_i 94,00 + g_i 94,25 + g_i 94,50 + g_i 94,75 + g_i 95,00 = 94,2757$$

$$g_i 93,75^2 + g_i 94,00^2 + g_i 94,25^2 + g_i 94,50^2 + g_i 94,75^2 + g_i 95,00^2 = 0,2925^2 + 94,2757^2$$

$$g_i > 0, \text{ para } i = 1,2,...,6$$

A primeira (restrição da média) e a última restrição (restrição de positividade) garantem a que a solução é uma medida de probabilidade neutralizadora do preço da incerteza. A segunda restrição garante a condição de não-arbitragem ou a hipótese de mundo com preço da incerteza neutralizado. A terceira restrição garante a forma de sino da medida de probabilidade. A solução é única, pois a função objetivo (entropia) é côncava e está definida sobre um conjunto convexo.

A solução analítica deste problema de otimização é determinada pelo método dos multiplicadores de Lagrange. A solução é uma medida de probabilidade discreta simétrica em forma de sino sobre Ω[26]:

$$q_i^* = \frac{\exp\left(\lambda_1 f_i + \lambda_2 f_i^2\right)}{\sum_{i=1}^{6} \exp\left(\lambda_1 f_i + \lambda_2 f_i^2\right)}$$

$$\text{para } i = 1,2,...,6$$

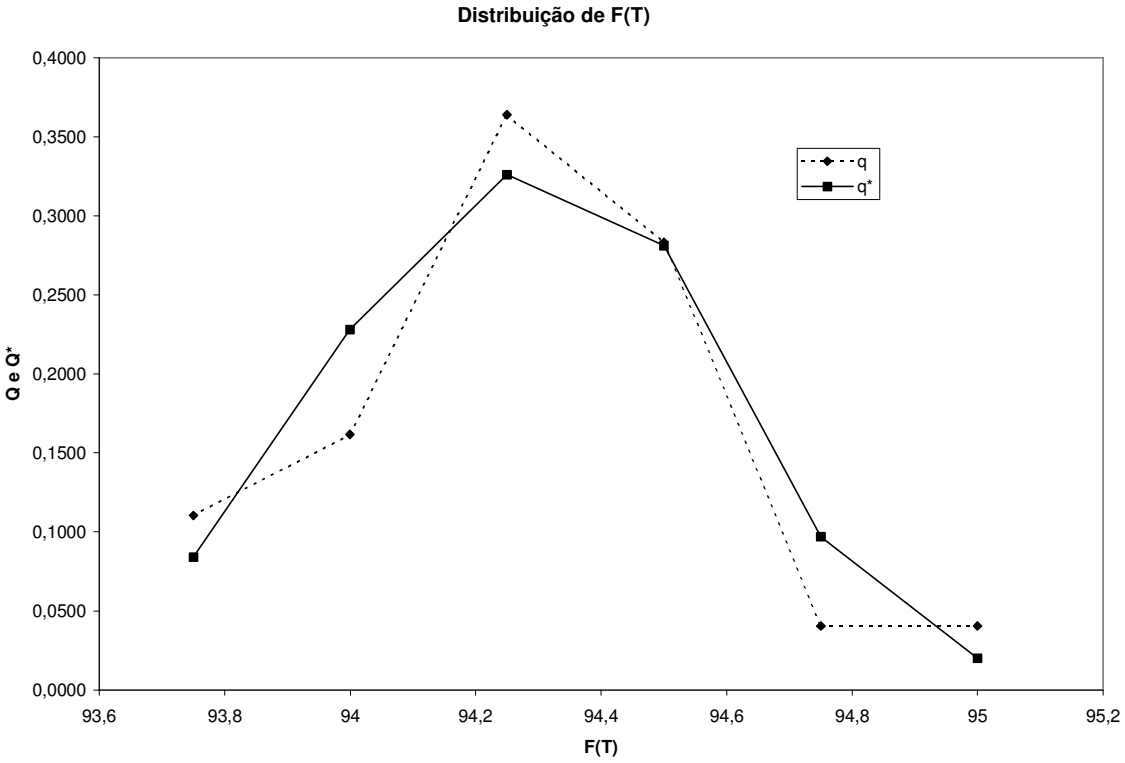
em que os multiplicadores de Lagrange λ_1 e λ_2 são determinados numericamente de tal modo que as restrições da média e da variância são satisfeitas. A distribuição encontrada é a normal discreta [Kemp (1997)].

Os valores numéricos da medida de probabilidade neutralizadora do preço da incerteza de máxima entropia $\mathbb{Q}^*$ são mostrados e plotados abaixo com a medida de probabilidade neutralizadora do preço da incerteza $\mathbb{Q}$ extraída diretamente a partir dos preços das opções. As duas distribuições são substancialmente similares. Em particular, a medida $\mathbb{Q}^*$ assemelha-se mais a um sino (mais simétrica e lisa), sendo assim consistente com a propriedade empírica 1, um

[26] A demonstração deste solução encontra-se em Gulko (1998, apêndice 1.1).

realístico desvio-padrão consistente com a propriedade empírica 2 e a propriedade de máxima entropia consistente com a propriedade empírica 3.

Figura 21: Distribuição de F_τ



Quadro 5: Estimativas das probabilidades neutralizadoras do preço da incerteza no mercado completo e entrópico

F_τ	93,75	94,00	94,25	94,5o	94,75	95,00
$\mathbb{Q}$ (mercado completo)	0,1103	0,1617	0,3640	0,2831	0,0404	0,0404
$\mathbb{Q}^*$ (mercado entrópico)	0,0840	0,2280	0,3260	0,2810	0,0970	0,0200

Quando os preços das opções calculados na medida de probabilidade neutralizadora do preço da incerteza são arredondados para o *tick* mais próximo, quase todos eles coincidem com os preços cotados.

Quadro 6: Preço racional da opção no mercado completo e no entrópico

	Preço de exercício (k)				
	93,75	94,00	94,25	94,50	94,75

Preços das opções de compra calculados na medida $\mathbb{Q}^*$ ($s^* = bq^* r_\tau$)	0,5199	0,2934	0,1235	0,0340	0,0050
Preço arredondado para o *tick* mais próximo	0,52	0,29	0,12	0,03	0,01
Preços das opções de compra cotados	0,52	0,30	0,12	0,03	0,01

Este exemplo canônico fornece três resultados adicionais:

(i) Um resultado positivo do teste – como deste do exemplo canônico – sustenta a propriedade de máxima entropia e a escolha do conjunto informacional *I.*

(ii) O conjunto informacional *I* deve incluir conhecimento disponível relevante sobre o preço do valor mobiliário subjacente. Este conjunto pode ser tão parcimonioso quanto o do exemplo canônico no qual o preço e sua volatilidade foram suficientes para a formação do preço da opção. No entanto, o conjunto informacional pode ser bem rico e incluir quaisquer dados relevantes adicionais na forma de restrições adicionais. Por exemplo, os dados adicionais poderiam ser os momentos de mais alta ordem da variável aleatória preço do ativo-base ou preços disponíveis das opções.

(iii) O mundo informacionalmente eficiente permite a formação de preços de opções na condição de mercado incompleto. Não foram necessárias seis equações para a obtenção da medida de probabilidade neutralizadora do preço da incerteza sobre um espaço de seis estados da natureza. Esta medida foi extraída de um conjunto parcimonioso de informações *I,* contendo apenas o domínio, o valor futuro do preço à vista pela taxa de juros certa e a variância do preço do ativo-base. Esta capacidade de formação de preço em mercado incompleto é um desvio da tecnologia de formação de preço vigente que se baseia na condição de mercado completo, mesmo sabendo-se que os mercados de capitais são incompletos.

Propriedade de geração (*property spanning*) dos preços das opções: diz-se que uma opção é gerada por ativos primários se o seu preço pode ser expresso como uma função apenas dos descritores dos ativos primários. Por exemplo, no modelo

de Black-Scholes-Merton, uma opção de ação é gerada por dois ativos primários: a ação subjacente e o ativo sem incerteza (*bond*) cujos descritores relevantes incluem o preço do *bond*, o preço da ação e a volatilidade do retorno da ação. Além disso, na economia de Black-Scholes-Merton, a geração de opção é equivalente à condição de mercado completo. Esta equivalência foi demonstrada, por exemplo, por Cox, Ross e Rubinstein em 1979[27].

A formação entrópica de preço goza da propriedade de geração se o conjunto informacional `I`, contendo apenas os descritores dos ativos primários, é suficiente para determinar preços mais realísticos de opções. Então um preço de opção pode ser expresso como uma função apenas dos ativos descritores dos ativos primários. O exemplo canônico apresentado exibe esta propriedade. Como a formação entrópica do preço da opção goza desta propriedade, então a geração de opção ocorre em mercado incompleto. Em tais casos, geração e condição de mercado completo são claramente não equivalentes quando comparados com os resultados clássicos de Black-Scholes (1973) e Merton (1973).

O que causa a divergência entre a abordagem entrópica e de Black-Scholes-Merton?

Na abordagem entrópica, a noção de geração reflete o montante de informação necessária para determinar um preço de opção, i.e., o montante de informação contido no conjunto informacional `I`. Este é a geração de opção informacional. A noção de geração de Black-Scholes-Merton, porém, baseia-se num conjunto de regras algébricas que são necessária para replicar fisicamente os valores intrínsecos (*payoffs*) da opção por uma carteira administrada, dinamicamente composta por ativo sem incerteza (*bond*) e ativo-base (ação) (carteira replicante). Isto é uma geração algébrica de preço da opção. A geração algébrica implica geração informacional. No argumento de Black-Scholes-Merton (argumento de replicação da opção), a informação fornecida pelos ativos primários determina os preços das opções. A geração informacional, por outro lado, não implica geração

---

[27] Cox, J.C.; Ross, S.A. & Rubinstein, M. 1979. "Option pricing: a simplified approach." *Journal of Financial Economics* 7:145-66

algébrica. Por exemplo, neste exemplo canônico os preços das opções, obtidos num mercado simples de eurodólar, utilizando a hipótese de mercado entrópico, não necessitaram de qualquer regra de replicação de opção. Portanto, a abordagem entrópica é menos restritiva que a abordagem de Black-Scholes-Merton. Em particular, a abordagem entrópica não se apóia na condição de mercado completo.

3.6 Determinação da medida de probabilidade neutralizadora do preço da incerteza de máxima entropia

Conforme Gulko (1999, p. 1-29:1-31), dado o conjunto informacional $I = \left\{\xi; r_f; s_0; \sigma^2\right\}$, a função de densidade de probabilidade neutralizadora do preço da incerteza de máxima entropia $q(s_\tau)$ da variável aleatória S_τ (preço do valor mobiliário na data de exercício $\tau = t_\tau$ sendo $t_0 = 0$) é a solução do seguinte programa:

$$q(s_\tau) = \arg\max_{f(s_\tau)} -\int_\xi f(s_\tau)\ln[f(s_\tau)]ds_\tau$$

restricoes:

$$(0)\ \int_\xi f(s_\tau)ds_\tau = 1$$

$$(1)\ \int_\xi s_\tau f(s_\tau)ds_\tau = s_0 \exp\left(r_f\tau\right)$$

$$(2)\ \int_\xi s_\tau^2 f(s_\tau)ds_\tau = \sigma^2 + \left[s_0 \exp\left(r_f\tau\right)\right]^2$$

$$(3)\ f(s_\tau) > 0,\ s_\tau \in \xi$$

Note-se que as restrições (0) e (3) garantem que a função seja uma densidade de probabilidade.

A restrição (1) garante que a medida de probabilidade seja neutralizadora do preço da incerteza. Ela expressa a hipótese de mundo com preço da incerteza neutralizado.

---

[28] Berger, J. 1985. *Statistical decision theory and bayesian analysis.* New York: Springer-Verlag.

A restrição (2) implementa a evidência empírica 1 (forma de sino) e pode ser relaxada para:

$$(2) \int_{\xi} s_{\tau}^{2} f(s_{\tau}) ds_{\tau} \leq \sigma^{2} + \left[s_{0} \exp\left(r_{f} \tau \right) \right]^{2}.$$

Conforme Gulko (1998, p. 1-28) a igualdade da restrição (2) significa que os investidores procuram um consenso sobre a volatilidade de S_{τ}. Já a desigualdade significa que os investidores procuram um consenso apenas sobre o limite superior da volatilidade. Matematicamente, ambas as restrições conduzem à mesma solução. Economicamente, a volatilidade pode ser considerada como um limite de volatilidade ao invés da volatilidade.

A função de densidade de probabilidade neutralizadora do preço da incerteza de máxima entropia $q(s_{\tau})$ é:

- Normal se $\xi = \mathbb{R} = (-\infty; +\infty)$, isto é, quando o preço pode assumir qualquer valor real.
- Normal truncada se $\xi = \mathbb{R}^{+} = [0; +\infty)$, isto é, quando o preço pode assumir valores positivos. Este é o domínio de variação do preço da ação.
- Normal duplamente truncada se $\xi = [0; 1]$, isto é, quando o preço pode assumir valores num intervalo fechado. Este é o domínio de variação do *bond* [Gulko (1998, p. 4-8)].

Gulko (1998, p. 1-30) chama a variável aleatória para a qual a função de densidade de probabilidade neutralizadora do preço da incerteza de máxima entropia de variável de estado. Portanto, a variável de estado anterior é o preço da ação ou do *bond*.

Este autor considera a variável aleatória retorno total da ação como:

$$R_{\tau} = \ln\left(\frac{S_{\tau}}{s_{0}} \right).$$

O domínio da taxa de retorno da ação é a reta real.

Para o *bond* o domínio é $(-\infty; -\ln s_{0}]$.

Desta forma, a função de densidade de probabilidade de máxima entropia do retorno total da ação é:

- Normal: este resultado é exatamente igual à hipótese assumida, em Finanças, de que a taxa de retorno da ação tem distribuição normal. No mercado entrópico, as distribuições normais são soluções para os problemas de inferência do investidor quando o conjunto de apoio é a reta real e a variância (volatilidade) é limitada superiormente.
- A função de densidade de probabilidade neutralizadora do preço da incerteza de máxima entropia do retorno total é obtida se seu valor esperado for $\mu_{R_\tau} = r_f - \frac{\sigma^2_{R_\tau}}{2}$.

Se a variável de estado - retorno total - possui distribuição normal, o preço da ação possui distribuição lognormal. Portanto, a função de densidade de probabilidade neutralizadora do preço da incerteza de máxima entropia do preço é a distribuição lognormal. Esta costuma ser outra hipótese utilizada em Finanças sobre a distribuição do preço da ação.

Conforme Gulko (1998, p. 1-30), se o desempenho do ativo é visto pelo investidor em termos espaço de preço ou do retorno total, trata-se de uma questão empírica. Note-se que esta escolha, por parte do investidor, determina se a medida de probabilidade neutralizadora do preço da incerteza é normal truncada ou lognormal.

Portanto, conforme Gulko (1998, p. 1-38), a versão entrópica generaliza o modelo de B-S-M. O modelo original de B-S-M admite que o preço do valor da ação obedece ao movimento browniano geométrico, a taxa de juros de curto prazo instantânea é constante e, também, que a volatilidade instantânea da taxa de retorno da ação é constante. A derivação entrópica relaxa estas restrições e, então, sugere que a fórmula de B-S-M pode ser muito mais geral e robusta do que usualmente se acredita.

Conforme Yang (1997, p. 35-6), Gulko (1995 e 1996) foi contraditório, pois, mesmo insistindo no fato de que o processo de difusão é irrelevante para a determinação da medida de probabilidade neutralizadora do preço da incerteza de

máxima entropia, ele os utiliza na determinação das medidas de probabilidade neutralizadora do preço da incerteza gama e beta.

Yang (1997, p. 36-44) utiliza o Sistema de Distribuições de Pearson [Johnson *et al.* (1994, p. 15-25)] e o relaciona com o princípio de máxima entropia para a obtenção das medidas de probabilidade neutralizadora do preço da incerteza.

Yang (1997, p. 34) e Gulko (1998, p. 3-10) ressaltam que, no caso da medida de probabilidade neutralizadora do preço da incerteza normal truncada no zero, a densidade não é zero, embora ela possa ser pequena. Isto implica que o preço pode ser nulo. Este resultado contradiz a restrição de que o preço deve ser estritamente positivo.

