

"A FEA e a USP respeitam os direitos autorais deste trabalho. Nós acreditamos que a melhor proteção contra o uso ilegítimo deste texto é a publicação online. Além de preservar o conteúdo motiva-nos oferecer à sociedade o conhecimento produzido no âmbito da universidade pública e dar publicidade ao esforço do pesquisador. Entretanto, caso não seja do interesse do autor manter o documento online, pedimos compreensão em relação à iniciativa e o contato pelo e-mail [bibfea@usp.br](mailto:bibfea@usp.br) para que possamos tomar as providências cabíveis (remoção da tese ou dissertação da BDTD)."

UNIVERSIDADE DE SÃO PAULO  
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE  
DEPARTAMENTO DE ADMINISTRAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO

NOVOS CRITÉRIOS DE PRECIFICAÇÃO EM  
MERCADOS INCOMPLETOS: CONSISTÊNCIA,  
RACIONALIDADE E DESVIO-PADRÃO  
NORMALIZADO POR VOLUME

Christian Johannes Zimmer

Orientador: Prof. Dr. José de Oliveira Siqueira

São Paulo  
Agosto de 2005

UNIVERSIDADE DE SÃO PAULO  
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE  
DEPARTAMENTO DE ADMINISTRAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO

**NOVOS CRITÉRIOS DE PRECIFICAÇÃO EM  
MERCADOS INCOMPLETOS: CONSISTÊNCIA,  
RACIONALIDADE E DESVIO-PADRÃO  
NORMALIZADO POR VOLUME**

Christian Johannes Zimmer

Orientador: Prof. Dr. José de Oliveira Siqueira

São Paulo  
Agosto de 2005

Prof. Dr. Adolpho José Melfi  
Reitor da Universidade de São Paulo

Prof<sup>a</sup>. Dr<sup>a</sup>. Maria Tereza Leme Fleury  
Diretora da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. Eduardo Pinheiro Gondim de Vasconcellos  
Chefe do Departamento de Administração

Prof. Dr. Isak Kruglianskas  
Coordenador do Programa de Pós-Graduação em Administração

---

CHRISTIAN JOHANNES ZIMMER

**NOVOS CRITÉRIOS DE PRECIFICAÇÃO EM  
MERCADOS INCOMPLETOS: CONSISTÊNCIA,  
RACIONALIDADE E DESVIO-PADRÃO  
NORMALIZADO POR VOLUME**

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

Orientador: Prof. Dr. José de Oliveira Siqueira

São Paulo  
Agosto de 2005

Tese defendida e aprovada no Departamento de Administração da Universidade de São Paulo- Programa de Pós-Graduação em Administração, pela seguinte banca examinadora:

Prof. Dr. José de Oliveira Siqueira – FEA/USP

Prof. Dr. Marcos Eugênio da Silva – FEA/USP

Prof. Dr. Oswaldo Luiz do Valle Costa – POLI/USP

Prof. Dr. Vladimir Belitsky – IME/USP

Prof. Dr. Richard Saito – FGV, São Paulo

### FICHA CATALOGRÁFICA

Elaborada pela Seção do Processamento Técnico do SBD/FEA/USP

Zimmer, Christian Johannes

Novos critérios de precificação em mercados incompletos: consistência, racionalidade e desvio-padrão normalizado por volume / Christian Johannes Zimmer. – São Paulo, 2005.

200 p.

Tese (doutorado) – Universidade de São Paulo, 2005

Bibliografia.

1. Martingais. 2. Finanças – Modelagem matemática. 3. Ordenação estocástica. 4. Otimização estocástica. I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade. II. Título.

CDD – 519.287

A minha família por me dar o sabor da vida.

Agradeço a todos que me ajudaram a concluir essa tese:

- minha esposa, Vanise, por me encorajar todos os dias e me agüentar nos momentos de muito trabalho,
- meus filhos, Rafael e Giovanni, por me motivarem a acabar em tempo,
- minha mãe, Maria Dolores, por ter me ensinado a buscar sempre o novo e a ser curioso,
- ao resto da minha família pelo apoio contínuo,
- meu orientador, Prof. Dr. José de Oliveira Siqueira, pela orientação, disposição e pelo apoio amigo,
- meus colegas de trabalho, especialmente José Euclides de Mello Ferraz, pela ajuda constante,
- a todos os amigos que colaboraram, mas cujas citações criariam uma lista quase infinita.



## RESUMO

O objetivo principal desse trabalho é propor novas alternativas de precificação de ativos em mercados financeiros incompletos. Para isso, introduzimos três abordagens originais:

1. A “dispersão relativa entrópica” (ERD) é uma medida de dispersão que normaliza o desvio-padrão de uma variável aleatória pela sua entropia para poder melhor comparar e ordenar variáveis de domínio diferente.
2. A “precificação consistente”, proposta em duas versões, aplica de forma recursiva o conceito de *perda esperada* e de *expected regret* para atribuir um valor único a uma variável aleatória.
3. O “equivalente monetário determinístico” (EMD) substitui o conceito de utilidade esperada por uma aproximação por momentos centrados da distribuição da variável, seguindo a idéia da expansão de Taylor, mas baseando-se no equivalente certo.

Para uma clara e completa introdução desses novos conceitos, serão necessários alguns passos preparatórios que contribuem para a literatura brasileira em finanças matemáticas:

- Um exemplo paradigmático de um mercado incompleto uniperiódico com três estados futuros será usado para visualizar alguns resultados intermediários e finais.
- As razões da incompletude serão diferenciadas em erros de estimação e erro de modelo. Na parte de erro de estimação contribuimos com uma análise dos métodos usados na literatura e a proposta de usar o método de funções de penalidade. Esse método, apesar de ser uma ferramenta conhecida, ainda não foi aplicado no problema de erro de estimação.
- A teoria de dualidade será exposta para esclarecer a conexão entre as otimizações de funções de utilidade e medidas de probabilidade. Como um dos primeiros textos no Brasil analisamos essa teoria do ponto de vista da sua aplicabilidade no mercado financeira, tanto para o caso paradigmático como para o caso de processos de preços gerais.
- Alguns dos métodos de escolha por medida de risco serão expostos, enquadrados e as suas conexões com as novas propostas serão discutidas.

## *ABSTRACT*

*The objective of this thesis is the development of some new alternatives for asset pricing in incomplete markets. We propose the following original approaches:*

- 1. The “entropic relative dispersion” (ERD) is a measure created by normalization of the standard deviation with a variable’s volume. This allows one to compare and ordenate random variables with different domains.*
- 2. The “consistent price”, for which we offer two slightly different versions, uses in a recursive way the approach of expected shortfall (respective expected regret) in order to attribute a unique value to a random variable.*
- 3. The “monetary certainty-equivalent” (EMD) substitutes the concept of expected utility with an approximation of a distribution by its central moments. The approach uses a Taylor expansion but it is based on the certainty-equivalent.*

*During the introduction of these new concepts we also contribute with the following analyses to the Brazilian literature in mathematical finance:*

- A basic example of an incomplete static market with three future states will be used to visualize the consequences of some of the approaches.*
- The origins of incompleteness are separated into estimation and model error. In the part of estimation error we contribute with a unifying analysis of the methods used in the literature as well as with the proposal of using penalty functions. The concept of penalty functions, although being already well-known, was not yet applied to the problem of estimation error.*
- The duality theory will be exposed in a detailed form in order to clarify the connection between optimization of utility functions and probability measures. We consider this analysis as one of the first in the Brazilian financial literature and show how it works for the basic model as well as for a market with a general price process.*
- A representative list of (so-called) risk measures will be exposed, compared and their connections with our new proposals for asset pricing will be discussed.*

# Sumário

<b>Lista de Acrônimos</b>	<b>3</b>
<b>Lista de Tabelas</b>	<b>5</b>
<b>Lista de Figuras</b>	<b>7</b>
<b>1 Introdução</b>	<b>9</b>
<b>2 Modelo básico</b>	<b>15</b>
2.1 O exemplo de uma árvore trinária . . . . .	15
2.2 Um mercado incompleto em tempo contínuo . . . . .	19
<b>3 Critérios de decisão sob incerteza</b>	<b>23</b>
3.1 Incerteza, risco e ambigüidade . . . . .	23
3.1.1 Medindo incerteza . . . . .	26
3.1.2 Medindo dispersão . . . . .	29
3.2 Medidas de risco . . . . .	31
3.2.1 Medidas de risco derivadas das distorções de valores . . . . .	34
3.2.2 Medidas de risco por equivalente certo e conjuntos de aceitação . . . . .	36
3.2.3 Classes de medidas de risco . . . . .	38
3.3 Função de utilidade . . . . .	41
3.3.1 Generalidades . . . . .	41
3.3.2 Estratégia eficiente, função de utilidade eficiente e controle ótimo . . . . .	45
<b>4 Métodos de alocação ótima</b>	<b>49</b>
4.1 Introdução . . . . .	49
4.2 A conexão entre o problema primal e dual . . . . .	51
4.2.1 Visualização da dualidade no exemplo paradigmático . . . . .	53
4.2.2 Caso completo com $\Omega$ finito e discreto . . . . .	57
4.2.3 Caso incompleto com $\Omega$ finito e discreto . . . . .	59
4.2.4 Dualidade no caso geral . . . . .	60
4.3 Algumas abordagens existentes na precificação em mercados incompletos . . . . .	69
4.4 Medindo distâncias entre conjuntos . . . . .	74
<b>5 Erro de estimação</b>	<b>79</b>
5.1 Introdução . . . . .	79
5.2 Abordagens existentes . . . . .	83
5.3 O uso da função de perda e estatística de Bayes . . . . .	84
5.4 O método de simulação ( <i>reamostragem</i> ) . . . . .	87
5.5 Abordagens alternativas usando funções de penalidade . . . . .	89

5.5.1	Aversão quadrática à incerteza . . . . .	90
5.5.2	Minimização do desvio da melhor escolha . . . . .	90
5.5.3	Minimização do desvio da escolha ótima . . . . .	91
5.5.4	A característica da mínima sensibilidade . . . . .	92
5.6	O problema de determinar valores esperados equivalentes no espaço dos valores médios . . . . .	94
5.7	Aplicação . . . . .	95
5.7.1	Os dados do mercado financeiro . . . . .	95
5.7.2	Os efeitos de diferentes abordagens . . . . .	96
<b>6</b>	<b>Erro de modelo</b>	<b>113</b>
6.1	Introdução: o problema de comparação de modelos . . . . .	113
6.2	A medida entrópica de dispersão relativa . . . . .	115
6.2.1	Introdução e definição da nova medida . . . . .	115
6.2.2	Aplicação da medida ERD na precificação em mercados incompletos	120
6.3	A precificação consistente . . . . .	127
6.3.1	Introdução . . . . .	127
6.3.2	Risco como valores não desejados . . . . .	129
6.3.3	Risco como valor do desvio do <i>benchmark</i> . . . . .	144
6.3.4	Aplicação da consistência na precificação em mercados incompletos	150
6.4	O Equivalente Monetário Determinístico (EMD) de grau $n$ . . . . .	156
6.4.1	Introdução . . . . .	156
6.4.2	Comportamento racional do investidor . . . . .	157
6.4.3	Utilidade esperada e preferência por momentos . . . . .	160
6.4.4	O <i>EMD</i> , a irracionalidade da fronteira eficiente e a decisão de alocação . . . . .	166
<b>7</b>	<b>Conclusão</b>	<b>169</b>
<b>A</b>	<b>Resultados de dualidade para espaços convexos</b>	<b>171</b>
<b>B</b>	<b>A ERD para algumas distribuições</b>	<b>175</b>
<b>C</b>	<b>Convexidade de ERD nas estratégias</b>	<b>179</b>
	Referências Bibliográficas	180

# Lista de Acrônimos

BM&F	Bolsa de Mercadorias e Futuros de São Paulo
BM	Valor de referência ( <i>Benchmark</i> )
CARA	Aversão ao risco absoluta constante
CVaR	Valor em risco condicionado ( <i>Conditional VaR</i> )
DBM	Desvio do <i>benchmark</i> .
DR	<i>Downside risk</i>
EMD	Equivalente Monetário Determinístico
EMM	Medida de Martingale Equivalente
ER	<i>Expected regret</i>
ERD	Dispersão Entrópica Relativa
ES	Perda esperada ( <i>expected shortfall</i> )
FTAP	Teorema Fundamental de Precificação de Ativos
HARA	Aversão ao risco absoluta hiperbólica
HJB	Hamilton-Jacobi-Bellman
IBOVESPA	Índice da bolsa de valores de São Paulo
ICAPM	Modelo de precificação intertemporal de ativos
IGP-M	Índice geral de preços do mercado
KL	Distância de Kullback-Leibler
LFT	Letra financeira do Tesouro Nacional
RF	Renda fixa
TCE	Esperança condicionada na cauda
WCE	Pior valor esperado condicionado
VaR	Valor em risco ( <i>Value at risk</i> )

# Lista de Tabelas

4.1	Conexão entre distâncias generalizadas e funções de utilidade . . . . .	76
4.2	Conexão entre <i>núcleos de informação</i> e funções de utilidade . . . . .	77
5.1	Ativos usados na otimização. . . . .	100
5.2	Carteira segundo a otimização média-variância simples. . . . .	101
5.3	Pesos dos ativos na carteira com restrições de aplicação. . . . .	101
5.4	Pesos dos ativos usando a abordagem de Black-Litterman com restrições de alavancagem. . . . .	102
5.5	Fronteira <i>reamostrada</i> com dados originais. . . . .	103
5.6	Fronteira <i>reamostrada</i> com dados alternativos. . . . .	104
5.7	Fronteira eficiente usando (5.3) com dados originais. . . . .	105
5.8	Fronteira eficiente usando (5.3) com dados alternativos. . . . .	106
5.9	Fronteira eficiente usando (5.5) com dados originais. . . . .	107
5.10	Fronteira eficiente usando (5.5) com dados alternativos. . . . .	108
5.11	Retornos determinísticos implícitos da fronteira eficiente de MARKOWITZ (1952) com restrições de não-alavancagem . . . . .	109
5.12	Retornos determinísticos implícitos da fronteira eficiente de BLACK; LIT-TERMAN (1992) com restrições de não-alavancagem . . . . .	110
5.13	Retornos determinísticos implícitos da fronteira eficiente de reamostragem com restrições de não-alavancagem. . . . .	111
6.1	Inconsistência entre preferência por momentos e utilidade esperada. As funções de utilidade são: $U_1(x) = 1 - \exp(-x)$ , $U_2(x) = \ln(1+x)$ , $U_3(x) = x^{\frac{1}{2}}$ . . . . .	161
6.2	Manutenção da decisão com expansão até ordem 3. . . . .	163
6.3	Inversão da decisão com expansão até ordem 4 para as funções de utilidade $U_1$ e $U_2$ . . . . .	163
B.1	As distribuições sob análise. . . . .	176
B.2	Entropias das distribuições. . . . .	177
B.3	<i>ERD</i> para as distribuições. . . . .	177