O resultado anterior sugere que as restrições sobre o conjunto de apoio devem ser mais bem definidas. Mesmo que o conjunto de apoio fosse definido como $\xi = (0;+\infty) = \mathbb{R}_*^+$, a solução continuaria sendo a normal truncada.

A autora argumenta que incorporando restrições com $g_k(x) = \ln x$, a medida de probabilidade neutralizadora do preço da incerteza correta é obtida.

3.6.1 Sistema de Distribuições de Pearson

Conforme Johnson *et al.* (1994, p. 15-6), entre 1890 e 1895 Pearson estabeleceu uma equação diferencial para geração de funções de densidade de probabilidade $f(x)$:

$$\frac{1}{f(x)}\frac{df(x)}{dx} = -\frac{a+x}{c_0 + c_1 x + c_2 x}.$$

Esta equação diferencial gera funções de densidade de probabilidade que dependem dos parâmetros a, c_0, c_1 e c_2.

Se $-a$ não é raiz da equação $c_0 + c_1 x + c_2 x = 0$, então $f(x)$ é finita quando $x = -a$. Além disso, $\left.\frac{df(x)}{dx}\right|_{x=-a} = 0$. Este resultado indica que o ponto $-a$ é de máximo ou moda.

Outro resultado evidente ocorre: $\left.\frac{df(x)}{dx}\right|_{f(x)=0} = 0$. Isto é, quando a função de densidade de probabilidade tem o valor nulo, isso significa que a curva tangencia a abscissa no ponto zero e/ou no infinito.

Se $x \neq -a$ e $f(x) \neq 0$, então $\frac{df(x)}{dx} \neq 0$.

A equação diferencial possui duas restrições:

$$(1) \int_{\xi} f(x)dx = 1$$
$$(2)\ f(x) \geq 0,\ x \in \xi$$

A maioria das distribuições possui uma forma semelhante a um sino, isto é, elas geralmente iniciam no zero, crescem até atingirem um valor máximo e, finalmente, decrescem até atingirem o zero. Isto é, elas costumam ser unimodais.

Se a taxa com a qual elas crescem é igual à taxa com a qual elas decrescem, a distribuição é simétrica. Se as taxas são diferentes, elas são assimétricas.

Pelo menos num dos extremos da distribuição há o contato com a abscissa. Geralmente, há o contato nos dois extremos.

Supondo que a variável aleatória é centrada no zero e que x_1 e x_2 são as raízes reais da equação $c_0 + c_1x + c_2x = 0$, as principais distribuições geradas pela equação diferencial de Pearson são [Yang (1997, p. 39-40) e Johnson *et al.* (1994, p. 15-25)] são:

- Tipo I e II ($x_1 \leq 0 \leq x_2$): beta
- Tipo III ($c_1 \neq 0$ e $c_2 = 0$): gama

3.6.2 Relação entre o Sistema de Distribuições de Pearson e o princípio da máxima entropia

Conforme Yang (1997, p. 40-44) há uma estreita relação entre o sistema de Pearson e o princípio da máxima entropia.

Para Kapur (1989), *apud* Yang (1997, p. 40), as distribuições de Pearson são apenas casos particulares do resultado derivado a partir do princípio da máxima

entropia quando são necessários os quatro primeiros momentos para determinar os quatro parâmetros que ocorrem na equação diferencial.

Nesta situação, teríamos a seguinte expressão derivada do princípio da máxima entropia:

$$\ln f^*(x) = -1 - \lambda_0^* - x\lambda_1^* - x^2\lambda_2^* - x^3\lambda_3^* - x^4\lambda_4^* .$$

Os lâmbdas são os multiplicadores de Lagrange.

Portanto:

$$\begin{aligned}
\frac{d}{dx}[\ln f^*(x)] &= \frac{d}{dx}\left(-1 - \lambda_0^* - x\lambda_1^* - x^2\lambda_2^* - x^3\lambda_3^* - x^4\lambda_4^*\right) \\
\frac{1}{f^*(x)}\frac{d}{dx}[f^*(x)] &= -\lambda_1^* - 2x\lambda_2^* - 3x^2\lambda_3^* - 4x^3\lambda_4^* \\
\frac{1}{f^*(x)}\frac{df^*(x)}{dx} &= -\lambda_1^* - 2x\lambda_2^* - 3x^2\lambda_3^* - 4x^3\lambda_4^*
\end{aligned}$$

Esta equação diferencial fornece distribuições trimodais sem restrição sobre o domínio.

Note-se que esta última equação é bem diferente da equação diferencial de Pearson.

Portanto, torna-se fácil perceber que o princípio de máxima entropia não reflete as restrições sobre o domínio (conjunto de apoio) da variável aleatória X. Já a equação diferencial é sensível às restrições sobre o domínio, incluindo as restrições sobre as suas extremidades.

Se a extremidade inferior da distribuição unimodal é conhecida, isto é, $\xi = (0;+\infty)$, então:

$$\frac{1}{f(x)}\frac{df(x)}{dx} = -\frac{a+x}{x\left(b_1 + b_2 x\right)} .$$

Desta forma, a equação diferencial gera apenas as distribuições do tipo I (beta) e III (gama).

Se $c_1 \neq 0$ e $c_2 = 0$, então:

$$\frac{1}{f(x)}\frac{df(x)}{dx} = -\frac{a+x}{c_0 + c_1 x} .$$

Desta forma, a equação diferencial gera apenas a distribuição do tipo III (gama). Note-se que o denominador é linear.

Se $c_1 = c_2 = 0$, então:

$$\frac{1}{f(x)}\frac{df(x)}{dx} = -\frac{a+x}{c_0}.$$

Desta forma, a equação diferencial gera apenas a distribuição normal. A normal não possui classificação no sistema de Pearson. O mesmo ocorre com a lognormal. Elas são casos limites de outras distribuições de Pearson.

Portanto, informações sobre o domínio contribuem para determinar os parâmetros da equação diferencial de Pearson.

3.6.3 Incorporação do sistema de Pearson no princípio de máxima entropia

A incorporação do sistema de Pearson no princípio de máxima entropia dá-se por meio de restrições que contenham informações sobre o domínio de variável aleatória.

Por exemplo, se é sabido que $x > 0$, pode-se utilizar $E_{\mathbb{P}}(\ln X)$, pois $\ln X$ está definido para este domínio.

Se $x < 1$, $E_{\mathbb{P}}[\ln(1-X)]$ pode ser utilizada.

Note-se que, se os dois primeiros momentos da variável aleatória são conhecidos, então $E_{\mathbb{P}}(\ln X)$ e $E_{\mathbb{P}}[\ln(1-X)]$ (médias geométricas) também são, pois, conforme Park & Sharp-Bette (1990, p. 398):

$$\begin{aligned}
E_{\mathbb{P}}(\ln X) &\approx \ln\left[E_{\mathbb{P}}(X)\right] - \frac{1}{2}\frac{Var_{\mathbb{P}}(X)}{E_{\mathbb{P}}^2(X)} \\
Var_{\mathbb{P}}(\ln X) &\approx \frac{Var_{\mathbb{P}}(X)}{E_{\mathbb{P}}^2(X)} \\
E_{\mathbb{P}}(\ln[1-X]) &\approx \ln\left[1-E_{\mathbb{P}}(X)\right] + \frac{1}{2}\frac{Var_{\mathbb{P}}(X)}{\left[1-E_{\mathbb{P}}(X)\right]^2} \\
Var_{\mathbb{P}}(\ln[1-X]) &\approx \frac{Var_{\mathbb{P}}(X)}{\left[1-E_{\mathbb{P}}(X)\right]^2}
\end{aligned}$$

3.7 *Determinação da medida de probabilidade neutralizadora do preço da incerteza de máxima entropia do preço do bond*

Portanto, para o preço do *bond*, tem-se o seguinte programa:

$$q(s_\tau) = \arg\max_{f(s_\tau)} -\int_\xi f(s_\tau)\ln[f(s_\tau)]ds_\tau$$

restricoes:

$$(0)\ \int_\xi f(s_\tau)ds_\tau = 1$$

$$(1)\ \int_\xi s_\tau f(s_\tau)ds_\tau = E_{\mathbb{Q}}(S_\tau) = s_0 \exp\left(r_f \tau\right)$$

$$(2)\ \int_\xi s_\tau^2 f(s_\tau)ds_\tau = E_{\mathbb{Q}}(S_\tau^2) = \sigma^2 + \left[s_0 \exp\left(r_f \tau\right)\right]^2$$

$$(3)\ \int_\xi \ln(s_\tau) f(s_\tau)ds_\tau = E_{\mathbb{Q}}[\ln(S_\tau)]$$

$$(4)\ \int_\xi \ln(1-s_\tau) f(s_\tau)ds_\tau = E_{\mathbb{Q}}[\ln(1-S_\tau)]$$

$$(5)\ f(s_\tau) > 0$$

Note-se que as restrições (3) e (4) representam, em forma de restrição, a informação sobre o domínio $s_\tau \in \xi = (0;1)$. Nos outros programas, esta informação foi ignorada pelo método dos multiplicadores de Lagrange. Agora, como ela foi explicitamente incorporada na forma de restrições, o domínio é considerado na obtenção da medida de probabilidade neutralizadora do preço da incerteza de máxima entropia.

Nesta situação, ter-se-ia a seguinte expressão derivada do princípio da máxima entropia:

$$\ln q(s_\tau) = -1 - \lambda_0^* - s_\tau\lambda_1^* - s_\tau^2\lambda_2^* - \ln(s_\tau)\lambda_3^* - \ln(1-s_\tau)\lambda_4^*.$$

Os lâmbdas são os multiplicadores de Lagrange.

Portanto:

$$\frac{1}{q(s_\tau)}\frac{dq(s_\tau)}{ds_\tau} = -\lambda_1^* - 2s_\tau\lambda_2^* - \frac{1}{s_\tau}\lambda_3^* + \frac{1}{1-s_\tau}\lambda_4^*$$

Esta equação diferencial fornece distribuições trimodais sem restrição sobre o domínio.

As restrições (1) e (2) são eliminadas para a determinação de uma distribuição unimodal, fazendo seus respectivos multiplicadores de Lagrange iguais a zero.

$$\frac{1}{q(s_\tau)}\frac{dq(s_\tau)}{ds_\tau} = -\frac{\lambda_3^* - (\lambda_3^* + \lambda_4^*)s_\tau}{s_\tau(1 - s_\tau)}$$

Ao derivar o lagrangiano em relação ao preço e exigir que a distribuição seja unimodal obtém-se a equação diferencial de Pearson. Portanto, existe uma relação direta entre o Sistema de Pearson e o princípio de máxima entropia.

Conforme Yang (1997, p. 43) a função de densidade de probabilidade neutralizadora do preço da incerteza de máxima entropia do preço do *bond* é uma beta cujos parâmetros dependem apenas dos multiplicadores de Lagrange, associados às restrições (3) e (4):

$$q(s_\tau \mid \beta_1;\beta_2) = \frac{1}{\mathrm{B}(\beta_1;\beta_2)} s_\tau^{\beta_1 - 1}(1 - s_\tau)^{\beta_2 - 1}$$
$$\beta_1 > 0$$
$$\beta_2 > 0$$

$\mathrm{B}(\beta_1;\beta_2) = \int_0^1 s_\tau^{\beta_1 - 1}(1 - s_\tau)^{\beta_2 - 1}\, ds_\tau$ é a função beta.

Novamente, Yang (1997, p. 44) relata que Gulko (1995 e 1996) assume processos de difusão tanto para o preço da ação quanto para o *bond* e que estas hipóteses tornam a obtenção da medida de probabilidade neutralizadora do preço da incerteza de máxima entropia uma tarefa difícil.

Apesar da aparente contradição metodológica apontada por Yang (1997) sobre o trabalho de Gulko até 1996, Gulko (1998, p. 4-17) afirma que, apesar de impor um processo de difusão para a variável de estado, a densidade de máxima entropia possui a propriedade útil de ser relativamente robusta com relação à escolha do processo de difusão da variável de estado.

Gulko (1998, p. 4-16) encontra uma medida de probabilidade neutralizadora do preço da incerteza de máxima entropia do preço do *bond* sem cupom.

Gulko (1998, p. 4-12) considera a existência de dois ativos num mercado perfeito e eficiente no intervalo de tempo $[0;\tau]$: um *bond* sem risco e outro com risco e sem cupom. O *bond* sem risco paga 1 na data de maturidade τ. O *bond* com risco

promete pagar 1 na data de maturidade $t > \tau$. O preço do *bond* com risco S_τ é uma variável aleatória cujo domínio é $\xi = [0;1]$.

Na data 0 os preços dos *bonds* e a variância do preço do *bond* com risco são conhecidos. Portanto, o conjunto informacional é $I = \left\{\xi = (0;1], r_f, s_0, \sigma^2\right\}$.

A função de densidade de probabilidade neutralizadora do preço da incerteza de máxima entropia é:

$$
\begin{aligned}
q\left(s_\tau \mid \beta_1;\beta_2\right) &= \frac{1}{\mathrm{B}(\beta_1;\beta_2)} s_\tau^{\beta_1-1}(1-s_\tau)^{\beta_2-1} \\
\mathrm{B}(\beta_1;\beta_2) &= \int_0^1 s_\tau^{\beta_1-1}(1-s_\tau)^{\beta_2-1}\, ds_\tau,\ \beta_1 > 0 \text{ e } \beta_2 > 0 \\
\beta_1 &= \omega\left[\frac{\omega(1-\omega)}{\sigma^2} - 1\right] \\
\beta_2 &= (1-\omega)\left[\frac{\omega(1-\omega)}{\sigma^2} - 1\right] \\
\omega &= s_0 \exp\left(r_f \tau\right)
\end{aligned}
$$

Note-se que:

$$
\begin{aligned}
E_{\mathbb{Q}}(S_\tau) &= \frac{\beta_1}{\beta_1+\beta_2} \\
&= \omega \\
&= s_0 \exp\left(r_f \tau\right)
\end{aligned}
\quad \text{e} \quad
\begin{aligned}
Var_{\mathbb{Q}}(S_\tau) &= \frac{\beta_1\beta_2}{\left(\beta_1+\beta_2\right)^2\left(\beta_1+\beta_2+1\right)} \\
&= \sigma^2
\end{aligned}
$$

Para simplificar este processo de obtenção, utilizar-se-á a idéia do Sistema de Pearson incorporando-a na forma de restrições.

3.8 *Determinação do preço único é livre de arbitragem da opção de compra européia simples ordinária sobre um bond*

Apesar de Yang (1997, p. 44) ter facilitado o processo de obtenção da medida de probabilidade neutralizadora do preço da incerteza , na página 52, a autora atinge uma fórmula *aproximada* do preço da opção européia simples ordinária de um *bond* sem cupom com risco.