# Lista de Figuras

2.1	Árvore trinomial visualizando o mundo paradigmático. . . . .	16
2.2	O valor do derivativo na data final $t_1 = T$ em dependência do valor do ativo subjacente $S$ . . . . .	18
3.1	A distinção entre risco e incerteza . . . . .	26
3.2	Duas variáveis com diferentes variedades e sem dominância estocástica. . .	29
3.3	A importância de modalidade para a compatibilidade da dominância estocástica e concentração. . . . .	31
5.1	Problema de substituição . . . . .	80
5.2	Gráfico das frequências absolutas (ordenadas) dos pesos do ativo IBOVESPA (abscissas) na carteira numa simulação com 1500 amostras. . . . .	89
5.3	Conjunto de fronteiras reamostradas. . . . .	96
6.1	A abordagem de consistência no caso de distribuição uniforme contínua. .	133
6.2	O <i>fatiamento</i> da distribuição na precificação por consistência. . . . .	137
6.3	Os preços intermediários na determinação do preço consistente com desvio do <i>benchmark</i> . . . . .	149
A.1	O cone $\mathcal{C} \subseteq \mathbb{R}_+^2$ e seu cone dual $\mathcal{C}^*$ . . . . .	172

# Capítulo 1

## Introdução

Um mercado financeiro é incompleto quando nem todo o pagamento numa data futura  $T$ ,  $D_T$ , pode ser replicado com os ativos do mercado. Um critério de completude num modelo Black-Scholes-Merton é que o número de fontes independentes de incerteza  $\{W^{(i)}\}_{i=1,\dots,d}$  seja igual ao número de ativos independentes  $\{S^{(j)}\}_{j=1,\dots,n}$ , i.e.,  $n = d$ ,<sup>1</sup> e que a matriz de volatilidade  $\{\sigma_{ij}\}_{i=1,\dots,n,j=1,\dots,d}$  seja inversível (KARATZAS; SHREVE, 1991).<sup>2</sup> O primeiro teorema de precificação de ativos (FTAP) estabelece que, no caso completo, sob a condição de não-arbitragem, o preço de qualquer pagamento futuro  $D_T$  é único (HARRISON; KREPS, 1979). No caso oposto, se não existir ativo nem para replicação parcial, é comum na teoria de finanças usar uma função de utilidade (geralmente a de von Neumann-Morgenstern) para precificar o ativo. Como essa função de utilidade não é única, a escolha depende do agente no mercado e, portanto, o preço determinado é subjetivo. Uma combinação dessas abordagens encontra-se na replicação de derivativos em mercados incompletos: a parte replicável por ativos negociados é precificada por arbitragem, enquanto a parte não replicável é precificada por utilidade e abre-se uma faixa de preços possíveis que são livres de arbitragem. Qualquer escolha de um preço dentro dessa faixa é subjetiva e alguns desses preços podem ser justificados por argumentos econômicos. Observamos que para analisar questões de determinação de preços em mercados incom-

---

<sup>1</sup>É importante lembrar que essa relação vale no mundo de um numerário, i.e., existe um ativo adicional em relação ao qual os outros  $n$  ativos foram normalizados.

<sup>2</sup>Alguns autores como MERTON (1990, cap.16.2) usam a relação entre estados da natureza e número de ativos para caracterizar completude. Isso implica uma diferenciação entre completude *dinâmica* e *estática*. A origem destas abordagens encontra-se nas idéias de replicação com formulação em equações diferenciais. Mas, como o caso geral pode ser analisado com a teoria dos martingales (vide HARRISON; KREPS, 1979), esta diferenciação não vai nos preocupa no presente trabalho.



pletos, não precisamos necessariamente adicionar derivativos ao nosso modelo. Porém, escolhemos usá-los pelas seguintes razões:<sup>3</sup>

1. Existe uma vasta literatura sobre a precificação em mercados incompletos usando derivativos cujas noções usaremos (um arcabouço teórico parecido facilita o entendimento).
2. Com a introdução de um pagamento futuro, podemos criar o caso de o valor da carteira do investidor atingir valores negativos, o que terá impactos na formalização dos resultados.

Um caso especial de mercado incompleto é o modelo de precificação intertemporal de ativos (CAPM intertemporal ou ICAPM), vide MERTON (1990), capítulo 15.<sup>4</sup> No ICAPM, o conjunto de oportunidades de investimentos (conjunto de consumo) não é constante. Existem flutuações intertemporais da volatilidade dos ativos, sendo que os preços dos  $n$  ativos  $S = (S_1, \dots, S_n)$  seguem um processo de Itô com valores esperados (*drifts*)  $\mu = (\mu_1, \dots, \mu_n)$  determinísticos<sup>5</sup>

$$\frac{dS_i(\omega, t)}{S_i(\omega, t)} = \mu_i(t) dt + \sigma_i(\omega, t) dW_i(t),$$

mas as volatilidades  $\sigma = (\sigma_1, \dots, \sigma_n)$  são processos aleatórios, dependendo de um processo de Wiener  $\tilde{W}$  independente de  $W$ , com *drifts*  $m$  e volatilidades  $s$  determinísticos:

$$\frac{d\sigma_i(\omega, t)}{\sigma_i(\omega, t)} = m_i(t) dt + s_i(t) d\tilde{W}_i(t).$$

Segue como resultado que o investidor não use somente a carteira do mercado, mas que ele inclua também um seguro contra futuras variações no consumo causadas pela volatilidade estocástica. A carteira de proteção (*hedging*) contra fontes de incerteza é composta por ativos que têm maior covariância com o fator de incerteza “volatilidade” (vide MERTON, 1990, equação 15.23 na seção 15.7).

No presente trabalho, vamos analisar duas formas de determinar o “preço racional” nos mercados com incerteza não *hedgeável*. A razão de usarmos essas propostas será explicada durante as análises dos problemas considerados. Os problemas de interesse são as incompletudes criados por:<sup>6</sup>

<sup>3</sup>Para uma análise econômica de mercados incompletos vide MAGILL; QUINZII (1996).

<sup>4</sup>O artigo original é MERTON (1973).

<sup>5</sup>A formulação matemática exata será deferida para a seção 2.2.

<sup>6</sup>Pode-se pensar em incluir também uma incerteza “estrutural” que captura o fato de não conhecer as fontes de incerteza.

1. incerteza sobre a forma das distribuições usadas (erro de modelo),
2. incerteza sobre parâmetros das distribuições usadas (erro de estimação).

Tipicamente, na análise de mercados incompletos, os processos de preços  $S$ , mesmo sendo de forma geral, são dados. No máximo, os parâmetros dos processos são desconhecidos, o que cria a incompletude.<sup>7</sup> Nesse quadro, não é incluída a falta de especificação do tipo dos processos usados, pois é a partir deles que é definido o conjunto de EMMs  $\mathcal{M}(S)$ . Imagine que não sabemos se  $S = S_1$  é um processo discreto ou se  $S = S_2$  é um processo contínuo. De qualquer forma, teremos que  $\mathcal{M}(S_1) \neq \mathcal{M}(S_2)$  e não faria muito sentido definir um critério de escolha entre medidas  $\mathbb{Q} \in \mathcal{M}(S)$ . Teríamos que escolher primeiramente entre  $S_1$  e  $S_2$ . A escolha de modelos e dos seus parâmetros são problemas antecedentes, mas LAKNER (1995) esclarece que, do ponto de vista abstrato, o tratamento dos problemas de incompletude é equivalente.

O **objetivo dessa tese** é teórico-normativo-quantitativo, de forma que as propostas visam a ajudar na criação de novos conhecimentos teóricos. Por natureza do problema, será necessário o uso de diversos conceitos matemáticos-quantitativos. Portanto, começamos o trabalho, no capítulo 2, com um **exemplo paradigmático**, por meio do qual ilustramos ao longo da tese as conseqüências das abordagens. No mesmo capítulo introduzimos também um mercado financeiro incompleto geral que usamos em diversas oportunidades para deduzir resultados teóricos.

O capítulo 3 tem como objetivo enquadrar alguns dos critérios existentes no trabalho presente. Depois de discutir e estabelecer a nossa definição da diferença entre risco e incerteza, analisamos algumas medidas de risco e funções de utilidade que podem ser usadas para precificar ativos (esse problema será chamado de primal no capítulo 4). Podemos fazer uma classificação das abordagens segundo o tipo de função-objetivo usada:

1. otimização de utilidade (esperada) da carteira
2. otimização de uma medida de incerteza ou risco da carteira
3. outras.

---

<sup>7</sup>Os processos de consumo e investimento são adaptados a uma filtração  $F$  maior que a filtração  $G$  criada por estes processos. O caso típico é que a filtração  $F$  é criada pelos preços observados  $S$ .

A relação entre as duas primeiras classes é, em muitos casos, bastante óbvia, sendo a função-objetivo da segunda classe a negativa da função-objetivo da primeira. Na classe “*outras*”, encontram-se abordagens como relação ótima entre risco e retorno ou critérios de estabilidade.

A conexão do problema primal com a escolha de uma **medida equivalente neutralizadora de risco (EMM)**<sup>8</sup>, ou seja, o problema dual, é o conteúdo do capítulo 4. Na seção 4.3, analisamos algumas das abordagens mais comuns e mostramos na seção 4.4 que, de forma abstrata, a escolha de uma função específica pode ser vista como a escolha de uma medida de distância entre conjuntos. Como a **abordagem por martingales** serve para resolver o problema dual e como não existem trabalhos brasileiros sobre este método em finanças, achamos importante expor o funcionamento em pormenores.

No ambiente de incerteza sobre os parâmetros da distribuição discutimos, no capítulo 5, primeiro, diversas abordagens existentes de uma forma geral e unificada. Depois, na seção 5.5, propomos usar **funções de penalização**, pois elas permitem implementar, de forma fácil, diversos objetivos de um investidor. Para mostrar a flexibilidade desse conceito, propomos duas novas medidas que visam a criar carteiras com mais estabilidade nas alocações de forma que as composições das carteiras são pouco sensíveis a alterações nos parâmetros. Analisamos os resultados de diversos critérios de tratamento dos erros de estimação numa aplicação prática no mercado financeiro brasileiro.

Para o problema de escolha de modelo, faremos, no capítulo 6, três outras propostas originais:

1. A **medida entrópica de dispersão relativa** (seção 6.2) supõe a existência do segundo momento de uma variável aleatória, pois o desvio-padrão da variável será normalizado pelo volume do conjunto que contém a maior parte da informação sobre a variável. Esse ajuste permite ordenar, de forma total, diversas distribuições: discretas e contínuas e de domínios diferentes. Porém, como a medida se baseia no desvio-padrão, ela não serve como medida de risco.
2. Na seção 6.3 será introduzida a **precificação consistente**. Essa abordagem é uma

---

<sup>8</sup>Para ser consistente com o nosso ponto de vista exposto na seção 3.1 sobre a diferença entre risco e incerteza, deveríamos chamar a medida de neutralizadora de (do preço da) incerteza. Como o termo técnico aplicado na literatura é o de neutralizadora de risco, ficamos com essa notação para facilitar a leitura.

forma genérica de avaliar um ativo, baseando-se na idéia de que risco é o desvio não desejável de um valor de referência (*benchmark*) e pressupondo que um agente econômico é consistente (no sentido a ser definido) no uso da sua função de avaliação de qualquer variável aleatória. Podemos, assim, atribuir um valor real a qualquer variável aleatória e escolher aquela de maior valor.

3. O  $EMD_n$  (equivalente monetário determinístico<sup>9</sup>) de grau  $n$  aproxima o equivalente certo de um ativo e tem uma forma “elegante” que deixa explícito os graus de aversão à incerteza (seção 6.4). Essa abordagem pode ser usada no problema de incerteza sobre o modelo quando a função de utilidade do investidor é definida, de forma que se escolhe a distribuição de maior  $EMD$ . De forma alternativa, quando o modelo (a distribuição) é dado, pode-se usar o  $EMD$  para escolher a utilidade mais “adequada”.

Finalmente, vale a pena mencionar que, ao longo do trabalho, discutiremos, entre outros, os seguintes pontos interessantes:

- A inconsistência entre a medida de risco *perda esperada* (seção 6.3.2) e a dominância estocástica.
- A diferença entre a formulação das medidas de risco *perda esperada* e *expected regret* (seção 6.3.3).
- A inconsistência entre a simples preferência por momentos e a utilidade esperada (seção 6.4.3).

---

<sup>9</sup>Para ser mais exato, deveríamos chamar o conceito de *equivalente monetário determinístico estático*, mas para não criar confusão com a literatura e como a diferença não é o foco do trabalho atual, usaremos o termo tradicional.

# Capítulo 2

## Modelo básico e apresentação do exemplo paradigmático

### 2.1 O exemplo de uma árvore trinária

Ao longo desse trabalho, será usado um exemplo simples para visualizar, de um lado, as diferenças entre as abordagens existentes e, do outro, os resultados das novas propostas. O exemplo é o de uma árvore trinária uniperiódica em  $\{t_0, t_1\}$  com dois ativos financeiros, um com incerteza  $S$  e outro sem,  $B$ . Os estados da natureza em  $t_1$  são dados por  $\Omega_{t_1} = \{\omega_1, \omega_2, \omega_3\}$  com as probabilidades  $\mathbb{P}_{t_1} = (p_1, p_2, p_3) = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ . Com  $\mathcal{F}_{t_1} = \sigma(\Omega_{t_1})^1$  e  $\mathcal{F}_{t_0} = \{\Omega_{t_1}, \emptyset\}$ , temos então o espaço probabilístico  $(\Omega_{t_1}, \mathcal{F}_{t_1}, \mathbb{P}_{t_1})$  e uma filtração  $\mathbb{F} = (\mathcal{F}_{t_0}, \mathcal{F}_{t_1})$ , o que pode ser visualizado na árvore trinária da figura 2.1.

Os valores dos ativos nos diferentes estados são:<sup>2</sup>

$$B_{t_0} = B_{t_1} = 1,$$

$$S_{t_0} = s_0 = 100,$$

$$S_{t_1} = S_{t_1}(\{\omega_1, \omega_2, \omega_3\}) = \begin{pmatrix} S_{t_1}(\omega_1) \\ S_{t_1}(\omega_2) \\ S_{t_1}(\omega_3) \end{pmatrix} = \begin{pmatrix} 120 \\ 100 \\ 90 \end{pmatrix}.$$

---

<sup>1</sup>Observe que nesse caso  $\sigma$  é a sigma-álgebra e não o desvio-padrão de uma variável aleatória.

<sup>2</sup>A condição de o valor do ativo  $B$  em  $t_1$  não depender do estado permite a interpretação como ativo “livre de incerteza/risco” e o fato de  $B_{t_0} = B_{t_1}$  implica que a taxa “livre de incerteza/risco” seja zero.