Em seguida, Yang (1997, p. 53-4) conclui que uma segunda aproximação gera a mesma fórmula exata encontrada por Gulko (1998, p. 4-18:4-19):

$$c_0 = s_0 \overline{\mathbb{Q}}(s_\tau \mid \beta_1 + 1; \beta_2; k_\tau) - k_\tau \exp(-r_f \tau) \overline{\mathbb{Q}}(s_\tau \mid \beta_1; \beta_2; k_\tau)$$

$$\overline{\mathbb{Q}}(s_\tau \mid \beta_1; \beta_2; k_\tau) = 1 - \mathbb{Q}(s_\tau \mid \beta_1; \beta_2; k_\tau)$$

$$\mathbb{Q}(s_\tau \mid \beta_1; \beta_2; k_\tau) = \int_0^{k_\tau} q(s_\tau \mid \beta_1; \beta_2) ds_\tau$$

Se o *bond* com risco promete pagar θ na data de maturidade $t > \tau$, conforme Gulko (1998, p. 4-19) tem-se:

$$c_0 = s_0 \overline{\mathbb{Q}}\left(\frac{k_\tau}{\theta} \mid \beta_1 + 1; \beta_2\right) - k_\tau \exp(-r_f \tau) \overline{\mathbb{Q}}\left(\frac{k_\tau}{\theta} \mid \beta_1; \beta_2\right).$$

3.9 *Determinação da medida de probabilidade neutralizadora do preço da incerteza de máxima entropia do preço de ação*

Portanto, para o preço da ação como variável de estado, tem-se o seguinte programa:

$$q(s_\tau) = \arg\max_{f(s_\tau)} -\int_\xi f(s_\tau) \ln[f(s_\tau)] ds_\tau$$

restricoes:

$$(0) \int_\xi f(s_\tau) ds_\tau = 1$$

$$(1) \int_\xi s_\tau f(s_\tau) ds_\tau = E_{\mathbb{Q}}(S_\tau) = s_0 \exp(r_f \tau)$$

$$(2) \int_\xi s_\tau^2 f(s_\tau) ds_\tau = E_{\mathbb{Q}}(S_\tau^2) = \sigma^2 + \left[s_0 \exp(r_f \tau)\right]^2$$

$$(3) \int_\xi \ln(s_\tau) f_{\mathbb{Q}}(s_\tau) ds_\tau = E_{\mathbb{Q}}[\ln(S_\tau)]$$

$$(4)\ f(s_\tau) > 0$$

Note-se que a restrição (3) representa, em forma de restrição, a informação sobre o domínio $s_\tau \in \xi = (0; +\infty)$. Nos outros programas, esta informação foi ignorada pelo método dos multiplicadores de Lagrange. Agora, como ela foi explicitamente incorporada na forma de restrições, o domínio é considerado na obtenção da medida de probabilidade neutralizadora do preço da incerteza de máxima entropia.

Nesta situação, ter-se-ia a seguinte expressão derivada do princípio da máxima entropia:

$$\ln q(s_\tau) = -1 - \lambda_0^* - s_\tau \lambda_1^* - s_\tau^2 \lambda_2^* - \ln(s_\tau) \lambda_3^*.$$

Os lâmbdas são os multiplicadores de Lagrange.

Portanto:

$$\frac{1}{q(s_\tau)}\frac{dq(s_\tau)}{ds_\tau} = -\lambda_1^* - 2s_\tau\lambda_2^* - \frac{1}{s_\tau}\lambda_3^*$$

Esta equação diferencial fornece distribuições bimodais sem restrição sobre o domínio.

A restrição (2) é eliminada para a determinação de uma distribuição unimodal, fazendo seu respectivo multiplicador de Lagrange igual a zero.

$$\frac{1}{q(s_\tau)}\frac{dq(s_\tau)}{ds_\tau} = -\frac{\lambda_3^* - \lambda_1^* s_\tau}{s_\tau}$$

Note-se que a solução da equação diferencial de Pearson é a distribuição gama, pois o denominador é linear. Gulko (1998) também atinge esta distribuição para o preço da ação.

Yang (1997, p. 56-9), contraditoriamente ao que vinha desenvolvendo, não utiliza o preço da ação como variável de estado.

Nesta tese, continua-se admitindo que a variável de estado é o preço da ação.

Gulko (1998, p. 3-12:3-13) também utiliza o preço da ação como variável de estado. A função de densidade de probabilidade neutralizadora do preço da incerteza de máxima entropia do preço da ação obtida por ele é a gama cujos parâmetros dependem apenas dos multiplicadores de Lagrange associados às restrições (1) e (3).

Na data 0 os preços da ação e do *bond* e a variância do preço do ação são conhecidos. Portanto, o conjunto informacional é $\mathsf{I} = \left\{\xi = (0;+\infty), r_f, s_0, \sigma^2\right\}$.

A função de densidade de probabilidade neutralizadora do preço da incerteza de máxima entropia é:

$$
\begin{aligned}
q\left(s_\tau \mid \beta_1;\beta_2\right) &= \frac{\beta_2^{\beta_1}}{\Gamma(\beta_1)} s_\tau^{\beta_1-1} \exp\left(-\beta_2 s_\tau\right) \\
\Gamma(\beta_1) &= \int_0^\infty s_\tau^{\beta_1-1} \exp(s_\tau)\, ds_\tau,\ \beta_1 > 0 \\
\beta_1 &= \frac{\omega^2}{\sigma^2} \\
\beta_2 &= \frac{\omega}{\sigma^2} \\
\omega &= s_0 \exp\left(r_f \tau\right)
\end{aligned}
$$

$\Gamma(\beta_1)$ é a função gama.

Note-se que:

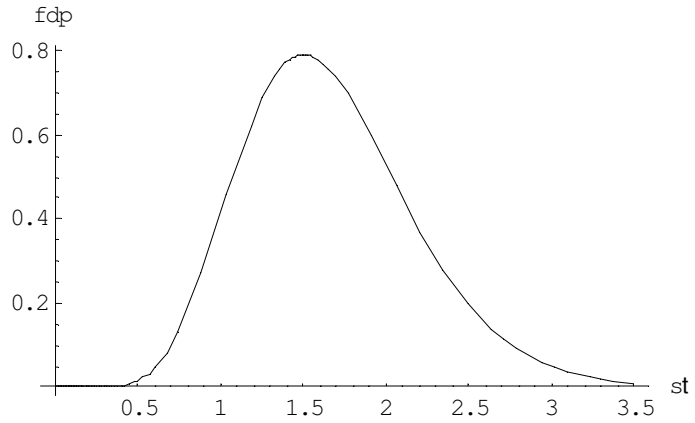
$$
\begin{aligned}
E_{\mathbb{Q}}(S_\tau) &= \frac{\beta_1}{\beta_2} \\
&= \omega \\
&= s_0 \exp\left(r_f \tau\right)
\end{aligned}
\quad \text{e} \quad
\begin{aligned}
Var_{\mathbb{Q}}(S_\tau) &= \frac{\beta_1}{\beta_2^2} \\
&= \sigma^2
\end{aligned}
$$

Conforme Gulko (1998, p. 3-13:3-14) a distribuição gama é mais flexível que a lognormal empregada no modelo de Black-Scholes-Merton. A gama é capaz de descrever a participação acionária nos vários estágios do ciclo de vida da corporação (semelhantemente a uma progressão biológica): (i) oferta pública inicial da ação (a distribuição exponencial expressa valor baixo da ação com alta volatilidade), (ii) maturidade e nome estável da ação (a distribuição gama assemelha-se a lognormal e reflete que o retorno esperado é maior que a volatilidade) e (iii) decadência da ação [distribuição semelhante ao estágio (i)].

A densidade $q\left(s_\tau = 0 \mid \beta_1;\beta_2\right) = 0$ se $\beta_1 > 1$. Neste caso, a gama e a lognormal são qualitativamente similares: ambas têm origem no zero e têm a forma de sino. A forma de sino descreve adequadamente, conforme Gulko (1998, p. 3-13), os valores aleatórios dos índices de ações (e.g. Dow Jones Industrial e S&P 500) e empresas estáveis e maduras (e.g. General Electric Co. e Wells Fargo Bank). O valor esperado do preço da ação $E_{\mathbb{Q}}(S_\tau)$ destas ações excede a sua volatilidade $\sqrt{Var_{\mathbb{Q}}(S_\tau)}$, isto é, $E_{\mathbb{Q}}(S_\tau) > \sqrt{Var_{\mathbb{Q}}(S_\tau)}$.

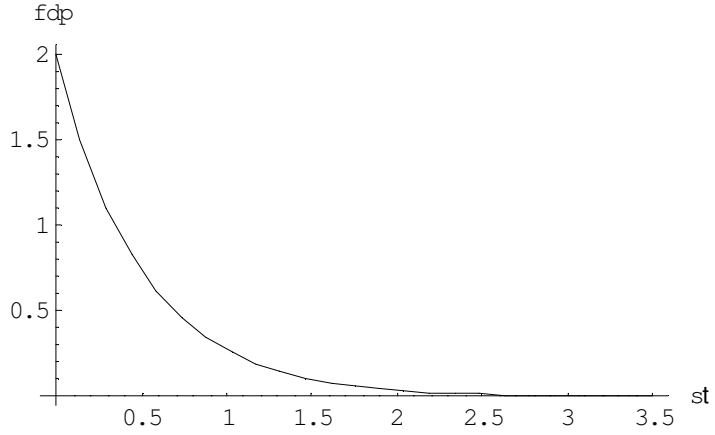
O gráfico a seguir representa uma função de densidade gama $q(s_\tau \mid \beta_1 = 10; \beta_2 = 6)$. Desta forma, a esperança do preço é $E_{\mathbb{Q}}(S_\tau) = \frac{10}{6} \approx 1{,}67$ e a volatilidade do preço é $\sqrt{Var_{\mathbb{Q}}(S_\tau)} = \sqrt{\frac{10}{36}} \approx 0{,}53$.

Figura 22: Gráfico da função de densidade de probabilidade gama(10;6)



A densidade $q(s_\tau = 0 \mid \beta_1; \beta_2) > 0$ se $0 < \beta_1 \leq 1$. Neste caso, a gama e a exponencial são qualitativamente similares: ambas possuem a massa de probabilidade concentrada próxima do zero e a esperança do preço é aproximadamente igual à volatilidade do preço. De fato quando a gama $q(s_\tau \mid \beta_1 = 1; \beta_2 = 2)$ coincide com a exponencial, $E_{\mathbb{Q}}(S_\tau) = \sqrt{Var_{\mathbb{Q}}(S_\tau)} = \frac{1}{2}$. Esta função de densidade descreve adequadamente, conforme Gulko (1998, p. 3-14), ações de baixo valor e alta volatilidade. Estas são freqüentemente as ações de empresas no início ou no final de seu ciclo de vida.

Figura 23: Gráfico da função de densidade de probabilidade gama(1;2)



3.10 *Determinação do preço único é livre de arbitragem da opção de compra européia simples ordinária sobre uma ação*

A fórmula exata encontrada por Gulko (1998, p. 4-18:4-19) para a determinação do preço único e livre de arbitragem para a opção européia simples ordinária sobre a ação é:

$$
\begin{aligned}
c_0 &= s_0 \overline{\mathbb{Q}}\big(s_\tau \mid \beta_1 + 1; \beta_2; k_\tau\big) - k_\tau \exp\big(-r_f \tau\big) \overline{\mathbb{Q}}\big(s_\tau \mid \beta_1; \beta_2; k_\tau\big) \\
\overline{\mathbb{Q}}\big(s_\tau \mid \beta_1; \beta_2; k_\tau\big) &= 1 - \mathbb{Q}\big(s_\tau \mid \beta_1; \beta_2; k_\tau\big) \\
\mathbb{Q}\big(s_\tau \mid \beta_1; \beta_2; k_\tau\big) &= \int_0^{k_\tau} q\big(s_\tau \mid \beta_1; \beta_2\big) ds_\tau
\end{aligned}
$$

Note-se que $\overline{\mathbb{Q}}(s_\tau \mid \beta_1; \beta_2; k_\tau)$ é a probabilidade de resultar em exercício a opção de compra.

4 Estimação da volatilidade implícita

"*A world is risk-neutral when the expected return on all assets is the risk-free rate of interest. What makes the risk-neutral world unique is that it places all assets on a level playing field. The question that arises then is: what determines the price of an asset? The answer is volatility.*"

Yang (1997, p. 4)

4.1 Introdução

Conforme Mendes & Duarte Jr. (1998, p. 63) volatilidade é hoje um dos conceitos mais importantes em Finanças. Na página 82, afirmam que, em termos de desenvolvimento em modelagem matemática, as instituições financeiras brasileiras estão em um estágio muito abaixo das européias, japonesas e norte-americanas, como presenciou um dos autores ao trabalhar nestes mercados mais desenvolvidos. Entender e melhorar os estimadores e modelos estatísticos usados para estimação da volatilidade de ativos e índices é uma necessidade para o desenvolvimento do mercado de ações brasileiro, da área de controle de mercado de carteiras de investimentos etc.

A volatilidade implícita difere da volatilidade de preço, conforme Mendes & Duarte Jr. (1998, p. 66). A volatilidade de preço usa essencialmente estimadores estatísticos diretamente sobre os preços observados do valor mobiliário subjacente. A volatilidade implícita usa os preços da opções sobre os valores mobiliários subjacentes e o modelo de determinação de preço da opção.

Conforme Mendes & Duarte Jr. (1998, p. 74) o problema da estimação da volatilidade é basicamente o de prevê-la na data de exercício da opção. Portanto, a volatilidade implícita pode ser entendida como uma previsão para a volatilidade de ativos e índices de mercado, de acordo com as expectativas do mercado de opções. Na página 70, afirma que um problema sério é o de combinar diferentes volatilidades implícitas em um único valor. Este é um problema em aberto.

Pela segunda propriedade empírica de Gulko (1998) e considerando a hipótese do mundo com preço da incerteza neutralizado, a volatilidade do preço do valor

mobiliário subjacente no mundo com risco é aproximadamente igual à do mundo com preço da incerteza neutralizado.

No entanto, a volatilidade do modelo é desvio-padrão do preço do valor mobiliário subjacente na data de exercício. Logo, ela é o único parâmetro não observável no mercado do modelo.

Uma das formas de estimar a volatilidade é utilizar a fórmula de determinação do preço da opção, considerando conhecido o preço da opção (por meio do preço de mercado) e considerar a volatilidade como incógnita da equação.

4.2 Exemplo paradigmático da estimação da volatilidade implícita do preço da ação Rctb RPN

Por exemplo, para uma opção de compra do mercado brasileiro RTCBJ12 sobre a ação Rctb RPN - uma das ações mais negociadas no mercado de capitais - de um determinado dia e horário de outubro 1999 tem-se:

- Preço à vista da ação: R$ 150,20;
- Preço de exercício: R$ 140,00;
- Taxa certa efetiva anual (SELIC/outubro de 1999): 0,195;
- Preço da opção de compra: R$ 12,80;
- Prazo: 21 dias úteis que corresponde a 1/12 do ano (período).

A taxa certa nominal anual obtém-se pela expressão, conforme Lemgruber (1992) e Luenberger (1998):

$$\begin{aligned} r_f &= \ln(1 + r_f^{\text{efetiva}}) \\ &= \ln(1{,}195) \\ &= 0{,}1781 \end{aligned}$$

Para a ação com o preço como variável de estado o modelo entrópico é o gama, isto é:

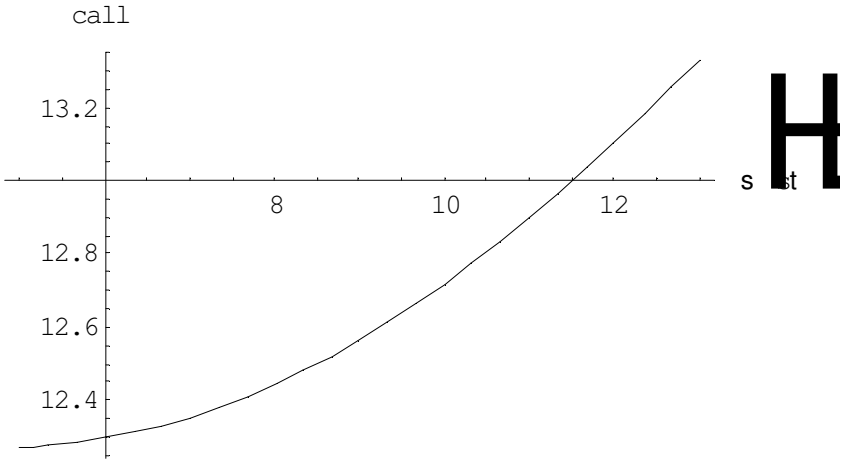
$$12,80 = 150,2 * \overline{\mathbb{Q}}\left(s_\tau \mid \beta_1 + 1; \beta_2; 140\right) - 140 * \exp\left(-0,1781 * \frac{1}{12}\right)\overline{\mathbb{Q}}\left(s_\tau \mid \beta_1; \beta_2; 140\right)$$

$$\beta_1 = \frac{\omega^2}{\sigma^2_{\mathbb{Q}}(S_\tau)}$$

$$\beta_2 = \frac{\omega}{\sigma^2_{\mathbb{Q}}(S_\tau)}$$

$$\omega = 150,2 * \exp\left(0,1781 * \frac{1}{12}\right)$$

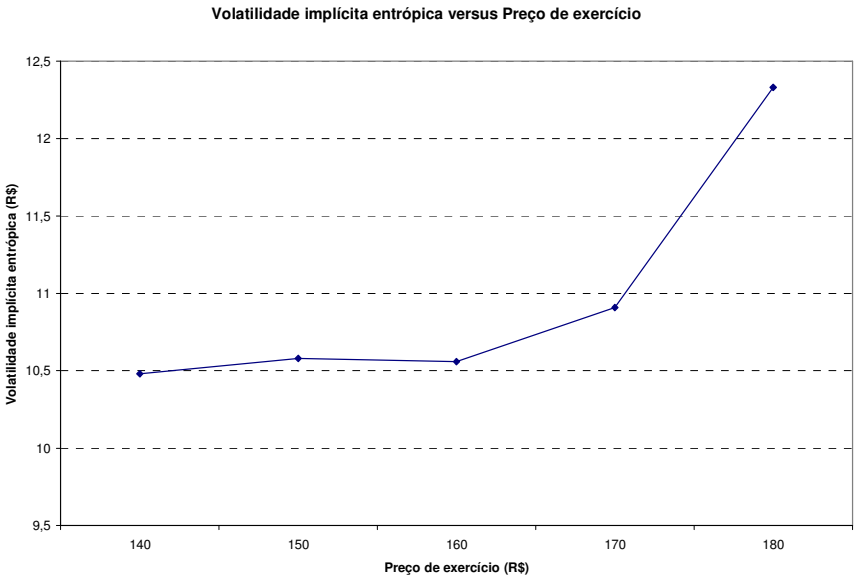
Figura 24: Preço racional da opção de compra *versus* a volatilidade implícita entrópica do preço da ação



Portanto, a volatilidade implícita entrópica é aproximadamente igual a 10,48 reais no prazo de 21 dias úteis para esta ação.