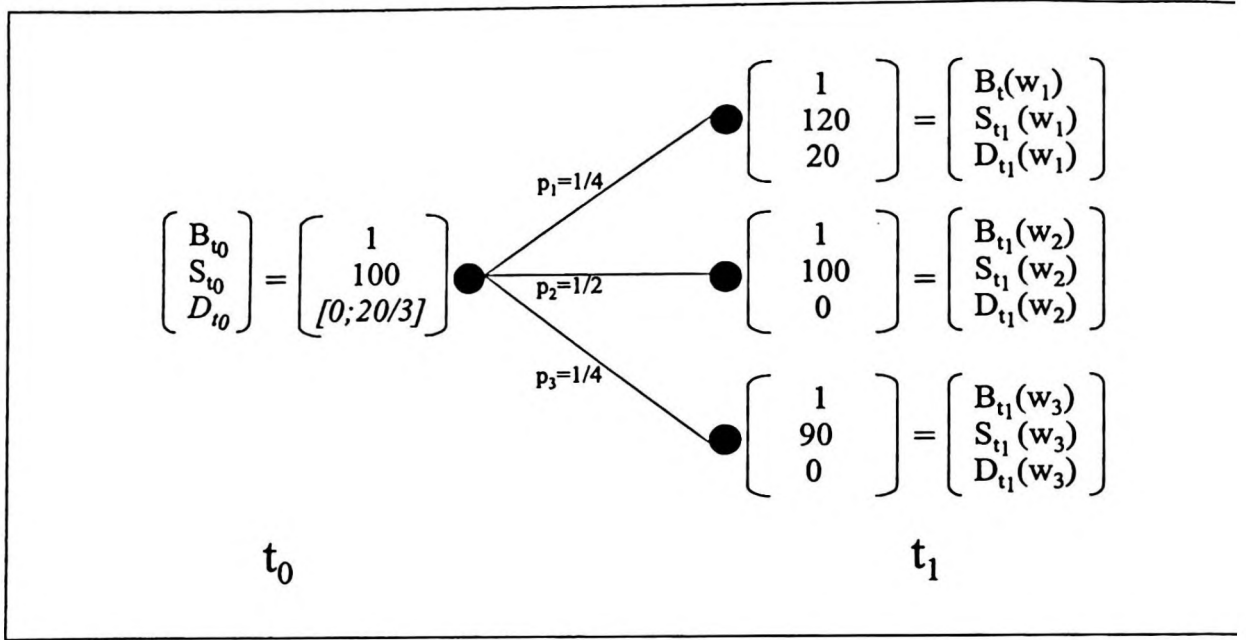


Figura 2.1: Árvore trinomial visualizando o mundo paradigmático.

Nesse mercado, queremos determinar o valor em  $t = t_0$ , do seguinte derivativo do  $S$

$$D_t = (S_t - 100)^+ = \begin{pmatrix} D_{t_1}(\omega_1) \\ D_{t_1}(\omega_2) \\ D_{t_1}(\omega_3) \end{pmatrix} = \begin{pmatrix} 20 \\ 0 \\ 0 \end{pmatrix}.$$

Para determinar o preço livre de arbitragem em  $t = t_0$  precisamos de uma medida neutralizadora de risco  $\mathbb{Q} = \mathbb{Q}_{t_1}$  com a qual podemos determinar

$$D_{t_0} = E_{\mathbb{Q}}[D_{t_1}].$$

O primeiro FTAP nos garante que existe ao menos uma medida  $\mathbb{Q}$  se o mercado é livre de arbitragem<sup>3</sup>. Isso é equivalente ao fato que o

$$\mathcal{M}(S) := \{\mathbb{Q} \sim \mathbb{P}_{t_1} \mid S_t = E_{\mathbb{Q}}[S_s \mid \mathcal{F}_t], \forall t \leq s\},$$

tem ao menos um elemento, ou seja a sua cardinalidade,  $|\cdot|$ , é ao menos 1:  $|\mathcal{M}(S)| \geq 1$ . Na linguagem da álgebra linear, a condição de não-arbitragem nos garante que o sistema de equações lineares (2.1) é *consistente*, enquanto o segundo FTAP nos garante que a solução do sistema é única se a matriz de futuros pagamentos  $[S_{t_1}, B]'$  tem *posto completo*.

<sup>3</sup>No nosso exemplo, básico a prova do FTAP 1 reduz-se à aplicação do lema de Farkas (vide ROSS, 1978) e no caso geral, à separação de espaços por hiperplanos, DUFFIE (1996).

O grande problema no mercado incompleto é escolher uma medida  $\mathbb{Q} \in \mathcal{M}(S)$ . Mas, mesmo que tenhamos infinitas alternativas, elas ainda são limitadas por criarem preços que excluem arbitragem. Calcularemos esses limites em seguida e já observamos que qualquer proposta terá que criar preços dentro desse intervalo.

Para que os processos de preço sejam martingales sob uma medida  $\mathbb{Q} = (q_1, q_2, q_3)$ , tem que valer

$$\begin{pmatrix} S_{t_1} \\ B_{t_1} \end{pmatrix}' \cdot q = \begin{pmatrix} S_{t_0} \\ B_{t_0} \end{pmatrix}, \quad (2.1)$$

$$\begin{pmatrix} 120 & 100 & 90 \\ 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} 100 \\ 1 \end{pmatrix}, \quad (2.2)$$

junto com a restrição de serem probabilidades, i.e.,  $q_i \geq 0, i = 1, 2, 3$ . Segue então que  $q_1 = \frac{1}{2}q_3, q_2 = 1 - \frac{3}{2}q_3, q_3 \in [0; \frac{2}{3}]$ . Escolhendo os valores “extremos” do intervalo, segue que  $D_{t_0} \in [0; \frac{20}{3}]$ . Se o agente achar que  $\frac{1}{2} = q_3$  é o preço de Arrow-Debreu *razoável* pelo estado  $\omega_3$ , ele determinará o preço *racional* do derivativo  $D$  como  $D_{t_0}(\{\frac{1}{4}; \frac{1}{4}; \frac{1}{2}\}) = \frac{15}{3} = 5$ . Por causa da maior importância atribuída às realizações não-desejadas da natureza (peso elevado atribuído ao estado  $\omega_3$ ), o preço estará mais perto do preço máximo de não-arbitragem.

A figura 2.2 visualiza o problema de escolha de uma medida de precificação na árvore trinária, com  $t_1 = T$ . Plotamos os pagamentos do derivativo nos três futuros estados e duas funções lineares de precificação. Se o pagamento do derivativo fosse uma função linear do ativo  $S$ , os pontos  $\{D_T(\omega_i)\}_{i=1,2,3}$  alinhar-se-iam numa reta e existiria uma única função (linear) de precificação. A inclinação da reta daria a razão de *hedge*<sup>4</sup>. Da mesma forma, se o mercado fosse completo (ou seja, o estado  $\omega_2$ , por exemplo, não existisse) usar-se-ia a reta que conecta os pagamentos dos dois outros estados.

---

<sup>4</sup>*Hedging*= Atividade de compôr uma carteira de ativos financeiros que visa criar uma proteção contra as flutuações provocadas pelas fontes de incerteza.

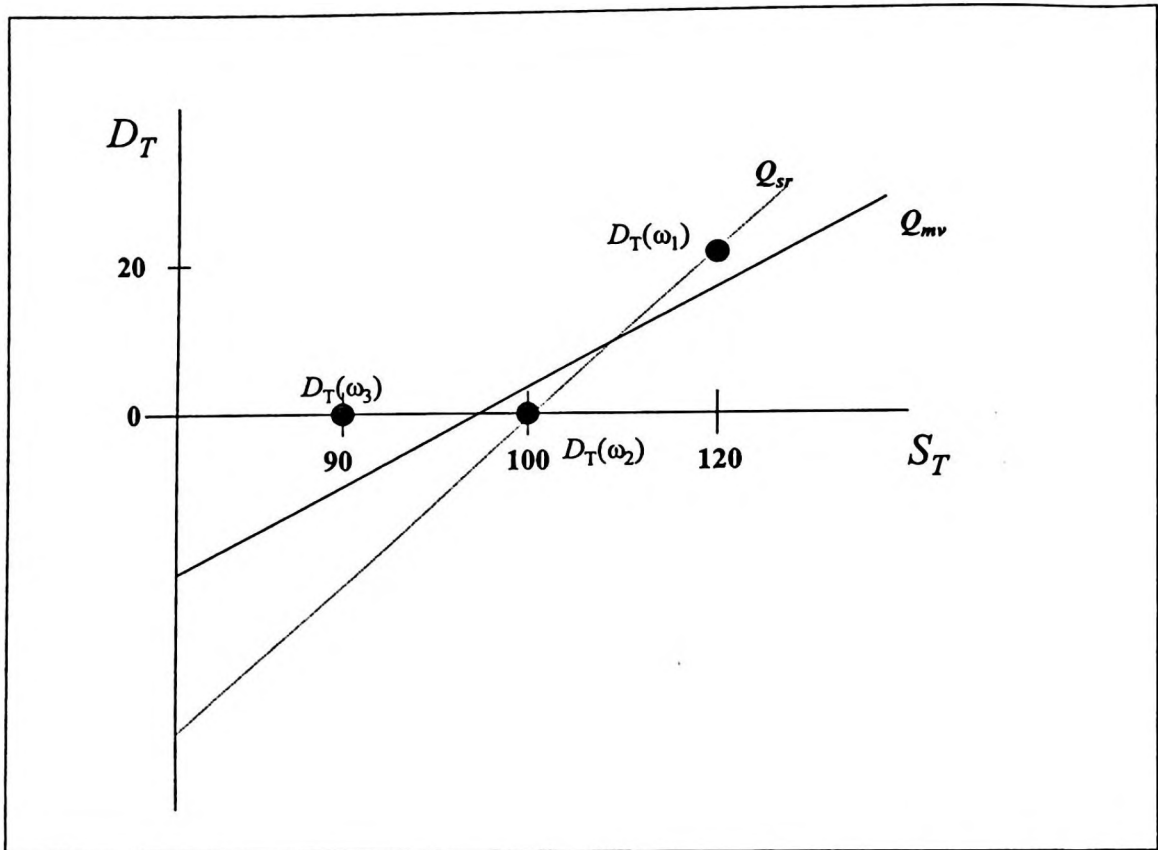


Figura 2.2: O valor do derivativo na data final  $t_1 = T$  em dependência do valor do ativo subjacente  $S$ .

No entanto, se o mercado é incompleto e o pagamento é não-linear, precisamos escolher uma função linear que “melhor aproxima” os pontos dos futuros pagamentos:

$$\min_{a, b \in \mathbb{R}} L(a + b \cdot S - D),$$

sendo  $L$  uma função de penalidade pelos desvios ocorridos.

Na figura 2.2, visualizamos os casos de super-replicação (norma máxima) e do *hedge* média-variância (norma quadrática).<sup>5</sup>

Para análises gerais das propostas, um exemplo não é suficiente. Por isso, no nosso modelo básico escolhido, os retornos relativos seguem uma distribuição normal. Existem duas razões para esta escolha:

1. O formalismo matemático no tempo contínuo é mais elegante.
2. Modelos não-markovianos precisam de cuidados matemáticos que desviam a atenção dos nossos problemas inerentes às abordagens propostas.

<sup>5</sup>Na seção 4.3, discutiremos diversas abordagens existentes.



## 2.2 Um mercado incompleto em tempo contínuo

Nessa seção, definimos os nossos objetos no mercado financeiro e introduzimos as noções importantes que serão úteis em diversas ocasiões ao longo do trabalho, vide KARATZAS; SHREVE (1991) para detalhes.

Num espaço probabilístico filtrado e ( $\mathbb{P}$ -) completo  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F} = \{\mathcal{F}_t\}_{t \in [0, T]})$ , definimos um processo livre de incerteza  $B$  (depósito financeiro, por exemplo) e os processos dos  $n \in \mathbb{R}^+$  ativos com incerteza  $S^{(1)}, \dots, S^{(n)}$  como:

$$\begin{aligned} dB_t &= B_t r_t dt, \quad B_0 = 1 \\ dS_t^{(i)} &= S_t^{(i)} \cdot \left( \mu_t^{(i)} + \sum_{j=1}^d \sigma_t^{(i,j)} dW_t^{(j)} \right), \quad S_0^{(i)} = s_i > 0, i = 1, \dots, n, \\ d &\geq n + 1. \end{aligned} \quad (2.3)$$

$W = (W^{(1)}, \dots, W^{(d)})'$  é um processo de Wiener  $d$ -dimensional (e  $\mathbb{F}$  é a  $\mathbb{P}$ -aumentação da filtração,  $\mathcal{F}_t^W := \sigma(W_s, 0 \leq s \leq t)$ ,  $0 \leq t \leq T$ , gerada pelo  $W$ ),  $r = \{r_t\}_{0 \leq t \leq T}$  a taxa livre de incerteza,  $\mu = \left\{ (\mu^{(1)}, \dots, \mu^{(n)})' \right\}_{0 \leq t \leq T}$  os retornos dos ativos com incerteza e  $\Sigma = (\sigma_t)_{0 \leq t \leq T} = \left\{ (\sigma_t^{(i,j)})_{1 \leq i \leq n, 1 \leq j \leq d} \right\}_{0 \leq t \leq T}$  uma matriz de volatilidade. Pressupomos que todos os processos são progressivamente mensuráveis com respeito à  $\mathbb{F}$ . Pressupomos também que existe uma solução forte única pela equação (2.3) e que a matriz de volatilidade é inversível quase em todo  $[0, T] \times \Omega$ . Se, além disso,  $r, \mu, \Sigma$  e  $\Sigma^{-1}$  são uniformemente limitados em  $(t, \omega) \in [0, T] \times \Omega$ , o processo do prêmio relativo de incerteza

$$\theta(t) := \Sigma^{-1}(\mu_t - r_t \mathbf{1}), \quad 0 \leq t \leq T,$$

com  $\mathbf{1} = (1, \dots, 1)' \in \mathbb{R}^n$  é limitado também. Esse modelo é livre de arbitragem se o posto de  $\Sigma(t, \omega)$  for  $d$  para quase todos os  $(t, \omega)$ , vide ØKSENDAL (2000), teorema 12.2.5 e comentário em seguida.

O processo  $Z = (Z_t)_{0 \leq t \leq T}$ , definido como

$$Z_t := \exp \left( - \int_0^t \theta'(s) dW(s) - \frac{1}{2} \int_0^t \|\theta(s)\|^2 ds \right), \quad 0 \leq t \leq T,$$

é um  $\mathbb{P}$ -martingale local e por ser positivo também é um  $\mathbb{P}$ -supermartingale. Mais ainda, como  $\theta$  é uniformemente limitado,  $Z$  é também um  $\mathbb{P}$ -martingale. Com esse processo de preço-estado, pode ser definida uma nova medida equivalente à  $\mathbb{P}$  da seguinte forma:

$$\mathbb{Q}(A) := E_{\mathbb{P}}[Z_T \cdot \mathbf{1}_A], \quad A \in \mathcal{F}.$$

Descontando os preços dos ativos com incerteza com o ativo livre de incerteza<sup>6</sup>, usando  $\gamma = (\gamma(t))_{0 \leq t \leq T}$  com

$$\gamma_t := \frac{1}{B_t}, \quad 0 \leq t \leq T,$$

temos que os  $\gamma S^{(1)}, \dots, \gamma S^{(n)}$  são  $\mathbb{Q}$ -martingales e que o processo  $\widetilde{W} = (\widetilde{W}_t)_{0 \leq t \leq T}$ , dado por

$$\widetilde{W}_t := W_t + \int_0^t \theta(s) ds, \quad 0 \leq t \leq T,$$

é um processo de Wiener em relação a  $(\mathbb{Q}, \mathbb{F})$  segundo o teorema de Girsanov (KARATZAS; SHREVE, 1991, teorema 5.1).

Para o restante do trabalho, limitar-nos-emos ao caso de  $Z$  ser um  $\mathbb{P}$ -martingale, o que será feito pressupondo que os coeficientes de (2.3) são constantes no tempo:  $\mu_t^{(i)} \equiv \mu^{(i)} \in \mathbb{R}$ ,  $\sigma_t^{(i,j)} \equiv \sigma^{(i,j)} \in \mathbb{R}$ . Uma análise detalhada desse modelo é comum na literatura, vide KARATZAS; SHREVE (1991), e não será feita aqui. A razão de usar coeficientes constantes é que será satisfeita a condição (mais fraca) de Novikov (KARATZAS; SHREVE, 1991, p. 199), que por sua vez já implica que a equação diferencial estocástica 2.3 tem solução (forte) única.<sup>7</sup> Nesse mercado, pressupomos que existe um *pequeno investidor*<sup>8</sup> que pode decidir em cada instante de tempo  $t \geq 0$ :

- quanto dinheiro  $\pi_t^{(i)}$  ele quer investir em cada ativo  $S^{(i)}$ ,  $i = 1, \dots, n$
- quanto dinheiro  $C_t$  ele quer consumir até  $t \geq 0$  (consumo cumulativo).