No entanto, o efeito chamado "sorriso da volatilidade" também ocorre. Isto é, ao variar apenas o preço de exercício, a volatilidade implícita varia, assemelhando-se a um sorriso.

Figura 25: Volatilidade implícita entrópica do preço da ação *versus* o preço de exercício



Quadro 7: Volatilidade implícita entrópica do preço da ação Rctb RPN para 21 dias úteis

Nome da opção	s_0 (R$)	k_τ (R$)	r_f (taxa de juros nominal composta continua mente ao ano)	τ (prazo em ano)	Período (ano)	Preço de mercado da opção $= c_0$ (R$)	$\sigma_{\mathbb{Q}}(S_\tau)$ Volatili-dade implícita entrópica (R$)
RTCBJ12	150,20	140	0,1781	1/12	1	12,80	10,48
RTCBJ14	150,20	150	0,1781	1/12	1	5,45	10,58
RTCBJ16	150,20	160	0,1781	1/12	1	1,50	10,56
RTCBJ18	150,20	170	0,1781	1/12	1	0,29	10,91
RTCBJ20	150,20	180	0,1781	1/12	1	0,08	12,33

Pela segunda propriedade empírica de Gulko (1998) tem-se ainda que:

$$\mathrm{var}_{\mathbb{Q}}(S_\tau) \approx \mathrm{var}_{\mathbb{P}}(S_\tau).$$

Desta forma, o valor da volatilidade do preço é aproximadamente igual à volatilidade implícita entrópica do preço do valor mobiliário subjacente.

No modelo de Black-Scholes-Merton, pode-se obter diretamente a volatilidade (desvio-padrão da variável de estado na data de exercício do preço) do ação por meio da seguinte fórmula [Luenberger (1998, p. 310)]:

$$\begin{aligned}\sigma_{\mathbb{Q}}(S_\tau) &= \sqrt{\mathrm{var}_{\mathbb{Q}}(S_\tau)} \\ &= s_0 \exp\left(r_f \tau\right)\sqrt{\left[\exp\left(\sigma_{\mathbb{Q}}^2 \tau\right) - 1\right]} \\ &\approx s_0 \exp\left(r_f \tau\right)\sigma_{\mathbb{Q}}\sqrt{\tau}\end{aligned}$$

sendo σ a volatilidade a termo.

A volatilidade a termo está relacionada com a volatilidade da taxa de retorno composta continuamente da seguinte maneira [Baxter & Rennie (1998)]:

$$\sigma_{\mathbb{Q}} = \sqrt{\frac{\mathrm{var}_{\mathbb{Q}}\left[\ln\left(\frac{S_\tau}{s_0}\right)\right]}{\tau}}$$

$$\sqrt{\mathrm{var}_{\mathbb{Q}}\left[\ln\left(\frac{S_\tau}{s_0}\right)\right]} = \sigma_{\mathbb{Q}}\sqrt{\tau}$$

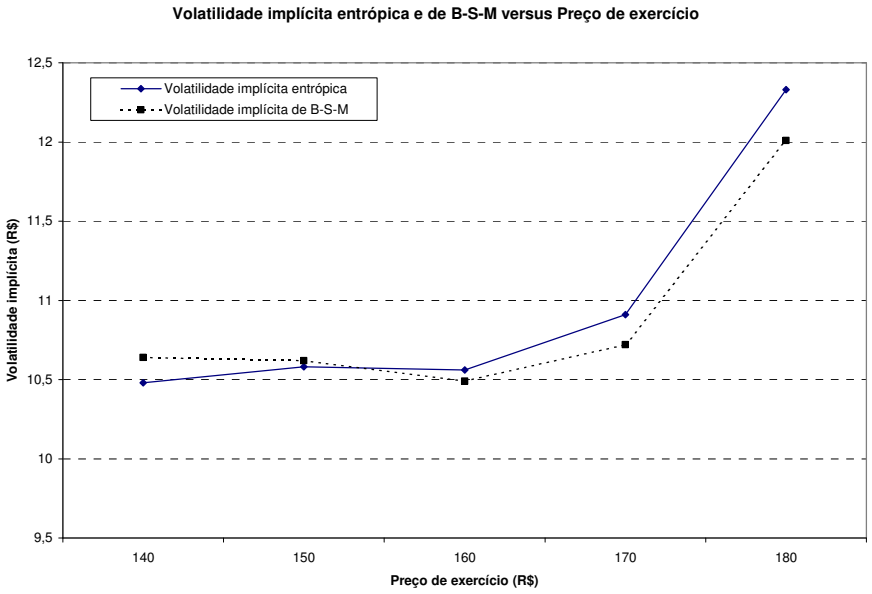
Note-se que a volatilidade da taxa de retorno é adimensional. A unidade da volatilidade a termo, porém, é o inverso da raiz quadrada do período, conforme Wilmott *et al.* (1997).

Se o modelo de Black-Scholes-Merton é utilizado, as volatilidades implícitas são:

Quadro 8: Volatilidade de Black-Scholes-Merton do preço da ação Rctb RPN para 21 dias úteis

Nome da opção	s_0 (R$)	k_τ (R$)	r_f (taxa de juros nominal composta continua mente ao ano)	τ (prazo em ano)	Período (ano)	Preço de mercado da opção $= c_0$ (R$)	$\sigma_{\mathbb{Q}}(S_\tau)$ Volatili-dade implícita por B-S-M (R$)
RTCBJ12	150,20	140	0,1781	1/12	1	12,80	10,64
RTCBJ14	150,20	150	0,1781	1/12	1	5,45	10,62
RTCBJ16	150,20	160	0,1781	1/12	1	1,50	10,49
RTCBJ18	150,20	170	0,1781	1/12	1	0,29	10,72
RTCBJ20	150,20	180	0,1781	1/12	1	0,08	12,01

Figura 26: Volatilidade implícita entrópica e de B-S-M do preço da ação *versus* o preço de exercício



A partir dos resultados anteriores, pode-se concluir que as volatilidades implícitas da variável de estado preço da ação são aproximadamente iguais.

É importante lembrar que o modelo de B-S-M ou lognormal é também equivalente a um modelo entrópico cuja variável de estado é a taxa de retorno composta continuamente.

Desta forma, este é um forte indício da irrelevância da escolha da variável de estado (preço ou taxa de retorno) para a determinação da volatilidade implícita do preço.

Sugere-se que seja utilizado um bom *software* que realize cálculo simbólico, como o Mathematica 3.0 ou superior.

Gulko (1998 e 1998a) testa o modelo beta para opção sobre *bond* e conclui que é superior a outros dois modelos concorrentes, baseados na normal e na lognormal.

4.3 Valor em risco entrópico

Conforme Duarte Jr. (1998, p. 10) o valor em risco é hoje a medida de risco de mercado absoluta mais usada nos mercados financeiros internacionais. O valor em risco de uma certa quantidade $n \in \mathbb{R}_*^+$ de valores mobiliários cujo preço unitário hoje é $s(t_0)$ é uma medida de quanto esta riqueza $ns(t_0)$ poderá depreciar durante um certo horizonte de tempo $\tau = t_\tau - t_0$, com certa probabilidade $\alpha \in \left]0;\frac{1}{2}\right[$. Uma definição formal (do ponto de vista probabilístico) de valor em risco é: se $S(t_\tau)$ é uma variável aleatória usada para representar o preço de um valor mobiliário com incerteza num período de tempo τ e α é uma probabilidade associada a valores baixos incomuns ou raros de preço do ativo [Berkeley (1965, p. 105)], então o valor absoluto da solução v para a equação seguinte:

$$\Pr\left(S(t_\tau) \leq \frac{v}{n}\right) = \alpha$$

é o valor em risco do valor mobiliário com incerteza, segundo Duarte Jr. (1998, p. 10).

4.4 Preço mínimo atual

O valor presente pela taxa de juros certa de v, isto é, $\exp(-r_f\tau)\frac{v}{n}$, é o valor que aqui se chamará de **preço mínimo atual** (*pma*) do valor mobiliário com probabilidade α no horizonte τ à taxa de juros certa r_f, isto é:

$$
\begin{aligned}
s_{\min}\left(\alpha;\tau;r_f;n\right) &= vp\left(\frac{v}{n} \mid r_f;\tau\right) \\
&= \exp(-r_f\tau)\frac{v}{n}
\end{aligned}
$$

4.5 Preço mínimo atual entrópico

O preço mínimo atual entrópico (*pmae*), portanto, consiste no preço mínimo atual na medida de probabilidade neutralizadora do preço da incerteza:

$$
\Pr\left(S^{\mathbb{Q}}(t_\tau) \leq \frac{v}{n}\right) = \alpha
$$

4.6 Perda máxima atual entrópica

Outro valor de interesse é a diferença entre o preço à vista e o *pmae* do valor mobiliário, que será chamado de **perda máxima atual entrópica** (*pmaxe*):

$$
\begin{aligned}
\nabla s &= s(t_0) - s_{\min}\left(\alpha;\tau;r_f;n\right) \\
&= s(t_0) - \exp(-r_f\tau)\frac{v}{n} \\
&> 0
\end{aligned}
$$

Normalmente usam-se os quantis de $\alpha = 0{,}05$, $0{,}03$ e $0{,}01$.

Retomando o exemplo anterior e definindo a volatilidade do preço de uma ação ($n = 1$) em 11 reais, tem-se que os *pmae* de 0,05, 0,03 e 0,01 no horizonte de 21 dias são R$132,83, R$130,49 e R$126,14. Isto é, as respectivas *pmaxe* nestes níveis de probabilidade são de R$17,37, R$19,71 e R$24,06.

4.7 Perda máxima atual entrópica padronizada

O *pmaxe* pode ser padronizado pela volatilidade implícita entrópica, originando a *pmaxe* padronizada:

$$
\begin{aligned}
\zeta &= \frac{\nabla s}{\sqrt{\mathrm{var}_{\mathbb{Q}}\left[S(t_\tau)\right]}} \\
&= \frac{s(t_0) - \exp(-r_f\tau)\frac{v}{n}}{\sqrt{\mathrm{var}_{\mathbb{Q}}\left[S(t_\tau)\right]}}
\end{aligned}
$$

Para Hull (1998, p. 338-9), o valor em risco (VaR) responde à seguinte pergunta: "Qual o máximo valor v que pode ser perdido com p% de chance nos próximos n dias?" Este conceito de VaR corresponde mais propriamente ao de *pmaxe*.

4.8 Comparação da medida de probabilidade neutralizadora do preço da incerteza da taxa de retorno (modelo entrópico lognormal) com a do preço da ação (modelo entrópico gama)

A função de densidade de probabilidade neutralizadora do preço da incerteza entrópica da variável de estado taxa de retorno $R_\tau = \ln\left(\frac{S_\tau}{s_0}\right)$ da ação ou do modelo de Black-Scholes-Merton é normal com parâmetros:

$$R_\tau \sim \text{normal}\left[\ln(s_0) + \left(r_f - \frac{\sigma_{\mathbb{Q}}^2}{2}\right)\tau; \sigma_{\mathbb{Q}}^2 \tau\right]$$

sendo a volatilidade a termo igual a:

$$\begin{aligned}
\sigma_{\mathbb{Q}} &= \sqrt{\frac{\ln\left[\exp\left(-2r_f\tau\right)\frac{\sigma_{\mathbb{Q}}^2(S_\tau)}{s_0^2} + 1\right]}{\tau}} \\
&\approx \frac{\exp\left(-r_f\tau\right)\sigma_{\mathbb{Q}}(S_\tau)}{s_0\sqrt{\tau}} \\
&= \frac{\exp(-0{,}1781 * (1/12)) * 11}{150{,}2 * \sqrt{(1/12)}} \\
&= 0{,}25
\end{aligned}$$

Fazendo $S_\tau = s_0 \exp(R_\tau)$, tem-se que a variável aleatória preço da ação na data de exercício possui distribuição lognormal com os seguintes parâmetros:

$$S_\tau \sim \text{lognormal}\left[s_0 \exp(r_f\tau); s_0^2 \exp(2r_f\tau)\left[\exp(\sigma_{\mathbb{Q}}^2\tau) - 1\right]\right].$$

A função de densidade de probabilidade neutralizadora do preço da incerteza de máxima entropia lognormal é:

$$\begin{aligned}
q(s_\tau) &= \frac{1}{s_\tau \sigma_{\mathbb{Q}}^2 \sqrt{2\pi\tau}} \exp\left\{-\frac{1}{2}\left[\frac{\ln\left(\frac{s_\tau}{s_0}\right) - \left(r_f - \frac{\sigma_{\mathbb{Q}}^2}{2}\right)\tau}{\sigma_{\mathbb{Q}}^2\sqrt{\tau}}\right]^2\right\} \\
&= \frac{1}{s_\tau\sqrt{\ln\left[\exp\left(-2r_f\tau\right)\frac{\sigma_{\mathbb{Q}}^2(S_\tau)}{s_0^2} + 1\right]}\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left[\frac{\ln\left(\frac{s_\tau}{s_0}\right) - \left(r_f\tau - \frac{\ln\left[\exp\left(-2r_f\tau\right)\frac{\sigma_{\mathbb{Q}}^2(S_\tau)}{s_0^2} + 1\right]}{2}\right)}{\sqrt{\ln\left[\exp\left(-2r_f\tau\right)\frac{\sigma_{\mathbb{Q}}^2(S_\tau)}{s_0^2} + 1\right]}}\right]^2\right\}
\end{aligned}$$

A função de densidade de probabilidade neutralizadora do preço da incerteza de máxima entropia gama é:

$$S_\tau \sim \text{gama}\left[\beta_1;\beta_2^{-1}\right]$$

$$q\left(s_\tau \mid \beta_1;\beta_2\right) = \frac{\beta_2^{\beta_1}}{\Gamma(\beta_1)} s_\tau^{\beta_1-1} \exp\left(-\beta_2 s_\tau\right)$$

$$\Gamma(\beta_1) = \int_0^\infty s_\tau^{\beta_1-1} \exp\left(s_\tau\right) ds_\tau,\ \beta_1 > 0$$

$$\beta_1 = \frac{s_0^2 \exp\left(2r_f\tau\right)}{\sigma_{\mathbb{Q}}^2\left(S_\tau\right)}$$

$$\beta_2 = \frac{s_0 \exp\left(r_f\tau\right)}{\sigma_{\mathbb{Q}}^2\left(S_\tau\right)}$$

Isto é:

$$q\left(s_\tau\right) = \frac{\left[\dfrac{s_0 \exp\left(r_f\tau\right)}{\sigma_{\mathbb{Q}}^2\left(S_\tau\right)}\right]^{\left[\frac{s_0^2 \exp\left(2r_f\tau\right)}{\sigma_{\mathbb{Q}}^2\left(S_\tau\right)}\right]}}{\displaystyle\int_0^\infty s_\tau^{\left[\frac{s_0^2 \exp\left(2r_f\tau\right)}{\sigma_{\mathbb{Q}}^2\left(S_\tau\right)}-1\right]} \exp\left(s_\tau\right) ds_\tau} s_\tau^{\left[\frac{s_0^2 \exp\left(2r_f\tau\right)}{\sigma_{\mathbb{Q}}^2\left(S_\tau\right)}-1\right]} \exp\left[-\frac{s_0 \exp\left(r_f\tau\right)}{\sigma_{\mathbb{Q}}^2\left(S_\tau\right)} s_\tau\right]$$

Note-se que as esperanças da variável aleatória preço da ação na data de exercício são exatamente iguais para nas duas distribuições, isto é:

$$E_{\mathbb{Q}}\left[S_\tau\right] = s_0 \exp\left(r_f\tau\right).$$

No entanto, as volatilidades da variável aleatória preço da ação na data de exercício são aproximadamente iguais para nas duas distribuições e são dependentes do modelo do preço da opção.