Como é realístico pressupor que ele pode tomar as decisões somente com as informações atuais, i.e., sem conhecer o futuro, especificamos os processos da seguinte maneira:

**Definição 2.1 (Processo de consumo cumulativo)** *Um processo  $C = (C_t)_{t \in [0, T]}$ ,  $\mathbb{F}$ -adaptado, com*

1. trajetórias crescentes e contínuas à direita, e
2.  $C_{t_0} = 0$  quase certamente (q.c.)

*é chamado de processo de consumo cumulativo.*

<sup>6</sup>Ou com qualquer numerário.

<sup>7</sup>Vide KARATZAS; SHREVE (1991, p. 281).

<sup>8</sup>Ou seja, um investidor que não influencia o preço do mercado.

**Definição 2.2 (Processo de investimento)** *Um processo  $\mathbb{F}$ -adaptado*

$\pi = \left( \pi_t^{(1)}, \dots, \pi_t^{(n)} \right)_{t \in [0, T]}$  *é chamado de processo de investimento se vale*

$$\sum_{i=1}^n \int_0^T \pi_s^{(i)} ds < \infty, \quad q.c.$$

A carteira criada por uma estratégia  $\pi = (\pi^{(1)}, \dots, \pi^{(n)})'$  tem um valor  $X_t$  em cada instante de tempo  $t \in [0, T]$  que é também um processo estocástico:

**Definição 2.3 (Processo de valor)** *Por um capital inicial  $x \in \mathbb{R}$  e um processo  $(\pi, C)$ , o processo  $X_t \equiv X_t^{x, \pi, C}$ ,  $t \in [0, T]$ , com  $X_0 = x$  e*

$$dX_t = \sum_{i=1}^n \pi_t^{(i)} \frac{dS_t^{(i)}}{S_t^{(i)}} + \left( X_t - \sum_{i=1}^n \pi_t^{(i)} \right) \frac{dB_t}{B_t} - dC_t \quad (2.4)$$

$$= r_t X_t dt + \pi_t \cdot [(b_t - r_t) dt + \sigma_t dW_t] - dC_t \quad (2.5)$$

*é chamado de processo de valor (da carteira) correspondente ao capital inicial  $x$ , o processo de investimento  $\pi$  e o consumo  $C$ .*

Por conveniência, chamamos o processo  $\varphi = (x, \pi, C)$  de processo de aplicação.

**Definição 2.4 (Investimento admissível e processo de ganho)** *Dizemos que um processo de investimento  $\pi$  é admissível se o processo associado de ganhos descontados*

$$G_t^\pi := \int_0^t \gamma_s \pi_s dS_s, \quad t \geq 0, \quad (2.6)$$

*é quase certamente limitado inferiormente por uma constante real  $K \in \mathbb{R}$ :*

$$\mathbb{P}(G_t^\pi \geq K, \forall t \geq 0) = 1.$$

A admissibilidade pode ser interpretada como a exigência de uma reserva de capital finita e ela evita estratégias de martingales. Neste caso, escrevemos  $\pi \in \mathcal{A}(x) := \mathcal{A}_K(x)$ .

**Observação 2.1** *No nosso caso de difusão, qualquer processo de investimento limitado e Borel-mensurável é admissível.*

Em um modelo geral, pode ser vantajoso basear a definição de  $\mathcal{M}(S)$  nos processos de ganho criados pelas estratégias  $\pi$ :

$$\mathcal{M}(S) := \{ \mathbb{Q} \sim \mathbb{P} \mid G \text{ é } (\mathbb{F}, \mathbb{Q}) - \text{martingale local} \}.$$

Dependendo de como se comporta o processo de preços  $S$  sob as medida equivalentes  $\mathbb{Q}$ , temos as seguintes condições para que o conjunto  $\mathcal{M}(S)$  não seja vazio (SCHACHER-MAYER, 2004a):

- Se  $S$  for um  $(\mathbb{F}, \mathbb{P})$  –semi-martingale<sup>9</sup>,  $S$  terá que ser um  $(\mathbb{F}, \mathbb{Q})$ -sigma-martingale<sup>10</sup> (vide DELBAEN; SCHACHERMAYER, 1998).
- Se  $S$  for (além de ser semi-martingale) localmente limitado,  $S$  terá que ser  $(\mathbb{F}, \mathbb{Q})$ -martingale local<sup>11</sup>.
- Se  $S$  for uniformemente limitado,  $S$  terá que ser  $(\mathbb{F}, \mathbb{Q})$  –martingale.

Mas, com o uso da condição de  $|\Omega| < \infty$  ou  $S$  limitado, as noções são iguais. A definição com o processo de ganho  $G$  é mais genérica, pois o conjunto de processos de investimentos  $\pi$  será ajustado à definição de  $S$  para que o ganho seja um martingale local.<sup>12</sup> No caso tratado nesse trabalho, de  $S$  sendo uma difusão, temos que  $G$  é um martingale. Mesmo assim preferimos introduzir as noções gerais porque:

- os resultados de dualidade para funções de utilidade com domínio  $\mathbb{R}$  são formulados em termos gerais (vide seção 4.2.4)
- a abordagem de dualidade em mercados financeiros incompletos é pouco difundida na literatura brasileira, o que nos levou a explicar a teoria em maiores detalhes usando conceitos gerais.

---

<sup>9</sup> $S$  é um semi-martingale se existe um martingale local  $M$  e um processo finito  $A$  tal que  $S = M + A$ .

<sup>10</sup> $S$  é um sigma-martingale se existe um martingale  $M$  e um processo  $\phi$  positivo,  $M$ –integrável e previsível tal que  $S = \phi \cdot M$ .

<sup>11</sup>Um martingale é dito local quando a condição de martingale vale somente para seqüências de tempo de parada. Obviamente, todo martingale é martingale local, mas não vale o contrário.

<sup>12</sup>Para uma análise geral de integrais estocásticas, fazemos referência a ROGERS; WILLIAMS (1987).

# Capítulo 3

## Critérios de decisão sob incerteza

Na precificação de ativos financeiros que não podem ser replicados por outros ativos no mercado, usam-se funções de utilidade e/ou de risco para atribuir um valor ao risco imanente dos ativos. Após uma discussão e diferenciação dos termos incerteza, risco e ambigüidade, seguimos a idéia de que a escolha de uma função de utilidade ou de uma medida de risco se reduz à escolha de uma medida de distância.

Além de introduzir uma definição coerente dos termos envolvidos, será proposto um ponto de vista abstrato e geral sobre a escolha do critério de decisão.

### 3.1 Incerteza, risco e ambigüidade

Quando se trata de incerteza e medidas de risco, surge o problema da falta de uma exata definição dos termos, o que se agrava no caso do termo ambigüidade. Tendo o problema da escolha entre variáveis aleatórias, o investidor que não gosta de incerteza escolheria a distribuição “mais degenerada”. Mesmo assim, o risco pode estar presente se o valor referencial ficar acima do centro de massa desta distribuição degenerada. Assim chegamos à seguinte conclusão heurística:

**Afirmção 3.1** *Quanto mais concentrada a distribuição, menos incerteza ela contém.*

Isso implicaria que a distribuição uniforme é a mais *incerta*, mas do ponto de vista de risco, ela não precisa ser a que contém mais *risco*. De fato, quando atentamos para uma variável aleatória sem ter um valor de referência, não podemos identificar o *risco*.

Tradicionalmente, um ambiente de incerteza é dado quando não existem valores objetivos para as probabilidades dos estados da natureza (e.g. KNIGHT, 1921). Um ambiente

de risco, nesse caso, seria identificado de forma que a variável sob análise é estocástica: decisões são tomadas sob incerteza quando não foram atribuídas probabilidades aos estados da natureza. Somente os chamados *atos*, que implicam *conseqüências*, são conhecidos. *Atos* são mapeamentos do espaço de estados para o espaço de conseqüências, vide SAVAGE (1954).

Observação 3.2 KARNI; SCHMEIDLER (1991, p. 1766) descrevem o modelo de decisão sob risco como

1. o conjunto de estados contém somente um elemento<sup>1</sup>, ou
2. todos os atos são constantes e o conjunto de conseqüências é constituído por medidas de probabilidades.<sup>2</sup>

Ao contrário, decisão sob incerteza ocorre quando o conjunto de atos inclui atos não constantes. Esta noção tradicional forçar-nos-ia a chamar a parte não desejável da distribuição de *perda*, por exemplo, e não de risco. Mas, do nosso ponto de vista, o termo *risco* é mais adequado pelas seguintes razões<sup>3</sup>:

1. Em finanças, a parte não desejável de uma distribuição é indentificada como *risco*. Se a variável aleatória é estritamente positiva, não teríamos perdas em termos absolutos, mas somente de forma relativa como desvio do valor de referência.
2. No nosso trabalho, começamos já com um espaço probabilístico  $(\Omega, \mathcal{F}, \mathbb{P})$  e não precisamos mais diferenciar o mundo sem medida de probabilidade. Dizendo que, no ambiente  $(\Omega, \mathcal{F}, \mathbb{P})$ , a decisão é sob risco, implica que não temos nenhum termo para descrever partes não desejáveis da distribuição.
3. Quando nos confrontarmos com instrumentos chamados de *livre de risco* em finanças, encontraremos um dilema, pois um ativo livre de risco pode ser visto como uma variável com distribuição degenerada. Mas, como existem probabilidades atribuídas, estamos num ambiente de decisão sob *risco*.

---

<sup>1</sup> *Singularidade* de um conjunto significa que ele contém somente um elemento.

<sup>2</sup> Ou, na linguagem dos autores, constituído pelas loterias no conjunto de resultados.

<sup>3</sup> Este conceito é seguido também por BALZER (1994) que chama o que nós consideramos como risco de "*risco de investimento*".

4. O termo “*Perda*” é usado para valores monetários, enquanto *risco* reforça melhor a conexão com a distribuição.

Por isso, adotamos aqui as definições que serão usadas no restante do trabalho:

- O termo ambigüidade é usado quando não sabemos qual a distribuição da variável aleatória (antiga definição do ambiente de incerteza).
- O termo de incerteza é usado quando as variáveis são estocásticas (antiga definição do ambiente de risco).
- O termo risco é usado quando há a possibilidade de obter resultados abaixo de um valor de referência (*benchmark*) estabelecido (essa definição não é existente na definição tradicional).

A abordagem típica para avaliar matematicamente e economicamente a incerteza é a de usar funções de utilidade, vide NEUMANN; MORGENSTERN (1953). No caso de ambigüidade, a referência clássica é a de SAVAGE (1954). A teoria de ANSCOMBE; AUMANN (1963) usa também ambigüidade, mas não pressupõe completa ignorância sobre as distribuições. A relação entre risco e incerteza foi abordada, entre outros, por HOLTON (1997), que define risco como “*a exposição à incerteza*” e diferencia risco de incerteza. Por ele, risco pode ser representado por uma variável aleatória que representa a exposição e um espaço probabilístico que representa incerteza. Essa notação foi também adotada por BOYLE; SIU; YANG (2002). Observa-se que, quando se trata do conceito de *risco*, foi popularizado por ARTZNER et al. (1997) a idéia de atribuir a uma distribuição um valor real, de forma que este valor pode ser interpretado como o dinheiro adicional necessário para se livrar do risco (seguro), vide seção 3.2.3 com detalhes.

A falta de conhecimento sobre a distribuição (ambigüidade) terá importância na análise de alocação ótima de ativos, criando uma forma de incompletude no mercado e duas formas desse problema serão tratadas no nosso trabalho (seção 5 e seção 6). Finalmente, pode-se visualizar a relação entre os conceitos de incerteza e risco segundo o nosso entendimento na figura 3.1.

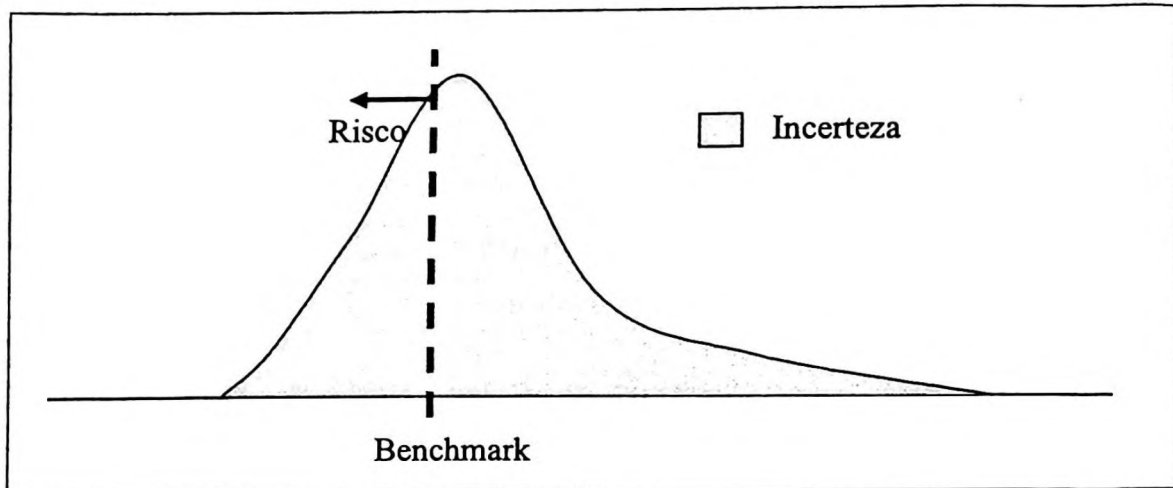


Figura 3.1: A distinção entre risco e incerteza

### 3.1.1 Medindo incerteza

#### Incerteza atribuída a um espaço mensurável

De um ponto de vista puramente probabilístico, incerteza é uma característica do espaço mensurável. A medida (probabilística) é a ferramenta natural para avaliar incerteza. Ela atribui a cada estado da natureza uma probabilidade de ser realizado. O caso extremo seria a existência de um estado único que teria peso total.

**Definição 3.1 (Existência de incerteza)** Dizemos que o nosso modelo  $(\Omega, \mathcal{F}, \mathbb{P})$  contém incerteza se existe **nenhum** átomo  $A \in \mathcal{F}$  tal que  $\mathbb{P}(A) = 1$ .

Lembramos que um elemento  $A \in \mathcal{A}$  é um átomo se  $\mathbb{P}(A) > 0$  e para todos  $A \supset B \in \mathcal{F}$ , com  $\mathbb{P}(B) > 0$ , vale que  $\mathbb{P}(A) = \mathbb{P}(B)$ . Se  $\Omega$  for átomo, o sistema é determinístico.

Pergunta-se então: qual seria a circunstância que nos levaria a dizer que uma “situação” tem mais incerteza que outra? Identificando as situações com  $\sigma$ -álgebras  $\mathcal{F}$  e  $\mathcal{G}$ , podemos pensar em calcular as cardinalidades delas e compará-las  $|\mathcal{F}| \stackrel{?}{=} |\mathcal{G}|$ . O problema é que a cardinalidade de uma  $\sigma$ -álgebra muitas vezes é infinita e não serve para comparações. Como alternativa, podemos pensar em considerar que se uma  $\sigma$ -álgebra é uma sub- $\sigma$ -álgebra de uma outra, i.e., podemos identificar se  $\mathcal{G} \subseteq \mathcal{F}$ . Infelizmente, esse caso é raro e teremos uma ordenação não-total.