Conhecidas as funções de densidade de probabilidade pode calcular os quantis de 0,01, 0,03 e 0,05.

	Perda máxima atual entrópica (*pmaxe*) (R$)		
Modelo entrópico do preço da ação	*1%*	*3%*	*5%*
Gama	24,06	19,71	17,37
Lognormal	23,51	19,38	17,13

Portanto, não há indícios, neste caso, de discrepâncias entre as *pmaxe* dos modelos, pois as funções de densidade são aproximadamente iguais. Esse fato pode ser observado no gráfico a seguir.

Figura 27: Função de densidade de probabilidade neutralizadora do preço da incerteza entrópica do preço da ação para a variável de estado taxa de retorno (lognormal - representada pela linha contínua) versus a da variável de estado preço da ação (gama - representada pela linha tracejada) com volatilidade implícita entrópica igual a 11 reais

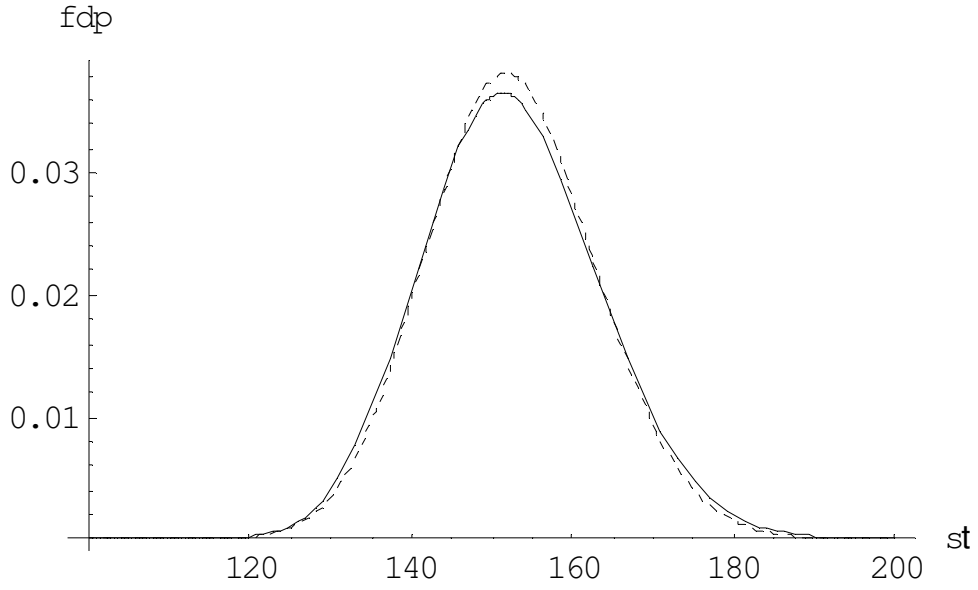
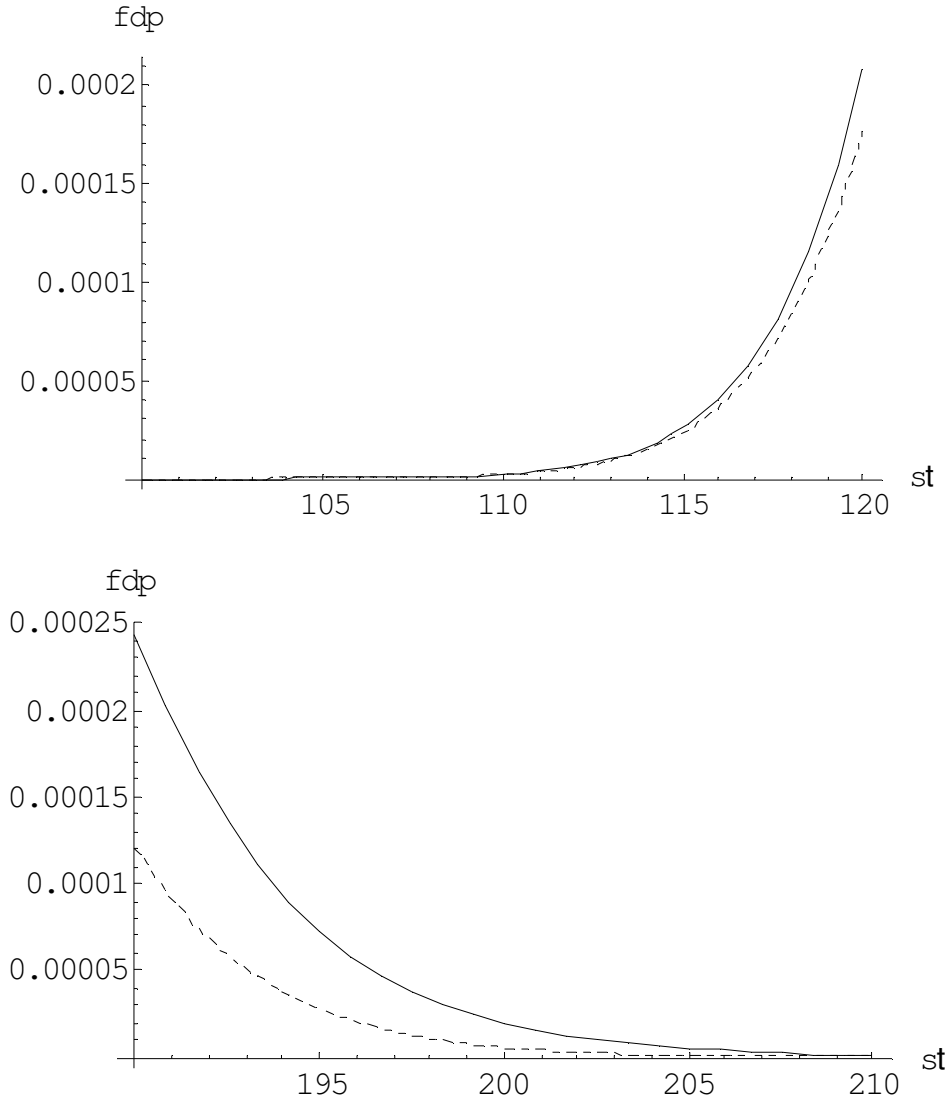


Figura 28: Caudas ampliadas da função de densidade de probabilidade neutralizadora do preço da incerteza entrópica do preço da ação para a variável de estado taxa de retorno (lognormal - representada pela linha contínua) e da variável de estado preço da ação (gama - representada pela linha tracejada) com volatilidade implícita entrópica igual a 11 reais



5 Conclusões, limitações, recomendações e temas para pesquisas

"É preciso coragem, Teeteto, sempre que se pode avançar, ainda que seja um pouquinho de cada vez. Quem desanimasse num caso desses, ante a escassez de resultados, como se comportaria em conjunturas mais sérias, em que não assinalasse nenhum avanço ou mesmo fosse obrigado a recuar? Nesse passo, conforme diz o provérbio, um tipo assim nunca tomará cidade alguma. Porém, agora, amigo, como superamos a dificuldade que formulaste, caiu em nosso poder a principal trincheira; tudo o mais será fácil e carece de importância."

Platão. 1980. *Diálogos*: *Sofista* ou *Do Ser*, 261bc.
Tradução de Carlos Alberto Nunes. Universidade Federal do Pará.

5.1 Conclusão da tese

A medida única de probabilidade neutralizadora do preço da incerteza $q(s_\tau)$ é obtida por meio da maximização da entropia diferencial da expectativa de mercado no mundo entrópico $h_{\mathbb{E}}(X_{\mathbb{E}}) = -\int_{\xi} f(s_\tau)\ln[f(s_\tau)]ds_\tau$, perfeito (esta condição garante a existência e a unicidade da taxa de juros certa r_f) e em equilíbrio competitivo (esta condição garante a existência e a unicidade preço à vista do valor mobiliário subjacente s_0) da variável aleatória preço S_τ do valor mobiliário subjacente na data de exercício τ sob o conjunto informacional $I = \left\{\xi; Var_{\mathbb{E}}(S_\tau) \approx Var_{\mathbb{P}}(S_\tau); E_{\mathbb{E}}(S_\tau) = \exp\left(r_f\tau\right)s_0\right\}$ no qual $\xi = (0;1)$ é o domínio do *bond*, $\xi = (0;\infty)$ é o domínio da ação e $\mathbb{P}$ é a medida de probabilidade real ou subjetiva do preço do valor mobiliário, isto é, $q(s_\tau) = \arg\max_{f(s_\tau)} h_{\mathbb{E}}(X_{\mathbb{E}} \mid I)$.

Como na condição de arbitragem o mundo com preço da incerteza neutralizado possui a propriedade da linearidade de preços, então $E_{\mathbb{Q}}\left(\frac{C_\tau}{c_0}\right) = \exp\left(r_f\tau\right)$ sendo $C_\tau = \max(0; S_\tau - k_\tau)$. Desta forma, o preço racional da opção de compra européia simples ordinária é $c_0 = \exp\left(-r_f\tau\right)E_{\mathbb{Q}}[\max(0; S_\tau - k_\tau)]$. Portanto, existe e é único o preço da opção de compra européia simples ordinária.

O mundo entrópico originou-se de um mundo sem arbitragem, com linearidade de preços, com preço da incerteza neutralizado, martingale e, finalmente, de um mundo informacionalmente eficiente. A aplicação do Princípio de Máxima Entropia em combinação com o Sistema de Distribuições de Pearson constitui

uma maneira alternativa de obter a medida de probabilidade neutralizadora do preço da incerteza no mundo entrópico. Desta forma, evidencia-se o fato da não necessidade da hipótese de mercado completo e eficiente e da medida de probabilidade real ou subjetiva do valor mobiliário subjacente da data de exercício para a determinação do preço racional da opção.

Quadro 9: Resumo das condições para necessárias para a obtenção dos modelos de Black-Scholes-Merton e entrópico

	Modelo Black-Scholes-Merton para ação	Modelo entrópico para ação
Variável de estado	Taxa de retorno da ação	Taxa de retorno ou preço da ação
Hipótese de mercado perfeito e do banco ideal constante	Necessárias	Necessárias
Hipótese de mercado em equilíbrio competitivo	Necessária	Necessária
Hipótese de mercado completo	Necessária	Desnecessária
Hipótese de mercado eficiente	Necessária	Desnecessária
Hipótese do processo estocástico que governa o movimento do preço do valor mobiliário subjacente	Necessária: movimento browniano geométrico	Desnecessária
Hipótese de mundo com preço da incerteza neutralizado	Desnecessária	Necessária
Hipótese da atitude do investidor em relação à incerteza	Desnecessária	Desnecessária
Método de estimação da medida de probabilidade	Teorema de Cameron-Martin-Girsanov	Princípio de Máxima Entropia ou de Mínima

	Modelo Black-Scholes-Merton para ação	Modelo entrópico para ação
neutralizadora do preço da incerteza		Divergência de Kullback-Leibler e Sistema de Distribuições de Pearson

5.2 Teste do modelo entrópico

O modelo entrópico para *bonds* tem sido testado por Gulko (1998 e 1998a) e tem se mostrado superior ao de Black-Scholes-Merton.

O modelo entrópico para ações é qualitativamente superior ao clássico modelo de Black-Scholes-Merton, no entanto ele necessita ser mais testado, apesar das fortes evidências teóricas a seu favor.

5.2.1 Os prós e contras da modelagem Black-Scholes-Merton versus a entrópica para ação

Os prós e contras da modelagem Black-Scholes-Merton *versus* a entrópica para ação são:

5.2.1.1 Prós da modelagem de Black-Scholes-Merton para o objeto subjacente ação:

1. A fórmula final é de fácil implementação computacional.
2. Todos os seus parâmetros ou são estimáveis (volatilidade) ou são observáveis (taxa de juros certa, preço à vista do ativo subjacente, preço de exercício e prazo).
3. O modelo é comunicável ao mercado (aos práticos) de maneira simples e prática por meio, principalmente da modelagem de árvore binária recombinante. Esta abordagem evita o uso de cálculo estocástico para determinar o preço de derivativos e ainda permite a implementação computacional com relativa facilidade.

4. Historicamente ela foi a primeira fórmula a considerar a existência da medida de probabilidade neutralizadora do preço da incerteza em vez da distribuição de probabilidade real ou subjetiva do ativo subjacente.

5. A análise da anatomia da fórmula de B-S-M permite o entendimento mais preciso da formação do preço racional de uma opção européia simples ordinária sobre uma ação.

5.2.1.2 Contras da modelagem de Black-Scholes-Merton para o objeto subjacente ação:

1. Esta modelagem parte da suposição de que o mercado é eficiente. Logo, o processo estocástico que descreve o movimento do preço da ação é o browniano geométrico. Esta modelagem culmina com a distribuição lognormal para o preço do valor mobiliário subjacente na data de exercício. No entanto, esta suposição de certa forma impõe uma medida de probabilidade real ou subjetiva à variável aleatória terminal do processo. Desta distribuição do preço na data de exercício é extraída a medida de probabilidade neutralizadora do preço da incerteza por meio de uma mudança de medida para transformar o processo estocástico num martingale. Esta medida é também uma lognormal com o outros parâmetros. Ora, sabe-se que a medida de probabilidade neutralizadora do preço da incerteza independe da medida de probabilidade real ou subjetiva do preço do ativo objeto. Logo, o princípio da parcimônia foi violado.

2. A lognormal segundo no modelo entrópico de Gulko (1998) é aplicável ao preço de ações de organizações estabelecidas e maduras no mercado. A gama é uma distribuição que modelo melhor as fases do ciclo de vida de uma ação. É possível construir uma gama cuja esperança seja um martingale.

3. O modelo de Black-Scholes-Merton parte da hipótese de que o *hedging* perfeito uniperíodo ou multiperíodo (estratégia de carteira replicante e autofinanciável) é possível. Isto implica a existência e unicidade da medida de probabilidade neutralizadora do preço da incerteza que, por sua vez,

implica existência de mercado completo e livre de arbitragem. Como conseqüência, obtém-se o preço único do derivativo. Conforme Cox & Rubinstein (1985), a existência de mercado completo é uma hipótese não realista. Do ponto de vista de Gulko (1998), a estratégia de *hedging* deve ser conseqüência do modelo para determinação do preço de derivativo, e não hipótese para sua obtenção.

4. Para obtenção da fórmula de B-S-M, há a suposição de que a taxa de juros de curto prazo certa durante a vida da opção e a volatilidade instantânea são constantes. Estas também são suposições fortes sobre os processos estocásticos que governam os movimentos destas potenciais fontes de risco.

5. A hipótese de haver um determinado processo estocástico do preço do ativo subjacente introduz a necessidade do entendimento de processos estocásticos (martingale etc.) e do cálculo estocástico (Lema de Ito etc.). Estas teorias são muito complicadas para um problema tão simples de formação de preço num mercado perfeito e em equilíbrio competitivo.