Por causa desses problemas, propomos analisar o conjunto  $\Omega$ . Podemos nos perguntar quantos elementos básicos de  $\Omega$  são necessários para conseguir probabilidade total,



i.e.,  $\text{Inc}(\Omega) = \inf \{\#\omega \in A : \omega \text{ é átomo e } \mathbb{P}(A) = 1\}$ . Podemos generalizar isso para  $\sigma$ -álgebras:

$$\text{Inc}(\mathcal{F}) := \{\#A \in \mathcal{F} : \mathbb{P}(A) = 1, A \text{ átomo}\}$$

A vantagem desse conceito é que podemos identificar se a situação dada por  $\Omega_1 = (0, 1]$  é mais ou menos incerta que  $\Omega_2 = (0, 2]$ .<sup>4</sup> A importância desse caso é que podemos comparar variáveis aleatórias definidas em espaços probabilísticos diferentes. Para podermos então comparar as incertezas, introduzimos a seguinte definição:

**Definição 3.2 (Incerteza de uma  $\sigma$ -álgebra e de espaços de eventos)** *Dizemos que o espaço de eventos  $\Omega_1$  é mais incerto que o espaço de eventos  $\Omega_2$  se*

$$\text{Inc}(\Omega_1) \geq \text{Inc}(\Omega_2).$$

*Dizemos que a  $\sigma$ -álgebra  $\mathcal{F}_1$  é mais incerta que a  $\sigma$ -álgebra  $\mathcal{F}_2$ , (as duas definidas no mesmo espaço  $\Omega$ ), se*

$$\text{Inc}(\mathcal{F}_1) \geq \text{Inc}(\mathcal{F}_2).$$

Na próxima seção, descreveremos como usar esse conceito para variáveis aleatórias.

### Incerteza de uma variável aleatória

Propomos usar o conceito de incerteza para caracterizar uma variável aleatória da seguinte forma:

#### Conjectura 3.1

1. *Quanto maior a incerteza do conjunto de eventos elementares, maior a incerteza de  $X$ .*
2. *Quanto maior a cardinalidade da  $\sigma$ -álgebra criada por  $X$ , mais incerta é  $X$ .*

Mas, além disso, existe um aspecto de variáveis aleatórias que não consideramos ainda: a concentração dos valores que a variável pressupõe, ou seja, a forma dela. Para explicar essa diferença, usamos o seguinte exemplo:

---

<sup>4</sup>No caso de uso com  $\sigma$ -álgebras, podemos dizer se  $\mathcal{B}(\Omega_1)$  é mais incerta que  $\mathcal{B}(\Omega_2)$ .

**Exemplo 3.1** (Incerteza de espaços probabilísticos e distribuições) *Considere as seguintes variáveis*

$$\begin{aligned} X_1 &: (\Omega_1, \mathcal{F}_1, \mathbb{P}_1) \rightarrow (\{1, 2, 3\}, \mathcal{P}(\{1, 2, 3\}), U_{\{1,2,3\}}), \\ X_2 &: (\Omega_2, \mathcal{F}_2, \mathbb{P}_2) \rightarrow (\{1, 2, 3\}, \mathcal{P}(\{1, 2, 3\}), U_{\{1,2,3\}}), \end{aligned}$$

sendo que  $U_{\{1,2,3\}}$  é a medida uniforme em  $\{1, 2, 3\}$ , e  $\Omega_1 = \{\omega_1, \omega_3\}$ ,  $\Omega_2 = \{\omega_1, \omega_2, \omega_3\}$ .

Definimos  $X_1$  e  $X_2$  como

$$\begin{aligned} X_1(\omega) &= \begin{cases} 1 & \omega = \omega_1, \\ 3 & \omega = \omega_3, \end{cases} \\ X_2(\omega) &= \begin{cases} 1 & \omega = \omega_1, \\ 2 & \omega = \omega_2, \\ 3 & \omega = \omega_3, \end{cases} \end{aligned}$$

com

$$\begin{aligned} \mathbb{P}_1(\{\omega \in \Omega_1 \mid X_1(\omega) = 1\}) &= \mathbb{P}_1(\{\omega \in \Omega_1 \mid X_1(\omega) = 3\}) = \frac{1}{2}, \\ \mathbb{P}_2(\{\omega \in \Omega_2 \mid X_2(\omega) = 1\}) &= \mathbb{P}_2(\{\omega \in \Omega_2 \mid X_2(\omega) = 2\}) \\ &= \mathbb{P}_2(\{\omega \in \Omega_2 \mid X_2(\omega) = 3\}) = \frac{1}{3}. \end{aligned}$$

Intuitivamente, essa forma de incerteza ajuda a comparar, por exemplo, um modelo de árvore binomial com um modelo de árvore trinomial.

Quando analisamos a incerteza contida na distribuição de uma variável aleatória  $X$ , não nos preocupamos diretamente com a distribuição  $\mathbb{P}$ , mas com a medida induzida pela variável  $X$ , a distribuição  $\mathbb{P}_X$ . Pressupondo  $X_1, X_2$  como variáveis aleatórias reais em  $(\Omega, \mathcal{F}, \mathbb{P})$ , o conceito de comparar  $\sigma$ -álgebras pode não ser muito sensível à “incerteza” da distribuição. Por exemplo, podemos ter facilmente  $\sigma(X_1) = \sigma(X_2) = \mathcal{F}$ , mas  $X_1$  pode ser uniforme contínua e  $X_2$  normal.

Uma forma tradicional de ordenar distribuições é a *dominância estocástica*, que resulta numa ordenação parcial das variáveis em  $(\Omega, \mathcal{F}, \mathbb{P})$ , usando as funções de distribuição das variáveis.<sup>5</sup>

<sup>5</sup>Vide, por exemplo, HUANG; LITZENBERGER (1988), capítulo 2, e as referências citadas.

### 3.1.2 Medindo dispersão

A dispersão descreve quanto os valores de uma variável aleatória  $X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (E, \mathcal{E}, \mathbb{Q})$  variam no espaço  $E$  dos valores. É natural considerar as duas seguintes características:

1. O tamanho do conjunto imagem - a variedade.
2. A variação nesse conjunto imagem - a variabilidade ou concentração.

O próximo exemplo serve para explicar melhor a diferença desses dois pontos.

**Exemplo 3.2** Na figura 3.2, temos duas distribuições,  $X_1$  (uniforme) e  $X_2$  (normal), sendo que nenhuma domina estocasticamente (de primeiro grau) a outra. A variedade, determinada pelo intervalo dos valores, é diferente para  $X_1$  e  $X_2$ , e segundo ela, a uniforme deve ser preferida, pois não realizará valores infinitos.

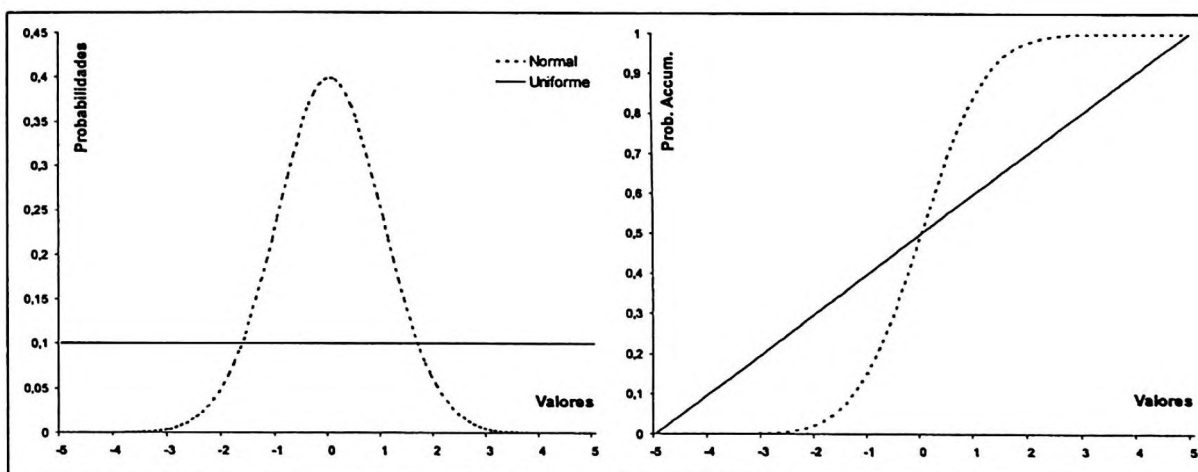


Figura 3.2: Duas variáveis com diferentes variedades e sem dominância estocástica.

Para analisar o tamanho do conjunto de imagem, usamos os conceitos de amplitude (tamanho do intervalo da imagem) e cardinalidade do conjunto de imagem. Para analisar a variabilidade, usamos o conceito de concentração.

**Definição 3.3 (Dispersão)** *Dispersão combina o tamanho do intervalo de uma variável  $X$ ,  $Im(X)$ , com a concentração dos valores.*

O objetivo do restante desta seção é o de formalizar essa idéia.

Um tamanho grande do intervalo da imagem da variável permite uma maior variedade de possíveis valores, o que é uma característica importante da dispersão.

**Definição 3.4 (Amplitude)** *A amplitude de uma variável aleatória  $X$  é definida por*

$$\text{rg}(X) := X_{\max} - X_{\min}.$$

Infelizmente, esta definição não captura ainda o fato de que os valores possam ser concentrados em alguns pontos, mas ajuda a discriminar duas variáveis,  $X_1$  e  $X_2$ , como no exemplo 3.2. Essa medida univariada não é sensível ao “*aumento de risco*” (na terminologia de ROTHSCHILD; STIGLITZ, 1970)<sup>6</sup>. Para conseguir essa sensibilidade, precisamos da noção de concentração de valores.

A concentração é provavelmente o conceito no qual pensa intuitivamente a maioria das pessoas quando pensa na “incerteza” de uma distribuição. Neste ponto, entra a forma da distribuição, que é muitas vezes capturada pela curvatura da função de densidade. Usamos a seguinte (meta-) definição:

**Definição 3.5 (Concentração)** *Dizemos que uma função distribuição  $F_{X_1}$  é mais concentrada que  $F_{X_2}$ , se  $F_{X_1}$  “desvia mais” da distribuição uniforme  $U$  que  $F_{X_2}$ .*

A dominância estocástica ocupa um espaço importante na teoria de comparação de incerteza, pois ela captura a massa nas caudas das distribuições. Mas, é importante lembrar que uma condição importante para a coincidência da concentração visual com o conceito de dominância estocástica é que a distribuição seja unimodal. Na figura 3.3, lado esquerdo, vemos uma figura seguindo o caso normal:  $X_1$  é mais concentrada em torno da média que  $X_2$  e a intuição induz que  $X_2$  é mais incerta que  $X_1$ . Portanto, no lado direito,  $X_1$  é mais concentrada que  $X_2$ , mas, facilmente, classifica-se  $X_1$  como mais incerta.

O ponto crítico é como definir a distância entre  $F_{X_1}$  ou  $F_{X_2}$  e a uniforme. Uma medida importante de concentração em torno da média é a variância, enquanto a entropia (diferencial) mede concentração sem ter uma medida de posição central como referência. A seguinte seção discute algumas maneiras de medir distâncias. Concluindo, é importante ter clareza sobre o que finalmente tem que ser analisado. Tendo a incerteza do espaço original e a variedade igual para duas variáveis aleatórias, falta medir a concentração.

---

<sup>6</sup>Na nossa terminologia, seria um aumento de variabilidade.

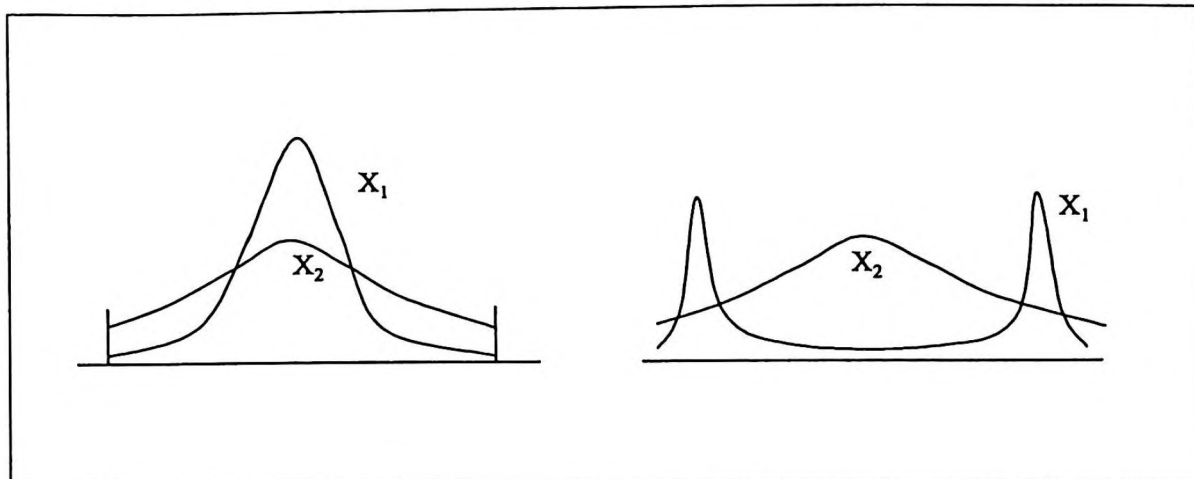


Figura 3.3: A importância de modalidade para a compatibilidade da dominância estocástica e concentração.

## 3.2 Medidas de risco

Nós já mencionamos a diferença entre os conceitos de mensurar risco e incerteza usando como critério a relação do valor da variável aleatória com um *benchmark* para caracterizar o risco. É também possível especificar a direção do desvio do *benchmark*. Essa distinção permite o uso dos termos de risco unilateral e bilateral. Finalmente, depende de como o agente econômico define o *benchmark* e qual parte do desvio ele considera como risco. O foco dessa seção é analisar quais características uma medida (de incerteza ou risco) deve ter para que nós a consideremos como “adequada”.

Começamos com uma variável aleatória  $X$  (interpretada como o retorno de um ativo) no espaço probabilístico  $(\Omega, \mathcal{F}, \mathbb{P})$ , e um funcional  $\rho : \mathcal{X} \rightarrow \mathbb{R}$ , interpretada como uma função de risco aplicada ao resultado  $X \in \mathcal{X}$ , sendo que  $\mathcal{X}$  é o espaço de distribuições em  $(\Omega, \mathcal{F}, \mathbb{P})$ . Chama-se  $\rho$  uma **medida** (ou funcional) **de risco** quando um decisor que é averso ao risco quer minimizá-la, enquanto o funcional  $s : \mathcal{X} \rightarrow \mathbb{R}$  chama-se **medida de segurança** e o decisor quer maximizá-la, segundo PFLUG (1999).

Para um agente econômico com capital inicial  $x \in \mathbb{R}$  confrontado com  $n$  ativos com retornos  $X = (X_1, \dots, X_n)$  e querendo aplicar os seus recursos, o conjunto aleatório

$$\mathcal{R}(\omega) := \left\{ x \cdot \sum_{i=1}^n X_i(\omega) \cdot \lambda_i \mid \sum_{i=1}^n \lambda_i = 1 \right\}$$

representa os retornos absolutos de todas as carteiras *admissíveis* para o agente. Ao contrário das mudanças *relativas*, o nível de capital inicial é importante.  $\mathcal{R}$  é um conjunto

convexo nos ativos que permanece convexo se introduzirmos outras restrições convexas e pode também ser visto como um sub-conjunto do espaço das funções de distribuição, i.e.,

$$\mathcal{R} = \left\{ F(\lambda) \left( \frac{\cdot}{x} \right) \mid X' \lambda \sim F(\lambda), \lambda' \mathbf{1} = 1, X \sim F \right\},$$

com  $X$  o vetor de retornos com função distribuição conjunta  $F_X$  em  $(\Omega, \mathcal{F}, \mathbb{P})$  (vide GIORGI, 2002; PFLUG, 1999). O agente agora aplica uma função  $f$  potencialmente multi-dimensional aos elementos do conjunto  $\mathcal{R}$ . Se a função é bi-dimensional de forma que  $f_1 = \mu$  e  $f_2 = -\rho$  e  $\rho$  é uma medida de risco/incerteza, ele usa somente uma abordagem retorno-risco/incerteza. Um outro exemplo seria o funcional de utilidade esperada:  $s : \mathcal{X} \rightarrow \mathbb{R}, X \mapsto E[U(X)]$ .

É importante lembrar a diferença com as funções de risco/segurança que mapeiam uma variável aleatória em uma outra:

$$R, S : L^p((\Omega, \mathcal{F}, \mathbb{P}); (\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)) \rightarrow L^p((\Omega, \mathcal{F}, \mathbb{P}); (\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)),$$

com  $p \in (0, \infty)$ ,  $\lambda$  a medida de Lebesgue nos conjuntos de Borel em  $\mathbb{R}$ . Se  $S$  satisfaz às condições necessárias, pode ser interpretada como função de utilidade. Este tipo de função distorce os valores das variáveis aleatórias e assim, implicitamente, também as funções distribuição.

Podemos também pensar em aplicar uma função à função distribuição diretamente. Nesse caso, a função aplicada é chamada de função de distorção.<sup>7</sup>

**Definição 3.6 (Função de distorção)** *Interpretamos a função  $g : [0, 1] \rightarrow [0, 1]$ , decrescente com  $g(0) = 1, g(1) = 0$ , quando aplicada à função distribuição  $F$  como uma função de distorção.*

Como  $g$  “distorce” as probabilidades atribuídas aos eventos, criamos uma outra medida no espaço probabilístico  $\mu(A) := g(\mathbb{P}(X \in A))$ , com  $A \in \mathcal{B}(\mathbb{R})$ . Podemos então medir a “incerteza distorcida de  $X$ ” por

$$M_g(X) := E_g[X] := E_\mu[X] = \int_{\mathbb{R}} g(\mathbb{P}(X \leq x)) dx.$$

Na área de seguros, o prêmio é calculado usando somente a parte negativa das realizações o que implica que

$$M_g(X) = \int_{\mathbb{R}_+} g(\mathbb{Q}(X \leq x)) dx,$$

$$\mathbb{Q}(\omega) = \mathbb{P}(\omega) \cdot \mathbf{1}_{\{X(\omega) \leq 0\}}(\omega).$$

<sup>7</sup>Vide, por exemplo, WANG (1996) ou WIRCH; HARDY (2000).

A diferença entre a abordagem de distorção das probabilidades e o uso de funções de risco é que distorcemos probabilidades e não valores criados pela carteira, assim como no caso das funções de utilidade. A conexão entre essas abordagens é formalizada pelo seguinte resultado:

**Lema 3.3 (Conexão entre funções de distorção e de risco)** *Seja  $g$  uma função de distorção. Então existe uma função  $R$  sublinear<sup>8</sup> tal que*

$$\int_{\mathbb{R}_+} g(\mathbb{Q}(X \leq x)) dx = \int R(X) dx.$$

O resultado é uma consequência do teorema da mudança de variável (vide BILLINGSLEY, 1995). A sublinearidade é necessária por causa da restrição aos valores não-negativos. Observamos que a idéia de uma função de distorção é também relacionada a medidas de risco espectrais como proposto por ACERBI (2002). Partindo da idéia de que uma combinação convexa de medidas de risco é também uma medida de risco, o autor introduz uma classe de “distorções” possíveis que possam ser usadas para criar novas medidas de risco.<sup>9</sup>

**Definição 3.7 (Espectro admissível de risco)** *Uma função  $\phi \in \mathcal{L}^1([0, 1])$  é chamada de espectro de risco admissível ou função de aversão à incerteza, se  $\phi$  é positiva, decrescente e tem norma 1.<sup>10</sup>*

**Definição 3.8 (Medida de risco espectral)** *A medida de risco espectral de  $X$  com respeito a  $\phi \in \mathcal{L}^1([0, 1])$  é dada por*

$$M_\phi(X) = - \int_0^1 F_X(x) \phi(x) dx.$$

<sup>8</sup>Uma função  $R : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$  é chamada de sublinear, se vale para qualquer  $X_1, X_2 \in (\Omega, \mathcal{F}, \mathbb{P})$  e  $\lambda \in \mathbb{R}^+$ , que  $R(X_1 + X_2) \leq R(X_1) + R(X_2)$ , e que  $R(\lambda X_1) = \lambda R(X_1)$ .

<sup>9</sup>Usamos a notação  $\mathcal{L}^1([a, b])$  para representar o espaço das funções integráveis no intervalo  $[a, b] \subseteq \mathbb{R}$ .

<sup>10</sup>No uso dos termos *positivo* e *decrescente* temos que ter cuidado, pois eles valem para funções. Nesse caso, usam-se as seguintes noções:

1.  $\phi \in \mathcal{L}^1([a, b])$  é positivo, se  $\forall I \subset [a, b]$

$$\int_I \phi(x) dx \geq 0.$$

2.  $\phi \in \mathcal{L}^1([a, b])$  é decrescente, se  $\forall c \in [a, b], \forall \varepsilon > 0$  com  $[c - \varepsilon, c + \varepsilon] \subset [a, b]$  vale que

$$\int_{c-\varepsilon}^c \phi(x) dx \geq \int_c^{c+\varepsilon} \phi(x) dx.$$

Como simples resultado da teoria de probabilidade segue (vide também lema 3.3):

**Lema 3.4 (Conexão entre funções de distorção e espectro de risco)** *Seja  $g$  uma função de distorção. Então existe um espectro de risco  $\phi \in \mathcal{L}^1([0, 1])$  tal que*

$$\int_{-\infty}^0 g(\mathbb{P}(X \leq x)) dx = - \int_0^1 \mathbb{P}(X \leq x) \phi(x) dx.$$

Esse resultado nos permite usar esquemas de classificação para funções de distorção e medidas espectrais de risco para analisar funções de risco se necessário.

### 3.2.1 Medidas de risco derivadas das distorções de valores

Anotamos aqui alguns casos comuns em que as preferências dos agentes são refletidas nas especificações da função  $R$ , seguindo o esquema de classificação de DANIELSSON et al. (2005):

#### 1. Medidas de discrepâncias favoráveis e desfavoráveis

- Variância:  $\rho(X) := E[X^2]$  (MARKOWITZ, 1952).
- Semi-variância:  $\rho(X) := E[(\min(X, 0))^2]$  (MARKOWITZ, 1959).
- Distância interquantil:  $F_X^{-1}(q_1) - F_X^{-1}(q_2)$ ,  $1 \geq q_1 \geq q_2 \geq 0$ .
- Risco de mercado:  $\rho(X) := \beta := \frac{E[(X - E[X]) \cdot (M - E[M])]}{E[(M - E[M])^2]}$ , sendo que  $M$  é uma variável aleatória que representa um “mercado” de referência.

#### 2. Medidas de discrepâncias desfavoráveis

- Momentos parciais inferiores e *downside risk*:

$$\rho(X) := DR(X) := E[(BM - X)^p \cdot 1_{\{X \leq \text{mar}\}}], p \geq 1,$$

sendo  $BM$  um *benchmark* constante e  $\text{mar}$  um valor mínimo aceitável (BAWA, 1975; FISHBURN, 1977).

- O critério de *value-at-risk*<sup>11</sup> é dado pelo valor esperado de uma função indicadora de  $X$  sendo menor que um valor restritivo,  $VaR$ , a um nível de confiança pré-definido  $\alpha$ :

$$\rho(X) := VaR_\alpha(X) := F_X^{-1}(q) := \inf \{q \in \mathbb{R} : \mathbb{P}(X \leq q) \geq \alpha\}.$$

<sup>11</sup>Para o caso dinâmico, vide FÖLLMER; LEUKERT (1999).



- A esperança condicionada na cauda (*Tail conditional expectation*) (vide EMBRECHTS; KLÜPPELBERG; MIKOSCH, 1997) é definida como o valor esperado das  $\alpha \cdot 100\%$  piores perdas da carteira (quantil à esquerda com probabilidade  $\alpha$ ):

$$\rho(X) := \text{TCE}_\alpha(X) := -E[X | X \leq \text{VaR}_\alpha(X)].$$

Essa abordagem é semelhante ao *downside risk*, após normalização. De fato, com  $\alpha = 1$ ,  $\text{VaR} = 0$ , temos

$$\text{TCE}_\alpha(X) \cdot \mathbb{P}(X \leq \text{VaR}) = -E[(X - \text{VaR}_\alpha(X))^p \cdot 1_{\{X \leq \text{VaR}_\alpha(X)\}}] = \text{DR}(X).$$

- Pior valor esperado condicionado (*Worst conditional expectation*): esta medida, proposta por ARTZNER et al. (1999), é definida como

$$\rho(X) := \text{WCE}_\alpha(X) := \inf \{E[X | A] : \mathbb{P}(A) > \alpha, A \in \mathcal{F}\}.$$

- VaR condicionado (*conditional VaR*): como o infimum na definição de *WCE* não é fácil de se usar em aplicações práticas, foi proposta o VaR condicionado (URYASEV; ROCKAFELLAR, 1999; PALMQUIST; URYASEV; KROKHNAL, 1999):

$$\rho(X) := \text{CVaR}_\alpha(X) := \inf \left\{ \frac{E[X - s]}{\alpha} - s : s \in \mathbb{R} \right\}.$$

No caso que vale

$$\mathbb{P}(X < \text{VaR}_\alpha(X)) > 0,$$

$$\mathbb{P}(X \leq \text{VaR}_\alpha(X)) = \alpha,$$

as medidas VaR condicionado e *TCE* são iguais (vide ACERBI; TASCHE, 2002).

- Perda esperada (*Expected Shortfall*): sendo  $\text{VaR}_\alpha(X)$  o maior valor que *WCE* deve pressupor, esta medida é idêntica ao *CVaR* (vide ACERBI; TASCHE, 2002)

$$\begin{aligned} \rho(X) &= \text{ES}_\alpha(X) \\ &= \frac{E[X \cdot 1_{\{X \leq \text{VaR}_\alpha(X)\}}] + \text{VaR}_\alpha(X) \cdot (\alpha - \mathbb{P}(X \leq \text{VaR}_\alpha(X)))}{\alpha}. \end{aligned}$$

Os autores propuseram esta forma para garantir que no caso de distribuições não-contínuas a medida ainda seja contínua no parâmetro de confiança  $\alpha$ . Se a distribuição for contínua usa-se a forma simplificada

$$\begin{aligned} \text{ES}_\alpha(X) &= -\frac{1}{\alpha} \cdot E[X \cdot 1_{\{X \leq \text{VaR}_\alpha(X)\}}] \\ &= -\frac{E[X \cdot 1_{\{X \leq \text{VaR}_\alpha(X)\}}]}{\mathbb{P}(X \leq \text{VaR}_\alpha(X))} \\ &= -E[X | X \leq \text{VaR}_\alpha(X)]. \end{aligned}$$

- *Expected regret*: fortemente relacionado com os conceitos anteriores de *DR*, *ES* e *CVaR* com  $\alpha = 1$ , o *ER* mede o valor esperado de uma perda acima de um *benchmark* *BM*

$$\rho(X) := \text{ER}_{\text{BM}}(X) := -E[(\text{BM} - X)^+].$$

**Observação 3.5** 1. No caso de  $X$  tendo uma distribuição contínua temos equivalência entre os conceitos *ES*, *TCE* e *WCE* (vide *ACERBI; TASCHE, 2002, corolário 5.3*).

2. Apesar de *ES*, *TCE* e *WCE* tendo algumas características desejáveis (vide seção 3.2.3), elas sofrem do problema de não incorporar o “desvio” do benchmark, contrário às medidas *ER* e *DR*. Isso significa que não é penalizado o “quanto” os valores não-desejáveis desviam do benchmark. Esta característica parece ser a razão porque *ES* não seja consistente com os conceitos de dominância estocástica de primeira e segunda ordem (vide *DANIELSSON et al., 2005*). Na seção 6.3 introduzimos a abordagem de precificação por consistência que formulamos com duas configurações: a primeira não penaliza o desvios dos valores não-desejados, seguindo assim o conceito da medida *ES*. A segunda formulação **penaliza** os desvios, sendo assim em analogia com as medidas *ER* e *DR*. Neste caso evitamos a incoerência com as dominâncias estocásticas de primeira e segunda ordens.

### 3.2.2 Medidas de risco por equivalente certo e conjuntos de aceitação

Interpretamos uma medida funcional como um mapeamento do espaço das variáveis aleatórias no espaço real. O risco associado com uma variável aleatória foi então visto como resultado de uma certa transformação a ela aplicada. Uma alternativa para entender

o risco embutido numa variável aleatória é se perguntar quantas entidades determinísticas são necessárias para tornar a variável aleatória aceitável. Essa idéia é originada do equivalente certo, i.e., o valor determinístico mínimo pelo qual o agente venderia a variável aleatória.

**Definição 3.9 (Equivalente monetário certo)** *O equivalente certo é o valor real  $\bar{x} \in \mathbb{R}$  tal que:*

$$U(\bar{x}) = E[U(X)], \quad X \in \mathcal{X}.$$

Suponhamos que um agente define um valor que ele pode perder na aplicação do seu dinheiro. O risco atribuído a uma variável aleatória  $X \in \mathcal{X}$  pode ser visto como o valor de ativos sem risco (e.g. caixa) que é necessários para tornar a carteira de ativos com risco  $X$  aceitável.

**Definição 3.10 (Medida de risco induzida por um conjunto)** *Seja  $\mathcal{A}$  um conjunto de distribuições. A medida de risco para uma distribuição  $F$ , induzida pelo conjunto  $\mathcal{A}$ , é dada por<sup>12</sup>*

$$R_{\mathcal{A}}^{(ec)}(F) := \inf \{x \mid T_x F \in \mathcal{A}\}.$$

Se um nível de risco determinístico  $c \in \mathbb{R}$  é dado, podemos nos perguntar quais seriam as distribuições de retornos (por exemplo, criados por uma composição de um carteira) que implicam, após aplicação de  $R^{(ec)}$ , valores  $c(X) \leq c$ . Para isso, usa-se a definição de um conjunto aceitável.

**Definição 3.11 (Conjunto aceitável)** *O conjunto de nível*

$$\mathcal{A}_{R^{(ec)}}(c) := \{F \in \mathcal{D} \mid R^{(ec)}(F) \leq c\}$$

*é chamado de conjunto aceitável ao nível  $c \in \mathbb{R}$ , induzido pela medida de risco  $R^{(ec)}$ .*

Observamos que um conjunto aceitável é um conjunto de distribuições e não de valores determinísticos. O número de valores determinísticos para que o investidor possa aceitar uma distribuição  $F$  encontra-se na definição da medida de risco. A conexão entre medidas de risco por equivalente certo e conjuntos aceitáveis é dado no seguinte resultado<sup>13</sup>.