6. Outra complicação desta modelagem é iniciar este processo com a equação diferencial estocástica da taxa de retorno instantânea do valor mobiliário subjacente. Isto inicia uma dificuldade que é o entendimento da unidade de medida da volatilidade da taxa de retorno da ação (volatilidade a termo), que é o tempo elevado a menos meio, i.e., $t^{-\frac{1}{2}}$, conforme Wilmott *et al.* (1987).

5.2.1.3 Prós da modelagem entrópico para o objeto subjacente ação:

1. A abordagem entrópica utiliza o princípio de máxima entropia ou de mínima divergência de Kullback-Leibler da Teoria da Informação para determinar a distribuição de probabilidade que melhor aproveita todas as informações disponíveis e apenas elas. Esta é uma forma não viesada e honesta de estimar a distribuição de probabilidade. Este método é útil para a resolução de problemas com indeterminação (*ill-posed*). Estes problemas não são solucionáveis, geralmente, pelo princípio de máxima verossimilhança.

2. O princípio de máxima entropia resolve diretamente o problema central da existência e a unidade da medida de probabilidade neutralizadora do preço da incerteza a partir do conjunto informacional (informações disponíveis dos ativos básicos contidas nas restrições), como o domínio do ativo objeto (zero a mais infinito), condição de não-arbitragem (a expectativa do preço do ativo subjacente na data de exercício é o valor futuro pela taxa certa do preço à vista da ação) e volatilidade (que faz com que a distribuição tenha a forma de sino).

3. Esta abordagem adota o preço como variável de estado do sistema em vez da taxa de retorno do preço do ativo objeto. O foco é o preço do ativo subjacente.

4. Esta abordagem também gera uma fórmula anatomicamente similar à de B-S-M. Portanto, o resultado é computacionalmente prático.

5. A hipótese de mercado eficiente não é necessária, pois, apenas com a suposição de que o mercado está em equilíbrio competitivo e é perfeito, pode-se supor a existência e unicidade dos preços de equilíbrio dos ativos deste mercado.

6. Não há mais a necessidade de supor a possibilidade de *hedging* perfeito. Logo, não há a necessidade de o mercado ser completo para a existência e unicidade da medida de probabilidade neutralizadora do preço da incerteza.

7. O método de máxima entropia garante a existência e unicidade da medida de probabilidade neutralizadora do preço da incerteza, pois a função de entropia é côncava e as restrições, que são basicamente os dois primeiros momentos, formam uma região convexa.

8. O método de máxima entropia permite a fácil extensão para a determinação do preço racional de portfólio de ações e *bond*s (modelo multivariado).

9. O método de máxima entropia não necessita do conhecimento da distribuição de probabilidade real ou subjetiva do preço do ativo subjacente na data de vencimento.

10. Com a eliminação da possibilidade do preço do ativo objeto na data de exercício ser nulo, obtém-se a distribuição gama. Esta distribuição é flexível. Ela pode gerar a exponencial (diretamente) e a lognormal (assintoticamente).

11. A explicação do modelo é simples, pois só é necessário o conhecimento de Teoria da Informação (princípio de máxima entropia), otimização não-linear (método dos multiplicadores de Lagrange) e algumas noções de Microeconomia (formação de preços em mercados competitivos). Isto é, não é mais necessário utilizar processos estocásticos e cálculo estocástico.

12. Há a necessidade de definir com rigor o que é um mundo com preço da incerteza neutralizado (*risk-neutral world*). No mundo com preço da incerteza neutralizado há incerteza que pode ser maximizada pelo princípio de máxima entropia e também há volatilidade do preço do valor mobiliário subjacente que é praticamente, segundo Gulko (1998), igual à do mundo com real.

5.2.1.4 Contras da modelagem entrópica para o objeto subjacente ação:

1. Há relativamente poucos estudos sobre Teoria da Informação aplicada em Economia. Exceções são os trabalhos de Theil (1967 e 71), Golan *et al.* (1996), Yang (1997) e Gulko (1998 e 1998a). No entanto, Yang (1997) critica e faz avanços sobre os resultados de Gulko (1995-6).

2. Parece haver divergências com relação à interpretação das hipóteses que justificam o emprego da função objetivo entropia discreta ou diferencial para a determinação da medida de probabilidade neutralizadora do preço da incerteza pelo método da máxima entropia. Gulko (1998) argumenta que o princípio deriva da hipótese de mercado informacionalmente eficiente, isto é, metade dos investidores a cada instante acham que o preço do ativo base irá subir e metade que irá descer no próximo instante. Yang (1997) evita este

argumento e afirma que apenas a existência do mundo com preço da incerteza neutralizado é suficiente para o emprego do método de máxima entropia. Nesta tese há uma tentativa de integrar essas duas hipóteses.

5.3 Limitações

A extensão da abordagem entrópica para a determinação de preço de opção americana [Yang (1997)], opção *path-dependent* (incluindo as exóticas) e de carteiras de valores mobiliários (modelo multivariado). Observe-se que em nenhum modelo foi considerado o dividendo não-constante da ação ou cupom para o *bond*.

5.4 Pesquisas futuras

A seguir são propostas algumas questões e conjecturas para promover futuras pesquisas:

- Até que ponto a hipótese de mercado perfeito pode ser relaxada de modo que ainda continue sendo possível estimar a medida única de probabilidade neutralizadora do preço da incerteza? Conjectura: a resposta deve estar em Yu (1998).
- Qual o processo estocástico browniano que gera uma gama ou uma beta? Conjectura: ele é um tipo especial de movimento browniano diferente do geométrico.
- Qual é a distribuição que maximiza a taxa de entropia de um processo estocástico markoviano sujeito a um conjunto informacional? Conjectura: a resposta a esta pergunta deve estar relacionada com a resposta da anterior.
- Por que a variância na medida de probabilidade neutralizadora do preço da incerteza é aproximadamente igual à da medida de probabilidade real ou subjetiva do preço do valor mobiliário subjacente? Conjectura: o mundo com preço da incerteza neutralizado ou entrópico é um mundo com a mesma quantidade de incerteza que o mundo real com preço da incerteza, i.e., eles têm a mesma entropia. Então as variâncias do preço futuro devem ser aproximadamente iguais nas duas medidas de probabilidade (real ou subjetiva

e a neutralizadora do preço da incerteza). No caso da distribuição normal, esta conjectura funciona perfeitamente.

- Por que há o sorriso da volatilidade? Conjectura: se a conjectura anterior for verdadeira, deve ser encontrada uma faixa de valores viáveis de volatilidade do preço.
- Qual a diferença entre a determinação do preço racional da opção financeira e da opção real? Conjectura: se o princípio de não-arbitragem for mantido, a diferença não está no método proposto na tese; ela está na não-unicidade do preço à vista do valor ativo subjacente que pode ser causada pela não-existência de equilíbrio competitivo ou simplesmente por ser não observável.
- Até que ponto há a necessidade de exigir a aplicação do princípio de não-arbitragem para a determinação do preço racional da opção real? Conjectura: parece ser necessária a aplicação deste princípio para que o preço possa denominar-se racional.
- Qual a relação entre o ganho informacional relativo da previsão da direção do preço no mundo real com preço da incerteza e o mundo entrópico? Conjectura: em mercados competitivos, os ganhos informacionais relativos da previsão tendem aos do mundo entrópico.
- Como integrar a teoria da decisão com a Teoria da Informação? Conjectura: há uma estreita conexão entre a função de utilidade e a divergência. Portanto, descortina-se um caminho de integração entre Teoria da Informação e teoria Estatística bayesiana da decisão. A tese de Samperi (1998) abre esta possível unificação de abordagens.
- Quais os critérios para a escolha da medida de probabilidade prévia para a aplicação do princípio de mínima divergência? Conjectura: no caso da opção sobre ação uma candidata seria a lognormal e no caso do *bond*, a uniforme. Consultar Newman (1999).
- Conjectura: Se dois valores mobiliários são dependentes, então o conjunto informacional do valor mobiliário subjacente deve ser

$I = \left\{ \xi; E_{\mathbb{E}}(S_\tau) = \exp\left(r_f \tau\right) s_0; \rho\left(S_\tau; S_\tau^{'}\right) = \dfrac{Cov\left(S_\tau; S_\tau^{'}\right)}{\sqrt{Var_{\mathbb{E}}(S_\tau) Var_{\mathbb{E}^{'}}\left(S_\tau^{'}\right)}} \right\}$ sendo o coeficiente de correlação linear de Pearson $\rho\left(S_\tau; S_\tau^{'}\right) \neq 0$, isto é, a informação mútua entre as duas variáveis aleatórias é diferente de zero.

6 Apêndice

6.1 Programação do modelo de Black-Scholes-Merton e entrópico em Visual Basic

Este programa em Visual Basic para o Excel que calcula o preço da opção européia simples ordinária sobre ação pelo modelo de B-S-M e entrópico foi escrito por Fernando Hideki Kato. O modelo entrópico teve que ser programado porque a função de distribuição gama do Excel apresentou problemas.

```
Option Explicit
Public Function BSEuro(ByVal op As String, _
                       ByVal callPut As String, _
                       ByVal S As Double, _
                       ByVal X As Double, _
                       ByVal t As Double, _
                       ByVal sigma As Double, _
                       ByVal r As Double, _
                       ByVal Period As Double _
                      ) As Double
   If UCase(op) <> "PRICE" Or _
      (UCase(callPut) <> "C" And UCase(callPut) <> "P") Or _
      S <= 0 Or _
      X <= 0 Or _
      t <= 0 Or _
      sigma <= 0 Or _
      r < 0 Or _
      Period <= 0 Then
      BSEuro = CVErr(xlErrValue)
      Exit Function
   End If

   On Error GoTo Error0

   Dim tt As Double
   tt = CDbl(t) / Period

   Dim d1 As Double
   d1 = (Log(S / X) + (r + 0.5 * sigma ^ 2) * tt) / (sigma * Sqr(tt))
   Dim d2 As Double
   d2 = d1 - sigma * Sqr(tt)

   Dim Nd1 As Double
   'Nd1 = Application.NormDist(d1, 0, 1, True)
   Nd1 = ALNORM(d1, False)
   Dim Nd2 As Double
   'Nd2 = Application.NormDist(d2, 0, 1, True)
   Nd2 = ALNORM(d2, False)

   If UCase(callPut) = "C" Then
      BSEuro = S * Nd1 - X * Exp(-r * tt) * Nd2
   Else
      BSEuro = X * Exp(-r * tt) * (1 - Nd2) - S * (1 - Nd1)
   End If

   Exit Function
Error0:
   MsgBox "Error in BSEuro." & Err
End Function

Public Function GMEuro(ByVal op As String, _
```

```
                        ByVal callPut As String, _
                        ByVal S As Double, _
                        ByVal X As Double, _
                        ByVal t As Double, _
                        ByVal sigma As Double, _
                        ByVal r As Double, _
                        ByVal Period As Double _
                       ) As Double
   If UCase(op) <> "PRICE" Or _
      (UCase(callPut) <> "C" And UCase(callPut) <> "P") Or _
      S <= 0 Or _
      X <= 0 Or _
      t < 0 Or _
      sigma <= 0 Or _
      r < 0 Or _
      Period <= 0 Then
      GMEuro = CVErr(xlErrValue)
      Exit Function
   End If

   On Error GoTo Error0

   Dim tt As Double
   tt = CDbl(t) / Period

   'Dim lambda As Double
   'lambda = X / sigma / sigma * Sqr(S / X / Exp(-r * tt))
   'S = S * lambda
   'X = X * lambda
   'sigma = sigma * lambda

   Dim u As Double
   u = S / Exp(-r * tt) / (sigma * sigma)
   Dim v As Double
   v = u * u * sigma * sigma

   Dim gammaDistAlfa As Double
   gammaDistAlfa = v
   Dim gammaDistBeta As Double
   gammaDistBeta = 1 / u

   Dim Gd1 As Double
   'Gd1 = Application.GammaDist(X, gammaDistAlfa, gammaDistBeta, True)
   'Gd1 = Application.GammaDist(X / gammaDistBeta, gammaDistAlfa, 1, True)
   Gd1 = GammaDistAc(X, gammaDistAlfa, gammaDistBeta)
   Dim Gd2 As Double
   'Gd2 = Application.GammaDist(X, gammaDistAlfa + 1, gammaDistBeta, True)
   'Gd2 = Application.GammaDist(X / gammaDistBeta, gammaDistAlfa + 1, 1, True)
   Gd2 = GammaDistAc(X, gammaDistAlfa + 1, gammaDistBeta)

   If UCase(callPut) = "C" Then
      GMEuro = S * (1 - Gd2) - X * Exp(-r * tt) * (1 - Gd1)
   Else
      GMEuro = X * Exp(-r * tt) * Gd1 - S * Gd2
   End If

   'GMEuro = GMEuro * lambda

   Exit Function
Error0:
   MsgBox "Error in GMEuro." & Err
End Function

Option Explicit

' This file includes the Applied Statistics algorithm AS 66 for calculating
' the tail area under the normal curve, and two alternative routines which
' give higher accuracy.   The latter have been contributed by Alan Miller of
' CSIRO Division of Mathematics & Statistics, Clayton, Victoria.   Notice
' that each function or routine has different call arguments.
'
```

```
'         Algorithm AS66 Applied Statistics (1973) vol22 no.3
'
'       Evaluates the tail area of the standardised normal curve
'       from x to infinity if upper is .true. or
'       from minus infinity to x if upper is false
'
Public Function ALNORM(ByVal X As Double, ByVal upper As Boolean) As Double
   Const ZERO = 0#
   Const ONE = 1#
   Const half = 0.5
   Const ltone = 7#
   Const utzero = 18.66
   Const con = 1.28
   Const P = 0.398942280444
   Const q = 0.39990348504
   Const r = 0.398942280385
   Const a1 = 5.75885480458
   Const a2 = 2.62433121679
   Const a3 = 5.92885724438
   Const b1 = -29.8213557807
   Const b2 = 48.6959930692
   Const c1 = -0.000000038052
   Const c2 = 0.000398064794
   Const c3 = -0.151679116635
   Const c4 = 4.8385912808
   Const c5 = 0.742380924027
   Const c6 = 3.99019417011
   Const d1 = 1.00000615302
   Const d2 = 1.98615381364
   Const d3 = 5.29330324926
   Const d4 = -15.1508972451
   Const d5 = 30.789933034

   Dim z, y As Double

   z = X
   If z < ZERO Then
      upper = Not upper
      z = -z
   End If
   If Not (z <= ltone Or upper And z <= utzero) Then
      ALNORM = ZERO
      If Not upper Then ALNORM = ONE - ALNORM
      Exit Function
   End If
   y = half * z * z

   If z > con Then
      ALNORM = r * Exp(-y) / (z + c1 + d1 / (z + c2 + d2 / (z + c3 + d3 / (z + c4 + d4 /
(z + c5 + d5 / (z + c6))))))
      If Not upper Then ALNORM = ONE - ALNORM
      Exit Function
   End If

   ALNORM = half - z * (P - q * y / (y + a1 + b1 / (y + a2 + b2 / (y + a3))))

   If Not upper Then ALNORM = ONE - ALNORM
End Function


Public Function ALoGam(ByVal X As Double, IFAULT As Integer) As Double
   ALoGam = Application.GammaLn(X)
End Function




'
'  ALGORITHM AS239  APPL. STATIST. (1988) VOL. 37, NO. 3
'
'  Computation of the Incomplete Gamma Integral
```

```
'
'  Auxiliary functions required: ALOGAM = logarithm of the gamma
'  function, and ALNORM = algorithm AS66
'
Public Function GammaD(ByVal X As Double, ByVal P As Double, IFAULT As Integer) As Double
   Const ZERO = 0#
   Const ONE = 1#
   Const TWO = 2#
   Const OFLO = 1E+37
   Const THREE = 3#
   Const NINE = 9#
   Const TOL = 0.00000000000001
   Const XBIG = 100000000#
   Const PLIMIT = 1000#
   Const ELIMIT = -88#

   Dim PN1, PN2, PN3, PN4, PN5, PN6, ARG, C, RN, A, B, AN As Double

   GammaD = ZERO

'  Check that we have valid values for X and P

   If (P <= ZERO Or X < ZERO) Then
     IFAULT = 1
     Exit Function
   End If
   IFAULT = 0
   If (X = ZERO) Then Exit Function

'  Use a normal approximation if P > PLIMIT

   If (P > PLIMIT) Then
      PN1 = THREE * Sqr(P) * ((X / P) ^ (ONE / THREE) + ONE / (NINE * P) - ONE)
      GammaD = ALNORM(PN1, False)
      Exit Function
   End If

'  If X is extremely large compared to P then set GAMMAD = 1

   If (X > XBIG) Then
      GammaD = ONE
      Exit Function
   End If

   If X <= ONE Or X < P Then
      '  Use Pearson's series expansion.
      '  (Note that P is not large enough to force overflow in ALOGAM).
      '  No need to test IFAULT on exit since P > 0.
      ARG = P * Log(X) - X - ALoGam(P + ONE, IFAULT)
      C = ONE
      GammaD = ONE
      A = P
      Do
         A = A + ONE
         C = C * X / A
         GammaD = GammaD + C
      Loop While (C > TOL)
      ARG = ARG + Log(GammaD)
      GammaD = ZERO

      If ARG >= ELIMIT Then GammaD = Exp(ARG)
   Else
      '  Use a continued fraction expansion
      ARG = P * Log(X) - X - ALoGam(P, IFAULT)
      A = ONE - P
      B = A + X + ONE
      C = ZERO
      PN1 = ONE
      PN2 = X
      PN3 = X + ONE
      PN4 = X * B
```

```
      GammaD = PN3 / PN4
      Do
         A = A + ONE
         B = B + TWO
         C = C + ONE
         AN = A * C
         PN5 = B * PN3 - AN * PN1
         PN6 = B * PN4 - AN * PN2
         If (Abs(PN6) > ZERO) Then
            RN = PN5 / PN6
            If (Abs(GammaD - RN) <= Application.Min(TOL, TOL * RN)) Then
               ARG = ARG + Log(GammaD)
               GammaD = ONE
               If (ARG >= ELIMIT) Then GammaD = ONE - Exp(ARG)
               Exit Function
            End If
            GammaD = RN
         End If
         PN1 = PN3
         PN2 = PN4
         PN3 = PN5
         PN4 = PN6
         If (Abs(PN5) >= OFLO) Then
            '  Re-scale terms in continued fraction if terms are large
            PN1 = PN1 / OFLO
            PN2 = PN2 / OFLO
            PN3 = PN3 / OFLO
            PN4 = PN4 / OFLO
         End If
      Loop
   End If
End Function

Public Function GammaDistAc(ByVal X As Double, ByVal alpha As Double, ByVal beta As
Double) As Double
   Dim ok As Integer
   If X = 0 Then
      GammaDistAc = 0
   Else
      GammaDistAc = GammaD(X / beta, alpha, ok)
   End If
   If ok <> 0 Then
      GammaDistAc = CVErr(xlErrNum)
   End If
End Function
```