<sup>12</sup>Usamos a definição de uma operação de translação  $(T_x F)(u) := F(x - u)$ ,  $\forall u \in \text{dom} F$  e a notação  $R^{(ec)}$  para destacar a conexão com o espírito do equivalente certo.

<sup>13</sup>Vide PFLUG (1999) para uma síntese de resultados.

**Proposição 3.6** *Suponha que  $F$  é invariante sob translação, i.e.,  $\forall F \in \mathcal{A}$  e  $x \geq 0$  vale que  $T_x F \in \mathcal{A}$ .*

1. *Vale que  $\forall F$*

$$R_{\mathcal{A}_{R^{(ec)}}(0)}^{(ec)}(F) = R^{(ec)}(F).$$

2. *Se  $\mathcal{A}$  é fechado<sup>14</sup>, vale que*

$$\mathcal{A}_{R_{\mathcal{A}}^{(ec)}}(0) = \mathcal{A}.$$

**Prova.**

1. Seja  $F \in \mathcal{A}_{R^{(ec)}}(0)$ , então

$$\begin{aligned} R_{\mathcal{A}_{R^{(ec)}}(0)}^{(ec)} &= \inf \{x : T_x F \in \mathcal{A}_{R^{(ec)}}(0)\} \\ &= \inf \{x : R^{(ec)}(T_x F) \leq 0\} \\ &= \inf \{x : R^{(ec)}(F) - x \leq 0\} \\ &= R^{(ec)}(F) \end{aligned}$$

2. “ $\supseteq$ ”: seja  $F \in \mathcal{A}$ . Então  $T_0 F = F \in \mathcal{A}$  e  $\inf \{x : T_x F \in \mathcal{A}\} \leq 0$ , o que implica que  $F \in \mathcal{A}_{R^{(ec)}}(0)$ .

“ $\subseteq$ ”: seja  $F \in \mathcal{A}_{R_{\mathcal{A}}^{(ec)}}(0)$  com  $\mathcal{A}$  fechado. Defina-se

$$x^* := \inf \{x : T_x F \in \mathcal{A}\}$$

e sabemos, por  $\mathcal{A}$  ser fechado, que  $T_{x^*} F \in \mathcal{A}$ . Isso implica que

$$F = T_{-x^*} T_{x^*} F \in \mathcal{A}.$$

■

### 3.2.3 Classes de medidas de risco

Pressupomos nessa seção que os eventos não desejáveis são aqueles  $\omega \in \Omega : X(\omega) \leq BM(X)$  e pressupomos que  $X \in L^1$  e  $BM(X) = E[X] = 0$ .

É interessante levar em conta na escolha de uma das medidas de risco não somente o aspecto matemático, mas também o significado econômico da abordagem. Um das

<sup>14</sup>Com respeito à topologia induzida pela convergência fraca.

primeiras questões é quais características uma medida de risco deve possuir. Por exemplo, SONDERMANN (1991) qualificou a aditividade como um requisito básico para qualquer sistema de preços razoável: “*additivity [is] a basic requirement for any economically reasonable price system in a market*”.

Os autores ARTZNER et al. (1997), ARTZNER et al. (1999) e DELBAEN (2000) foram os primeiros a desenvolver um esquema de classificação de medidas de risco, propondo algumas características desejáveis para medidas de risco a serem usadas no ambiente financeiro-econômico. Esses trabalhos deram início a outras pesquisas generalizadoras, como CVITANIC; KARATZAS (1999), e análises sobre quais medidas de risco podem ser classificadas dentro do esquema proposto (vide ACERBI; TASCHE, 2002). Uma medida de risco  $\rho$  é chamada de **coerente** se possuir as seguintes características:

1. Invariância sob translação - a inclusão de um instrumento sem incerteza numa carteira de instrumentos com risco  $X$  diminui o valor em risco da carteira pelo valor do ativo sem risco:<sup>15</sup>

$$\rho(X + \alpha \cdot B_T) = \rho(X) - \alpha, \quad \alpha \in \mathbb{R}$$

2. Subaditividade - existe a possibilidade de diversificação do risco. Juntar dois ativos de risco numa carteira implica num risco da carteira menor (ou igual) que a soma dos riscos individuais, ou seja, para todos  $X, Y \in \mathcal{X}$  vale

$$\rho(X + Y) \leq \rho(X) + \rho(Y).$$

3. Homogeneidade positiva - o risco de um ativo alavancado por uma constante é o mesmo que o risco alavancado do ativo:  $\forall \lambda \geq 0 : \rho(\lambda X) = \lambda \rho(X)$ .
4. Monotonicidade - se  $X$  é menor que  $Y$ , então o risco de  $Y$  tem que ser menor que o risco de  $X$  :

$$X \leq Y \Rightarrow \rho(X) \geq \rho(Y).$$

5. Relevância - qualquer ativo com risco tem um valor positivo de risco:

$$\forall X \leq 0, X \neq 0 : \rho(X) > 0.$$

---

<sup>15</sup>Usamos  $B_T$  para identificar o ativo livre de risco.

Esse esquema classifica algumas medidas tradicionais como inaceitáveis:<sup>16</sup>

- Variância não é coerente, pois não é monótona.
- Semi-variância não é coerente, pois não é subaditiva.
- $VaR$  não é coerente, pois não é subaditiva.
- $TCE$  não é coerente porque não é subaditiva (DELBAEN, 2000).
- $ES(X) := E_{\mathbb{P}} [X \cdot 1_{\{X \leq BM\}}]$  não é medida coerente, pois não é invariante sob translação. Mas, essa condição pode ser substituída por uma característica parecida (SCHULMERICH; TRAUTMANN, 2001):

$$B_T^{-1} \cdot E_{\mathbb{P}} [X + \alpha \cdot B_T] \leq \rho(X) - \alpha.$$

A interpretação é que se o valor  $ES[X]$  for injetado na carteira atual e aplicado no ativo livre de risco, a posição total terá um valor positivo *na média*. A invariância sob translação garante um valor positivo em todos os casos. Por causa da caracterização mais fraca, porém, economicamente ainda interessante, SCHULMERICH; TRAUTMANN (2001) chamaram a medida de **quase-coerente**.

Para verificar se uma medida de risco (estática)  $\rho$  é coerente, pode ser usada a proposição 4.1 de ARTZNER et al. (1997):  $\rho$  é coerente se existe uma família de probabilidades  $\mathcal{P}$  em  $\Omega$  tal que

$$\rho(X) = \sup_{\mathbb{Q} \in \mathcal{P}} \{E_{\mathbb{Q}}[-X]\}. \quad (3.1)$$

Portanto, quando trabalhamos com medidas de risco, podemos verificar se existe um conjunto  $\mathcal{P}$  tal que<sup>17</sup>

$$E_{\mathbb{P}} [R(X)] = \sup_{\mathbb{Q} \in \mathcal{P}} \{E_{\mathbb{Q}}[-X]\}.$$

Usando o lema 3.3, sabemos como conectar uma função de risco  $R$  com uma função de distorção  $g$  ou um espectro de risco  $\phi$ . Com a condição de  $\phi$  ser admissível, a medida de risco  $M_{\phi}$  é coerente, vide ACERBI (2002), teorema 2.5. Esse fato foi provado para funções de distorção por DENNEBERG (1994), no qual ele mostra que, no caso  $\mathbb{P}(X > x) =$

<sup>16</sup>Veja ARTZNER et al. (1999).

<sup>17</sup>As medidas identificadas com (3.1) podem ser interpretadas como criadas a partir de cenários, pois toda  $\mathbb{Q} \in \mathcal{P}$  pode ser vista como cenário. Um exemplo seria a medida de risco criada pela  $BM\&F$  para o cálculo das margens nos contratos de futuros de renda fixa.

$1 - F_X(x)$ ,  $g$  decrescente em  $x$  e côncava em  $\mathbb{P}$ , a medida distorcida  $M_g$  é uma medida de risco coerente. FRITELLI (2002) conseguiu mostrar que uma condição suficiente para uma medida de risco  $\rho$  ser representável como em (3.1) é a sua sublinearidade.

Como uma extensão das medidas coerentes, a noção de medidas de risco **convexas** foi introduzida por FÖLLMER; SCHIED (2002). A diferença é a substituição das condições de subaditividade e homogeneidade positiva pela condição (mais fraca) de convexidade. Ou seja, para qualquer variável aleatória  $X \in \mathcal{X}$ , uma funcional  $\rho : \mathcal{X} \rightarrow \mathbb{R}$  é chamada de medida de risco convexa se valem as condições:

1. Convexidade -  $\forall X, Y \in \mathcal{X}, \forall \lambda \in [0, 1]$  :

$$\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y).$$

2. Monotonicidade - se  $X \leq Y$ , então  $\rho(X) \geq \rho(Y)$ .

3. Invariância sob translação - se  $m \in \mathbb{R}$ , então  $\rho(X + m) = \rho(X) - m$ .

Definindo uma função de penalidade  $\alpha$  nas medidas de probabilidade  $\mathcal{P}$ , os autores mostraram que existe uma representação equivalente ao caso de medidas coerentes da forma (3.1):

$$\rho(X) = \sup_{\mathbb{Q} \in \mathcal{P}} \{E_{\mathbb{Q}}[-X] - \alpha(\mathbb{Q})\}.$$

Foi provado por FRITELLI; GIANIN (2002), teorema 6, que a condição de sublinearidade é necessária e suficiente para uma medida de risco convexa ser representável dessa forma.

## 3.3 Função de utilidade

### 3.3.1 Generalidades

A precificação por utilidade considera mapeamentos  $U : \mathbb{R}^+ \rightarrow \mathbb{R}$ , sendo que valores negativos “não são úteis”. A função pode ser aplicada em fluxos financeiros intertemporais (como consumo, por exemplo) e/ou ao valor final da carteira criada.

Contrário a muitas propostas-padrão (vide seção 4.3 para referências), no trabalho atual, temos que considerar o resultado final  $V_T = X_T^{x, \pi, C} - D_T \in \mathbb{R}$  e definir uma função de utilidade no espaço imagem da variável aleatória. Começamos com a introdução de

algumas noções, pois o problema da utilidade ser definida em todos os números reais implica uma particularidade quando consideramos a análise do problema dual.

A utilidade de um investidor é normalmente diferenciada entre uma utilidade dos resultados intertemporais, especialmente dos consumos possíveis,  $\{C_t\}_{0 \leq t < T} \subseteq \mathcal{C}$  e a utilidade criada a partir o valor final da carteira,  $X_T \in \mathbb{R}$ . Para modelar o primeiro caso, definimos<sup>18</sup>

$$\mathcal{U}_1 := \left\{ \begin{array}{l} U : \mathcal{C} \rightarrow \mathbb{R} \mid U \in C^2(\mathcal{C}, \mathbb{R}), \\ \text{crescente, (estritamente) côncava e satisfazendo crescimento polinomial} \end{array} \right\},$$

$$\mathcal{U}_2 := \left\{ \begin{array}{l} U : \mathbb{R} \rightarrow \mathbb{R} \mid U \in C^2(\mathbb{R}^+, \mathbb{R}), \\ \text{crescente, (estritamente) côncava e satisfazendo crescimento polinomial} \end{array} \right\}.$$

Obviamente, um investidor pode ter uma utilidade a partir dos dois conjuntos o que significa que precisamos definir

$$\mathcal{U} := \mathcal{U}_1 \cup \mathcal{U}_2 \subseteq C^2(\mathcal{C} \times \mathbb{R}, \mathbb{R}).$$

Como queremos escolher uma função de utilidade específica, é importante considerar as estratégias de investimento que são admissíveis, de forma que as esperanças das utilidades criadas sejam limitadas.

**Definição 3.12 (Estratégia admissível considerando função de utilidade)** *Chamam uma estratégia  $\varphi = (x, \pi, C)$  de admissível considerando a classe de funções de utilidades  $\mathcal{U}$  e escrevemos - abusando da notação por conveniência -  $\varphi \in \mathcal{A}(x)$ , se as seguintes condições são satisfeitas:*

1.  $X_t^{x, \pi, C} \geq k$  q.c para todo  $t \in [0, T]$ ,  $k \in \mathbb{R}$ ,
2. para toda  $U_1 \in \mathcal{U}_1$ :  $E_{\mathbb{P}} \left[ \int_0^T U_1 \left( X_T^{x, \pi, C} \right) dt \right] < \infty$ ,
3. para toda  $U_2 \in \mathcal{U}_2$ :  $E_{\mathbb{P}} \left[ \left| U_2 \left( X_T^{x, \pi, C} \right) \right| \right] < \infty$ .

As possíveis combinações de funções de utilidade no mercado incompleto podem ser abrangidas pelos seguintes casos:

---

<sup>18</sup>Uma função  $f : \mathbb{R} \rightarrow \mathbb{R}$  satisfaz a condição de crescimento polinomial se existem constantes  $c_1$  e  $c_2$  tal que

$$|f(x)| \leq c_1 \cdot (1 + |x|)^{c_2}.$$



**Caso 1:** A estratégia  $\varphi$  não precisa ser auto-financeável, i.e.,  $C \neq 0$ . No caso e, que  $C_t \geq 0$  quase certamente (q.c.) para todos  $t \in [0, T]$ , o investidor retira recursos do seu valor acumulado - ele consome. Assim, ele cria uma certa utilidade  $U$  ao longo do período de aplicação:

$$U \in \mathcal{U}_1.$$

Como no mercado incompleto ele não pode replicar o pagamento final com uma estratégia auto-financeável, ele também não pode fazer isso retirando continuamente recursos:  $X_T^{x,\pi,C} \leq X_T^{x,\pi,0} \leq H_T$  q.c. por  $C_t \geq 0$  q.c. Assim ocorrerá uma perda futura:

$$V_T^{\pi,x,C,-D} := X_T^{x,\pi,C} - D_T \leq 0 \quad \text{q.c.}, \quad (3.2)$$

com

$$\mathbb{P} \left( X_T^{x,\pi,C} - D_T < 0 \right) > 0, \quad \text{por } t \in [0, T].$$

Para essa perda, o investidor não estabelece uma utilidade, pois o valor é negativo (ele sofre pela perda), o que nos permite dizer que a utilidade de um valor negativo é um sofrimento. Definindo  $\Delta_T^- := D_T - X_T^{x,\pi,C}$  se  $X_T^{x,\pi,C} \leq D_T$  (e  $\Delta_T^+ := X_T^{x,\pi,C} - D_T$  se  $X_T^{x,\pi,C} \geq D_T$ ), poderíamos também associar uma utilidade à perda  $\Delta_T^- \in \mathbb{R}^+$  usando a utilidade

$$Z \in \mathcal{U}_2.$$

**Caso 2:** Por causa de custos de transação, é bem comum que o investidor tenha que injetar recursos adicionais ao longo do período de negociação, i.e.,  $C_t \leq 0$  q.c. para  $t \in [0, T]$ . Esses pagamentos com fluxo negativo criam também sofrimento  $S$  ao longo do período:

$$S \in \mathcal{U}_1,$$

$$C \mapsto S(-C).$$

**Caso 3:** O investidor pode combinar os casos acima de várias formas:

1. Aceita um sofrimento pequeno, vindo da injeção de recursos ao longo do período, com a finalidade de somente ter um sofrimento pequeno de um pagamento final:  $S > 0, Z > 0$ .