6.2 Programação do modelo de Black-Scholes-Merton e entrópico de determinação do preço racional da opção de compra européia simples ordinária sobre uma ação no Mathematica 3.0

6.2.1 Fórmula do preço racional da opção de compra européia simples ordinária sobre uma ação (modelo de Black-Scholes-Merton):

```
Norm[x_] := 1/2*(1 + Erf[x/Sqrt[2]])

done[s_, k_, σ_, r_,t_] := (r*t + Log[s/k])/(σ*Sqrt[t]) +
(σ*Sqrt[t])/2

dtwo[s_, k_, σ_, r_,t_] := done[s, k, σ,  r, t] - σ*Sqrt[t]

BlackScholesCall[s_, k_, σ_, r_,t_] :=
   s*Norm[done[s, k, σ, r, t]] - k*Exp[-r*t]*Norm[dtwo[s, k, σ, r, t]]
```

6.2.2 Fórmula do preço racional da opção de compra européia simples ordinária sobre uma ação (modelo de entrópico):

```
<<Statistics`ContinuousDistributions`
w[s_,r_,t_]:=s*Exp[r*t]
b1[w_,v_]:=(w/v)^2
b2[w_,v_]:=w/(v^2)
gama[x_,y_]:=GammaDistribution[x,1/y]
call[s_,k_,v_,r_,t_]:=
    s*(1-CDF[gama[b1[w[s,r,t],v]+1,b2[w[s,r,t],v]],k])-
    k*Exp[-r*t]*(1-CDF[gama[b1[w[s,r,t],v],b2[w[s,r,t],v]],k])
```

7 Bibliografia

"E, com efeito, este é o comportamento do especialista. Em política, em arte, nos usos sociais, nas outras ciências tomará posições de primitivo, e ignorantíssimo; mas as tomará com energia e suficiência, sem admitir - e isto é paradoxal - especialistas dessas coisas. Ao especializá-lo a civilização o tornou hermético e satisfeito dentro de sua limitação; mas essa mesma sensação íntima de domínio e valia o levará a querer predominar fora de sua especialidade. E a conseqüência é que, ainda neste caso, que representa um maximum *de homem qualificado - especialismo - e, portanto, o mais oposto ao homem-massa, o resultado é que se comportará sem qualificação e como homem-massa em quase todas as esferas da vida."*

José Ortega y Gasset. 1984. "A barbárie do especialismo" *Humanidades* 2(6): 147-9. Brasília: Editora Universidade de Brasília.

"DPL 4.0: professional decision anaylsis software - academic edition". 1998. Applied Decision Analysis LLC - PriceWaterhouse&Coopers. Belmont, CA: Duxbury.

"NORMAS do Programa de Pós-graduação em Administração do Departamento de Administração da FEA-USP" 1994. *Caderno de Pesquisas em Administração* 51-67.

AMRAM, M. & KULATILAKA, N. 1999. *Real options: managing strategic investment in an uncertain world.* Massachusetts: Press: Harvard Business School.

AVERBUKH, V. Z. 1997. "Pricing american options using Monte Carlo simulation" (Tese de doutorado) Orientador: David Heath. Graduate School of Cornell University.

BAXTER, M. & RENNIE, A. 1998. *Financial Calculus: an introduction to derivate pricing.* Segunda reimpressão. Cambridge: Press: Cambridge University.

BEKMAN, O. R. & COSTA NETO, P. L. O. 1993. *Análise Estatística da Decisão.* 1ª reimpressão. São Paulo: Edgard Blücher.

BENISH, W. A. 1999. "Relative entropy as measure of diagnostic information" *Medical Decision Making.* 19(2): 202-6.

BENNINGA, S. 1997. *Financial modeling.* Cambridge, Massachusetts: Press: The MIT.

BERKELEY, E. C. 1965. *Sorte ou azar? Uma introdução à Probabilidade e Estatística por meio de experiências.* Coleção Cientistas do Amanhã. Brasília: Instituto Brasileiro de Educação, Ciência e Cultura - IBECC e Universidade de Brasília.

BERNSTEIN, P. L. 1997. *Desafio aos deuses: a fascinante história do risco.* 2ª edição. Rio de Janeiro: Campus.

BERRY, D. A. 1996. *Statistics: a Bayesian perspective.* Belmont, CA: Duxbury.

BIERMAN JR., H. & SMIDT, S. 1993. *The capital budgeting decision: economic analysis of investment projects.* 8ª ed. New York: Macmillan.

BIERMAN JR., H.; BONINI, C. P. & HAUSMAN, W. H. 1991. *Quantitative analysis for business decisions.* 8ª ed. Boston: Irwin.

BLACK, F. & SCHOLES, M. 1973. "The pricing of options and corporate liabilities." *Journal of Political Economy* 81: 637-59.

BOYER, C. B. 1974. *História da Matemática.* Tradução da Profa. Elza F. Gomide. 1ª reimpressão. São Paulo: Edgar Blücher & EDUSP.

BREALEY, A. B. & MYERS, S. C. 1996. *Principles of Corporate Finance.* 5ª ed. New York: McGraw-Hill.

_____. 1998. *Princípios de Finanças Empresariais.* 5ª ed. Lisboa: McGraw-Hill.

BROWN, C. 1995. *Chaos and Catastrophe Theories.* Sage University Paper series on Quantitative Applications in the Social Sciencies, 07-107. Thousand Oaks, CA: Sage.

BRUNI, A. L.; FAMÁ, R. & SIQUEIRA, J. de O. 1998. "Análise do risco na avaliação de projetos de investimento: uma aplicação do método de Monte Carlo." São Paulo: *Caderno de Pesquisas em Administração* 1(6): 62-74.

CAGLIARI, L. C. "A origem das letras do alfabeto" *Ciência Hoje* 17(98): 20-7.

CASTRO, F. M. de O. 1992. *A Matemática no Brasil.* Campinas: UNICAMP.

CHEUNG, W-Y T. 1997. “A macroscopic model of capital asset pricing and its application to option valuation” (Tese de doutorado) Orientador: Edison Tse. Department of Engineering-Economic Systems and Operations Research – Stanford University.

CLEMEN, R. 1996. *Making hard decisions: an introduction to Decision Analysis*. 2ª ed. Belmont, CA: Duxbury.

COMON, P. 1994. "Independent component analysis. A new concept?" *Signal Processing* 36: 287-314.

CONTANDRIOPOULOS, A.-P.; CHAMPAGNE, F.; POTVIN, L.; DENIS, J.-L. & BOYLE, P. 1994. *Saber preparar uma pesquisa*. São Paulo: Hucitec-Abrasco.

COVER, T. M. & THOMAS, J. A. 1991. *Elements of Information Theory.* New York: John Wiley.

_____. 1992. *Elements of Information Theory: solutions manual.* New York: John Wiley.

COX, J. C. & RUBINSTEIN, M. 1985. *Options markets*. New Jersey: Prentice Hall.

DE RIVERA, D. P. S. 1997. *Estadística: modelos e métodos*. Volume 1: fundamentos. 2ª ed. Madri: Alianza.

DERSHOWITZ, A. L. 1997. “The effect of options on pilot decisions making in the presence of risk” (Tese de doutorado) Orientador: Robert John Hansman, Jr. Department of Aeronautics and Astronautics – Massachusetts Institute of Technology.

DIAS, M. A. G. 1996. “Investimento sob incerteza em exploração e produção de petróleo” Volumes 1 e 2. (Dissertação de mestrado) Orientador: José Paulo Machado. Departamento de Engenharia Industrial de Pontifícia Universidade Católica do Rio de Janeiro.

DIXIT, A. K. & PINDICK, R. S. 1994. *Investment under uncertainty.* Princeton, NJ: Press: Princeton University.

_____. 1995. *The options approach to capital investment.* Massachusetts: Harvard Business Review.

DOS ANJOS, M. P. & PAIVA, A. C. 1998. "Modelos binomiais para o cálculo de opções: uma revisão crítica" *Resenha BM&F* 127: 19-54.

DOWNES, J. & GOODMAN, J. E. 1993. *Dicionário de termos financeiros e de investimento.* São Paulo: Nobel.

DUARTE Jr., A. M. 1998. *Gerenciamento de riscos corporativos no mercado financeiro brasileiro.* 8° Escola de Séries Temporais e Econometria - ESTE. R.J.: Associação Brasileira de Estatística - ABE e Sociedade Brasileira de Econometria - SBE.

DUNDER, C. 1998. "Portfólios eficientes incluindo opções." (Dissertação de mestrado). IME-USP.

EATWELL, J.; MILGATE, M. & NEWMAN, P. 1990. *Utility and probability.* New York: The New Palgrave.

EBRAHIMI, N.; MAASOUMI, E. & SOOFI, E. S. 1999. "Ordering univariate distributions by entropy and variance" *Journal of Econometrics* 90: 317-36.

EDLESON, M. E. 1994. "Real options: valuing managerial flexibility (A)" Massachusetts: Press: Harvard Business School. Número 9-294-109.

EDWARDS, E. 1971. *Introdução à Teoria da Informação.* Tradução do inglês por Leônidas Hegenberg e Octanny Silveira da Mota. Título original*: information transmission: an introductory guide to the application of the theory of information to the human sciences* publicado em 1964 pela Chapman & Hall em Londres. São Paulo: Cultrix-EDUSP.

EFRON, B. & TIBSHIRANI, R. J. 1993. *An introduction to the bootstrap.* New York: Chapman & Hall.

EPSTEIN, I. 1986. *Teoria da Informação.* Série Princípios. São Paulo: Ática.

EPSTEIN, R. A. 1977. *The theory of gambling and statistical logic.* New York: Press: Academic.

EZAWA, K. E. 1998. “Evidence propagation and value of evidence on influence diagrams” *Operations Research/INFORMS* 46(1).

FERREIRA, Aurélio B. de H. 1986. *Novo dicionário da Língua Portuguesa.* 2ª ed. Rio de Janeiro: Nova Fronteira.

FIRMO, J. M. 1993. *Dicionário de negócios Páginas Amarelas.* Rio de Janeiro: Expressão e Cultura.

FORST, A. J. & PRECHTER, Jr., R. R. 1995. *Elliott Wave Principle: key to market behavior.* Georgia: New Classics Library.

FRIMPONG, S. 1992. “Evaluation of mineral ventures using modern financial methods” (Tese de doutorado) Universidade de Alberta, Canadá.

GALINDO-FLORES, J. 1998. “Statistical and machine learning frameworks for Economics: analysis of error curves and applications to derivatives pricing and credit risk assessment” (Tese de doutorado) Departamento de Economia - Harvard University.

GÉLÉDAN, A. & BRÉMOND, J. 1988. *Dicionário econômico e social.* Lisboa: Livros Horizonte.

GILBERT, E. N. 1958. “An outline of Information Theory” *The American Statistician*

GLEISER, I. 1994. “Um teste para verificação da teoria do caos determinístico no mercado de capitais brasileiros” *Anais do XVI Congresso Brasileiro de Econometria* - Sociedade Brasileira de Econometria: 452-73.

GOLAN, A.; JUDGE, G. & MILLER, D. 1996. *Maximum entropy Econometrics: robust estimation with limited data.* New York: John Wiley.

GREENE, W. H. 1997. *Econometric analysis.* 3ª ed. New Jersey: Prentice Hall.

GROSJEAN, J. H. 1998. "Beating the Showcase Showdown" *Chance* 11(1): 14-9.

GULKO, B. (Les). 1998. “The entropic pricing theory” (Tese de doutorado) Orientador: Jonathan E. Ingersoll, Jr. Yale University.

_____. 1998a. “The entropic theory of bond option pricing” Internet.

HAYKIN, S. 1999. *Neural networks: a comprehensive foundation.* 2ª ed. New Jersey: Prentice Hall.

HERATH, S. B. H. 1998. "On application of pricing theory to engineering economic decisions under uncertainty" (Tese de doutorado) Universidade Auburn, Alabama.

HOLLOWAY, A. C. 1979. *Decision making under uncertainty: models and choices.* New Jersey: Prentice Hall.

HOWARD, R. A. 1996. "Options" In Zeckhauser, R. J.; Keeney, R. L. & Sebenius, eds., *Wise Choices. Decisions, Games, and Negociations.* Boston, Massachusetts: Press: Harvard Business School.

HUIZINGA, J. 1980. Homo ludens*: o jogo como elemento da cultura.* 2ª ed. Tradução do alemão por João Paulo Monteiro. São Paulo: Perspectiva.

HULL, J. C. 1998. *Futures and options markets.* 3ª ed. New Jersey: Prentice Hall.

_____. 1997. *Options, futures, and other derivatives.* 3ª ed. New Jersey: Prentice Hall.

HUTCHINSON, M.J.; LO, A.W. & POGGIO, T. 1994. "A nonparametric approach to pricing and hedging derivative securities via learning networks" *The Journal of Finance* 49(3): 851-89.

JAYNES, E. T. 1993. *Probability theory: the logic of science.* Internet.

JOHNSON, N. L.; KOTZ, S. & BALAKRISHNAN, N. 1994. *Continuous univariate distributions.* Volumes 1 e 2. 2ª ed. New York: John Wiley.

JOHNSON, N. L.; KOTZ, S. & KEMP, A. W. 1992. *Univariate discrete distributions.* John Wiley.

KARATZAS, I. & SHREVE, S. E. 1991. *Brownian motion and stochastic calculus.* New York: Springer-Verlag.

KELLY, Jr., J. L. 1956. "A new interpretation of information rate" *Bell System Technical Journal* 35: 917-926.

KEMP, A. W. 1997. "Caracterization of a discrete normal distribuition." *Journal of Statistical Planning and Inference* 63: 223-9.

KIMURA, H. *et al.* 1999. "Produtos financeiros tradicionais com características de contratos de opções" *Anais do IV Seminários em Administração.* 20 a 22 de outubro de 1999. Departamento de Administração da Faculdade de Economia, Administração e Contabilidade da USP. Internet.

KINNEY, J. J. 1997. *Probability: an introduction with statistical applications.* New York: John Wiley.

KNOPF, O. 1948. *Cálculo de probabilidades.* 3ª ed. Manuales técnicos Labor nº 21. Traduzido do alemão por Francisco Cebrián. Barcelona: Editoral Labor.

KRIPPENDORFF, K. 1986. *Information Theory: structural models for qualitative data.* Sage University Paper Series on Quantitative Applications in the Social Sciences, nº 07-062. Califórnia: Sage.

KULLBACK, S. 1997. *Information Theory and Statistics.* New York: Dover.

LAUAND, L. J. 1986. *Educação, teatro e Matemática medievais.* Coleção Elos. São Paulo: Perspectiva.

_____. 1987. "O que é uma Universidade?" São Paulo: Perspectiva e EDUSP.

LEE, R. 1996. "Personal distinctions for decisions modeling" (Tese de doutorado) Orientador: Ross D. Shachter. Department of Engineering-Economic Systems and Operations Research – Stanford University.

LEE, Y-K. L. 1997. "Extensions of options pricing theory to the analysis of new business opportunities" (Tese de doutorado) Orientador: Samuel S. Chiu. Department of Engineering-Economic Systems and Operations Research – Stanford University.

LEMGRUBER, E. F. 1992. *Avaliação de contrato de opções.* São Paulo: BM&F.

LENCE, S. H. & MILLER, D. J. 1998. "Recovering output-specific inputs from aggregate input data: a generalized cross-entropy approach" *American Journal of Agricultural Economics* 80: 852-67.

LEON, O. G. 1995. *Analisis de Decisiones: técnicas y situaciones aplicables a directivos e profesionales.* México: McGraw-Hill.

LEWIS, H. W. 1997. *Why flip a coin? The Art and Science of good decisions.* New York: John Wiley.

LIMA, E. L. 1985. *Logaritmos.* Rio de Janeiro: Sociedade Brasileira de Matemática.

LISBOA, M. de B. 1999. "Mercados seqüenciais e ativos financeiros em modelos de equilíbrio geral" *Revista de Econometria* 18(1): 115-58.

LUEHRMAN, T. A. 1995. "Capital projects as real options: an introduction" Massachusetts: Press: Harvard Business School. Número 9-295-074.

_____. 1998. "Investment opportunities as real options: getting started on the numbers" *Harvard Business Review* Reprint number: 98404.

_____. 1998. "Strategy as a portfolio of real options" *Harvard Business Review* Reprint number: 98506.

LUENBERGER, D. G. 1998. *Investment Science.* New York: Press: Oxford University.

_____. 1998. *Solutions manual for Investment Science.* New York: Press: Oxford University.

MACKAY, D. J. C. 1995. *A short course in Information Theory - outline.* Internet.

_____. 1999. *Information Theory, inference, and learning algorithms.* Versão preliminar. Internet.

MAGEE, J. F. 1964. "Decision trees for decision making" *Harvard Business Review* Reprint number: 64410.

MALINAK, S. M. 1996. "The value of searching for better alternatives" (Tese de doutorado) Orientador: Ronald A. Howard. Department of Engineering-Economic Systems and Operations Research – Stanford University.

MALUF, U. M. M. 1979. *Aspectos algébricos da entropia: sua relevância para o estudo do comportamento humano.* R.J.: Instituto de Seleção e Orientação Profissional do Centro Brasileiro de Ergonomia e Cibernética da Fundação Getúlio Vargas - ISOP/FGV.

MARKOWITZ, H. M. 1967. *Portfolio selection: efficient diversification of investments*. 3ª reimpressão. Cowles Foundation for Research in Economics at Yale University. New York: John Wiley.

Marriam-Webster´s pocket guide to synonyms. 1995. Springfield, Massachusetts: Merriam-Webster.

MEGALE, J. F. 1992. "Os clássicos das Ciências Humanas: perfil de um clássico/1" São Paulo: FEA-USP.

MENDES, B. V. M. & DUARTE Jr., A. M. 1999. *Modelos estatísticos aplicados ao mercado financeiro brasileiro.* 13º Simpósio Nacional de Probabilidade e Estatística - SINAPE.

MERTON, R. C. 1973. "Theory of rational option pricing." *Bell Journal of Economics and Management Science* 4: 141-83.

MOORE, P. G. 1986. *The business of risk*. Cambridge: Press: Cambridge University.

MOTA, L. *Adagiário brasileiro.* 1987. Belo Horizonte: Itatiaia.

NEFTCI, S. N. 1996. *An introduction to the mathematics of financial derivatives*. San Diego, CA: Press: Academic.

NETER, J.; KUTNER, M. H.; NACHTSHEIM, C. J. & WASSERMAN, W. 1996. *Applied linear statistical models.* 4ª ed. Irwin.

NEWMAN, J. N. 1999. "Model calibration in mathematical finance" (Tese de doutoramento) Orientador: Marco Avellaneda. Departamento de Matemática - New York University.

NICHOLS, N. A. 1994. "Scientific management at Merck: an interview with CFO Judy Lewent" *Harvard Business Review* 89-99.

NORTH, D. W. 1968. "A tutorial introduction to decision theory" *IEEE Transations on Systems Science and Cibernetics* vol. SSC-4(3): 200-10.

NUMESMAIA, M. das M. S. 1977. "Análise de dados categóricos usando Teoria da Informação" (Dissertação de mestrado) Orientador: Clóvis de Araújo Peres. Departamento de Estatística do IME-USP.

PANEL discussion: downstream decisions (options) and dynamic modeling. 1998. Seattle INFORMS Meeting, Sessão SE01.

PARK, C. S. & SHARP-BETTE, G. P. 1990. *Advanced Engineering Economics.* New York: John Wiley.

PATSIS, N. T. 1998. "Pricing american-style exotic options using ordinal optimization" (Tese de doutorado) Orientador: Yu-Chi Ho. Harvard University.

PAULI, R. U.; MAUAD, F. C. & HEILMANN, H. P. 1980. *Física 4: eletricidade, magnetismo, física moderna e análise dimensional.* São Paulo: E.P.U.

PETERS, E. E. 1994. *Fractal market analysis: applying chaos theory to investment and economics.* New York: John Wiley.

PIEPER, J. 1960. *Virtudes fundamentais.* Tradução de Narino e Silva e Beckert da Assumpção. Coleção Éfeso. Lisboa: Aster.

PINDYCK, R. S. & RUBINFELD, D. L. 1994. *Microeconomia.* 3ª ed. Makron Books.

PLATÃO. 1980. "Diálogos: Leis e Epínomis". Tradução do grego por Carlos Alberto Nunes. Universidade Federal do Pará.

PLISKA, S. R. 1998. *Introduction to mathematical finance: discrete time models.* Massachusetts: Blackwell.

POINCARÉ, H-J. 1984. *A ciência e a hipótese.* Brasília: Universidade de Brasília.

POIRIER, D. J. 1995. *Intermediate Statistics and Econometrics: a comparative approach.* Cambridge, Massachusetts: Press: The MIT.

POULTON, E. C. 1994. *Behavioral Decision Theory: a new approach.* Cambridge: Press: Cambridge University.

PRATT, J. W; RAIFFA, H. & SCHLAIFER, R. 1996. *Introduction to Statistical Decision Theory.* 2ª reimpressão. Cambridge, Massachusetts: Press: The MIT.

RAIFFA, H. 1977. *Teoria da Decisão: aulas introdutórias sobre escolhas em condições de incerteza.* Petrópolis, RJ: Vozes-EDUSP.

REZA, F. M. 1994. *An introduction to Information Theory.* Reimpressão da publicação de 1961 pela *McGraw-Hill Electrical and Electronic Engineering Series.* New York: Dover.

RICIERI, A. P. 1990. *Fractais e caos: a Matemática de hoje.* São Paulo: Prandiano.

_____. 1994. *Otimização natural: a genética dos máximos e mínimos.* São Paulo: Prandiano.

ROSS, S. M. 1996. *Stochastic Processes.* 2ª ed. New York: John Wiley.

_____. 1997. *Introduction to probability models.* 6ª ed. San Diego, CA: Press: Academic.

_____. 1998. *A first course in probability.* 5ª ed. New Jersey: Prentice Hall.

RUBINSTEIN, M. 1994. "Implied binomial trees" *The Journal of Finance* 69(3): 771-818.

SALOMON, D. V. 1972. *Como fazer uma monografia: elementos de metodologia do trabalho científico.* 2ª ed. Belo Horizonte: Interlivros de Minas Gerais.

SAMPERI, D. 1998. "Inverse problems, model selection and entropy in derivative security pricing" (Tese de doutorado) Orientador: Marco Avellaneda. Departamento de Matemática da *Graduate School of Arts and Sciences* - New York University.

SASAKI, K. 1968. *Statistical for modern business decision making.* Belmont, CA: Wadsworth.

SASSATANI, R. & SIQUEIRA, J. de O. 1998. "Precificação de opções européias e exóticas por simulação de Monte Carlo." São Paulo: *Anais do III SEMEAD* (Seminários de Administração) promovido pelo Programa de Pós-Graduação em Administração do Departamento de Administração da FEA/USP.

SASSATANI, R. 1999. "Uma análise empírica do preço da incerteza nos contratos futuros de índice Bovespa da BM&F" (Dissertação de mestrado) Orientador: José Roberto Securato. Departamento de Administração da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo.

SCHLAIFER, R. 1961. *Introduction to Statistics for business decisions.* New York: McGraw-Hill.

SEARS, F. W. 1955. *An introduction to Thermodynamics, the Kinetic Theory of Gases, and Statistical Mechanics.* 2ª ed. Cambridge: Addison-Wesley.

SECURATO, J. R. 1996. *Decisões financeiras em condições de risco.* São Paulo: Atlas.

_____. (Org.) 1999. *Cálculo financeiro das tesourarias: bancos e empresas.* São Paulo: Saint Paul Institute of Finance.

SHANNON, C. E. 1948. "A mathematical theory of communication" *Reprinted with corrections from The Bell System Technical Journal* 27: 379-423 e 623-656. Internet.

SHARPE, R. V. 1998. "Investment in human capital: an option for financial aid" (Tese de doutorado) Orientador: Cecilia Conrad. Claremont Graduate University.

SHARPE, W. F. & ALEXANDER, G. J. *Investments.* 4ª ed. NJ: Prentice Hall.

SHAW, W. 1998. *Modelling financial derivatives with Mathematica.* Cambridge: Press: Cambridge University.

SHREVE, S. 1997. *Stochastic Calculus and Finance.* Internet.

SILVA, P. A. L. da. 1990. *Fundamentos da Teoria da Decisão.* 9º Simpósio Brasileiro de Probabilidade e Estatística - SINAPE.

SIQUEIRA, J. de O. 1997. "Mensuração da estrutura de preferência do consumidor: uma aplicação de *conjoint analysis* em Marketing." (Dissertação de mestrado) Orientador: Hiroo Takaoka. apresentada à Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo.

SIQUEIRA, J. de O. & SASSATANI, R. 1999. "Formação do preço da opção baseada no valor da informação: uma abordagem bayesiana" *Anais do IV Seminários em Administração.* 20 a 22 de outubro de 1999. Departamento de Administração da Faculdade de Economia, Administração e Contabilidade da USP. Internet.

_____. 1999. "Determinação bayesiana do preço racional de opção" Departamento de Administração da Faculdade de Economia, Administração e Contabilidade da USP. Não publicado.

SMIDTS, A. 1997. "The relationship between risk attitude and strength of preference: a test of intrinsic risk attitude" *Management Science* 43(3): 357-70.

SMITH, J. E. & NAU, R. F. 1995. "Valuing risky projects: options pricing theory and Decision Analysis" *Management Science* 41(5): 795-816.

SNEE, R. D. 1998. "Discussion: increasing the value of the statistics profession" *Technometrics* 40(3): 207-210.

STEINER NETO, P. J. 1998. "A percepção dos resultados esperados pelos beneficiários como fator de influência no processo decisório" (Tese de doutorado) Orientador: Abraham Sin Oih Yu. Departamento de Administração da Faculdade de Economia, Administração e Contabilidade da USP.

STIGLER, S. M. 1986. *The history of Statistics: the measurement of uncertanty before 1900.* Cambridge, Massachusetts: Press: Harvard University.

STONEBRAKER, J. S. 1993. "Applying decision analysis to sequencial decisions under uncertainty with continuos decisions variables" (Tese de doutorado) Arizona University State.

STYAN, E. M. "Chance in Aristotle´s *Physics*" *Chance* 11(4): 11-6.

SUBRAMANIAN, A. R. 1998. "Applications of Stochastic Control to Mathematical Finance" (Tese de doutorado) Cornell University.

TEISBERG, E. O. 1995. "Methods for evaluating capital investiment decisions under uncertainty" *Real options in capital investment: models, strategies, and applications*. Edited by Lenos Trigeorgis. Connecticut: Praeger.

The new Marriam-Webster dictionary. 1989. Springfield, Massachusetts: Merriam-Webster.

THEIL, H. 1967. *Economics and Information Theory*. Amsterdam: North-Holland.

_____. 1971. *Principles of Econometrics*. New York: John Wiley.

TRIGEORGIS, L. 1998. *Real options: managerial flexibility and strategy in resource allocation*. Cambridge, Massachusetts: Press: The MIT.

TRIPPI, R. R. (*editor*). 1995. *Chaos & nonlinear dynamics in the financial market*s. Chicago: Irwin.

TURBAN, E. & MEREDITH, J. R. 1994. *Fundamentals of Management Science*. 6ª ed. Burr Ridge, Illinois: Irwin.

VAPNIK, V. N. 1995. *The nature of Statistical Learning Theory.* New York: Springer-Verlag.

VARIAN, H. R. (Editor). 1996. *Computational Economics and Finance: modeling and analysis with Mathematica*. New York: TELOS/Springer-Verlag.

_____. 1987. "The arbitrage principle in financial economics." *Economic Perspectives* 1(2): 55-72.

VINCE, R. 1999. *Cálculo e análise de riscos no mercado financeiro*. São Paulo: Makron Books.

WILMOTT, P.; HOWISON, S. & DEWYNNE, J. 1997. *The Mathematics of financial derivatives*. Cambridge: Press: Cambridge University.

WINSTON, W. 1998. *Financial models using simulation and optimization: a step-by-step guide with Excel and Palisade´s DecisionTools Software*. New York: Palisade.

WOLF, D. R. 1996. "Information and correlation in Statistical Mechanical Systems." http://dino.ph.utexas.edu/~wolf/wd.

WOLFRAM, S. 1991. *Mathematica: a system for doing Mathematics by computer.* 2ª ed. Illinois: Addisson-Wesley.

WONNACOTT, T. H. & WONNACOTT, R. J. 1990. *Introductory Statistics for Business and Economics.* 4ª ed. New York: John Wiley.

_____. 1990. *Introductory Statistics.* 5ª edição. New York: John Wiley.

YANG, Y. 1997. "Maximum entropy option pricing" (Tese de doutorado) Orientador: Paul Beaumont. College of Social Science - The Florida State University.

YU, D. 1998. "Frictional arbitrage valuation" (Tese de doutorado) Princeton University. Departamento de Economia.

YUNES, R. A. 1995. *A organização da matéria: acaso ou informação.* Florianópolis: UFSC.

ZUCKERMAN, M. *et al.* 1980. "Sensation seeking and its biological correlates" *Psychological Bulletin.* American Psychological Association. 88(1): 187-214.

ZURBANO, R. V. 1976. *Teoría de la Decisión Empresarial.* Bilbao: Ediciones Deusto.