2. Aceita um sofrimento maior de um pagamento maior final em compensação por uma utilidade intermediária devido a um aumento do consumo ao longo do período:  $U > 0, Z > 0$ .
3. Aceita um sofrimento ainda maior, decorrente das injeções intermediárias, para poder aproveitar de uma utilidade aumentada no final do período por causa de um superávit criado:  $S > 0, V > 0$ .

**Observação 3.7** 1. *Combinação 2 sem consumo  $C \equiv 0$ , e portanto  $U = 0$  e  $Z > 0$ , é a idéia subjacente da abordagem de esperança-variância.*

2. *Combinação 3 sem perda final, i.e.,  $X_T = D_T$  q.c., e portanto  $S > 0, V = 0$  é a idéia subjacente da abordagem de minimização de risco.*

3. *Uma função de utilidade atribui valores econômicos às realizações que podem ser utilizadas para punir os valores mais distantes do benchmark; por exemplo, a função de utilidade quadrática  $U(x) = ax - \frac{b}{2}x^2, a, b \in \mathbb{R}$  penaliza os desvios do valor esperado com a variância, i.e.  $E[U(X)] = aE[X] - \frac{b}{2}Var(X)$ . A reformulação equivalente em forma de penalidade - comum na literatura, vide por exemplo BENTAL (1985)- será usada na seção 5.*

**Condição 3.8** *Por razões de simplicidade, pressupomos nas nossas análises que  $C \equiv 0$ .*

Sem maiores discussões, usamos a teoria de utilidade esperada para caracterizar o comportamento do investidor. Existem violações famosas dessa teoria como o paradoxo de Allais (ALLAIS, 1953), e foram propostos diversos ajustes, vide CAMERER (1995), para uma compilação. Também surgiram teorias alternativas sobre o comportamento do investidor, como a teoria de prospecto de KAHNEMANN; TVERSKY (1979). Além dessas propostas, pode se pensar em incorporar a incerteza inerente nas preferências do investidor, o que resulta na teoria de utilidade esperada estocástica (vide BLAVATSKYY, 2005). Entendemos que é importante usar um sistema axiomático compatível com o comportamento real do investidor e livre de inconsistências. Porém, por razões de tratabilidade e de alta divulgação, continuamos usar a teoria de utilidade esperada de NEUMANN; MORGENTERN (1953). De fato, o foco desse trabalho não é analisar sistemas de axiomas, mas apresentar novas formas de precificação em mercados incompletos.

Com o uso de uma função de utilidade  $U$ , o agente tenta achar a estratégia de investimento  $\tilde{\pi} \in \mathcal{A}(x)$  tal que

$$\tilde{\pi} = \arg \max_{\pi \in \mathcal{A}(x)} E_{\mathbb{P}} \left[ U \left( V_T^{\pi, x, C, D} \right) \right]. \quad (\text{PInicial})$$

### 3.3.2 Estratégia eficiente, função de utilidade eficiente e controle ótimo

A abordagem tradicional para o problema de otimização da utilidade do agente é a do controle ótimo estocástico que foi aplicada, por exemplo, por MERTON (1990) em mercados financeiros de difusões. A idéia principal é derivar uma equação de Hamilton-Jacobi-Bellman, em geral uma equação diferencial parcial, e resolvê-la.

No caso de um sistema de preços que segue um modelo markoviano, podemos usar a abordagem de programação dinâmica estocástica para achar a estratégia ótima. Se vale a hipótese de utilidade esperada

$$U_2 \left( V_T^{\pi, x, C, D} \right) = U \left( V_T^{\pi, x, C, D} \right) = E_{\mathbb{P}} \left[ U \left( V_T^{\pi, x, C, D} \right) \right],$$

então escrevemos

$$U_t \left( V_T^{\pi, x, C, D} \right) := E_{\mathbb{P}} \left[ U \left( V_T^{\pi, x, C, D} \right) \middle| \mathcal{F}_t \right], \quad t \in [0, T]$$

e chamamos  $u_t$  o desempenho esperado do sistema ou utilidade esperada no instante  $t$ .<sup>19</sup>

A função

$$u_t(x) := \sup_{\pi \in \mathcal{A}(x)} U_t \left( V_T^{\pi, x, C, D} \right), \quad t \in [0, T]$$

é chamada função de utilidade indireta no tempo  $t$ .

Chamam-se estratégias eficientes aquelas estratégias admissíveis  $\varphi \in \mathcal{A}(x)$  que otimizam um problema associado de otimização de utilidade.

**Definição 3.13 (Estratégia eficiente)** *Uma estratégia admissível  $\pi^{\text{eff}} \in \mathcal{A}(x)$  é chamada de eficiente, se existe uma função de utilidade  $U \in \mathcal{U}$ , tal que  $\pi^{\text{eff}}$  resolve*

$$\pi^{\text{eff}} = \arg \sup_{\pi \in \mathcal{A}(x)} U \left( V_T^{\pi, x, C, D} \right). \quad (3.3)$$

<sup>19</sup>Se a hipótese de utilidade esperada não vale, o desempenho do sistema é simplesmente a utilidade.

Isso significa que  $\pi^{\text{eff}}$  é a estratégia ótima pela função de utilidade  $U$ , a qual chamamos de função de utilidade eficiente. O conjunto das estratégias eficientes é dado por:

$$\mathcal{E}(x) := \left\{ \pi^{\text{eff}} \in \mathcal{A}(x) \mid \exists U \in \mathcal{U} : \pi^{\text{eff}} = \arg \sup_{\pi \in \mathcal{A}(x)} U \left( V_T^{\pi, x, C, D} \right) \right\}. \quad (3.4)$$

A abordagem de controle ótimo é comum em finanças e foi popularizada pelos trabalhos de Merton (vide MERTON, 1990). A escolha de uma estratégia ótima  $\tilde{\pi}$  cria um preço  $x$  usando uma certa função de utilidade  $U$ . O desempenho esperado do sistema é dado por

$$\begin{aligned} u_s(x) &:= \sup_{\pi \in \mathcal{A}(x)} U_s \left( V_T^{\pi, x, C, D} \right) \\ &= \sup_{\pi \in \mathcal{A}(x)} E_{\mathbb{P}} \left[ \int_s^T U \left( V_T^{\pi, x, C, D} \right) \middle| \mathcal{F}_s \right], \quad s \in [0, T]. \end{aligned}$$

O uso de uma estratégia ótima  $\tilde{\pi}$  leva a

$$u_t(x) = U_t \left( V_T^{\tilde{\pi}, x, C, D} \right), \quad \forall t \in [0, T].$$

Se  $u_t$  é contínuo no domínio, chegamos às equações de controle estocástico ótimo de Hamilton-Jacobi-Bellman. Para todo  $s \in [0, T]$  é válido que<sup>20</sup>

$$0 = \frac{\partial u_t(x)}{\partial t} \Big|_{t=s} + \max_{\varphi \in \mathcal{A}(x)} [\mathcal{L}(t) u_t(x)] \Big|_{t=s} \quad (3.5)$$

$$u \left( T, V_T^{\pi, x, C, D} \right) = U \left( V_T^{\pi, x, C, D} \right) \quad (\text{condição de contorno}). \quad (3.6)$$

O teorema de verificação VI.4.1 de FLEMING; RISHEL (1975) fornece condições suficientes para que  $u$  e  $\varphi$  asseguram a atingibilidade de um mínimo em (3.5). O problema da existência de  $I$  é resolvida com a imposição de alguma condição que garante a parabolicidade uniforme do sistema (3.5). O teorema VI.6.1 de FLEMING; RISHEL (1975) resolve essa questão e o teorema de seleção VI.6.3 implica diretamente a existência de uma estratégia ótima  $\tilde{\pi}$ .

Apesar do fato de que resolver, mesmo numericamente, a equação em dimensões maiores seja extremamente complicado, a abordagem necessita que os preços dos ativos sigam um processo markoviano. Se não for o caso, não poderemos otimizar em cada passo a alocação de forma independente da trajetória percorrida.

Uma abordagem alternativa que evita a suposição de processos markovianos baseia-se na caracterização dual das carteiras por um conjunto de martingales que fornece uma formulação com respeito às medidas de probabilidades neutralizadoras de risco. A conexão

<sup>20</sup>Como o nosso objetivo aqui é simplesmente mostrar as conexões das abordagens, usamos de forma abstrata e geral  $\mathcal{L}$  para identificar a função geradora do processo.

entre o problema primal e dual depende tanto dos processos de preços (que podem ser até semi-martingales de forma geral) como da função de utilidade que precisa satisfazer algumas condições de regularidade. A discussão dessa conexão será o tema do próximo capítulo. Usaremos a dualidade para formular os problemas equivalentes da nossa abordagem proposta na seção 6.4.

# Capítulo 4

## Métodos de alocação ótima

### 4.1 Introdução

Queremos mostrar como o problema de escolha de uma função de utilidade para determinar um preço dentro da faixa dos preços livres de arbitragem é equivalente ao problema de escolha de uma medida equivalente martingale (EMM).<sup>1</sup>

A grande vantagem da abordagem por martingales é que esta evita alguns problemas da abordagem por programação estocástica com equações de Hamilton-Jacobi-Bellman (HJB), que são, entre outros:

- equações diferenciais, especialmente em dimensões maiores, poucas vezes têm soluções, nem mesmo numéricas
- sem a hipótese de um processo markoviano, a abordagem recursiva não é aplicável.

A abordagem por martingales, por outro lado, não está limitada a processos markovianos. Pode ser usado um procedimento em dois passos, como na abordagem HJB em modelos markovianos, mas para qualquer tipo de processo:<sup>2</sup>

---

<sup>1</sup>A abordagem de martingales foi usada por PLISKA (1986), COX; HUANG (1989,9) e KARATZAS; LEHOCZKY; SHREVE (1987) no caso de mercado completo. No caso incompleto em tempo discreto com sigma-álgebra  $\Omega$  discreta por HE; PEARSON (1991). No caso de uma difusão por KARATZAS et al. (1991), e, no caso geral de semi-martingales, por KRAMKOV; SCHACHERMAYER (1999).

<sup>2</sup>As noções serão introduzidas com cuidados matemáticos e esclarecidas nas futuras seções. A abordagem por martingales foi usada, por exemplo, em COX; HUANG (1989); KARATZAS et al. (1991); PLISKA (1986).

1. Problema de otimização estática: determina-se uma variável de valor da carteira  $\tilde{X} \in \mathcal{X}$  (vide seção 4.2.4) tal que, para uma medida de risco  $R$  (seção 3.2) e um pagamento aleatório final  $D_T$  (seção 2), vale

$$\tilde{X} = \arg \min_{X \in \mathcal{X}} E_{\mathbb{P}} [R(X_T - D_T)].$$

2. Problema de representação: determina-se uma estratégia de (super-) replicação para  $\tilde{X}$ .

Apesar dessas vantagens no caso geral, não é nosso objetivo investigá-las. Usamos essas explicações somente para deixar claro que a abordagem dual tem uma justificativa relevante (além da sua relevância matemática). No presente trabalho, estamos interessados somente na forma do problema dual e na sua conexão com o problema primal, mas não no método de solução.

Na teoria de dualidade, são tratados dois problemas de otimização que são considerados duais: o problema dual de minimização deve ser de tal forma que, quando o problema primal de maximização tem uma solução, o dual também tem, e os valores ótimos das funções-objetivo são iguais. Classicamente, em espaços finitos, o resultado é estabelecido pelo teorema de Kuhn-Fourier (KUHN, 1956). No caso de sistemas lineares homogêneos, ele pode ser substituído pelos teoremas de transposição, vide GORDAN (1873) e STIEMKE (1915), apud STOER; WITZGALL (1970). Quando a função-objetivo e as restrições são lineares, seguem os teoremas da programação linear de Gale-Kuhn-Tucker (GALE; KUHN; TUCKER, 1951). No nosso caso geral, porém, precisamos de resultados mais abrangentes, pois minimizamos funções convexas definidas em conjuntos convexos. Para esse tipo de problema, é necessário usar os conceitos das funções conjugadas de FENCHEL (1949). Generalizando mais ainda, qualquer teorema de existência de um ponto de inflexão pode ser interpretado como um teorema de dualidade, o que permite relaxar mais um pouco as condições de convexidade. Esses teoremas baseiam-se no uso de multiplicadores de Lagrange e são, às vezes, chamados de problemas de mín-máx (NEUMANN, 1928).<sup>3</sup>

Após uma breve introdução geral, mostraremos na seção 4.2.1 os passos no exemplo paradigmático para esclarecer qual seria o preço escolhido quando se usa uma escolha específica, sendo que, no nosso caso, uma função de utilidade exponencial, ou, equivalente,

<sup>3</sup>A justificativa teórica dessa abordagem, já descrita para o caso geral de conjuntos infinitos, é resumida no apêndice A.

a medida de mínima entropia relativa com respeito à medida original  $\mathbb{P}$ . Depois, vamos introduzir formalismo matemático, adaptando os trabalhos de KRAMKOV; SCHACHERMAYER (1999), KARATZAS (1996) e ROGERS (2001), nos casos mais gerais: o primeiro passo é o caso completo com discretos e finitos eventos (seção 4.2.2), depois seguem os casos de discretos e finitos eventos num mercado incompleto (seção 4.2.3) e infinitos eventos num mercado incompleto (seção 4.2.4), seguindo a estrutura de SCHACHERMAYER (2004a,0).

## 4.2 A conexão entre o problema primal e dual

Sendo o objetivo da decisão maximizar uma função  $U$ , estritamente côncava, na variável  $x$  sob a restrição  $Ax = b$ , temos o programa

$$\text{sujeito a } \begin{cases} \sup_x U(x) \\ Ax = b \\ x \geq 0. \end{cases} \quad (4.1)$$

Escrevendo a função lagrangiana com o multiplicador de Lagrange  $y$ , temos uma formulação sem restrições:<sup>4</sup>

$$\mathcal{L}(\tilde{x}, \tilde{y}) := \inf_y \sup_x [U(x) - y \cdot (Ax - b)].$$

**Observação 4.1 (Interpretação)** *O lagrangiano tem a interpretação tradicional de dar um preço<sup>5</sup> para cada unidade monetária em cada estado  $\omega \in \Omega$ . Se, por exemplo, o agente quer um ganho  $G_{t_1}(\omega) \neq \pi_{t_0} \cdot (S_{t_1}(\omega) - S_{t_0})$ ,  $i = 1, 2, 3$ , seguindo uma estratégia  $\hat{\pi}$  que não resulte em  $G_{t_1}(\omega)$ , então, ele compra (vende) a quantidade de unidades de dinheiro adicional necessária para atingir a sua meta por estado. Como o valor desse dinheiro adicional depende do estado  $\omega$ , o preço é dado pelo preço de Arrow-Debreu  $q(\omega) = y(\omega)$ . A unidade do preço de Arrow-Debreu é o incremento marginal em utilidade por incremento marginal de dinheiro em estado  $\omega_i$ . Assim, quanto mais o fluxo de caixa aumenta em  $\omega$ , maior será a utilidade de uma unidade de dinheiro.<sup>6</sup>*

<sup>4</sup>Continuamos usar o símbolo  $\tilde{x}$  para identificar a variável  $x$  como ótima, em vez de usar o asterisco,  $x^*$ , pois precisamos dele para identificar o espaço dual e os seus elementos.

<sup>5</sup>Nesse caso, a unidade do preço é “utilidade”.

<sup>6</sup> $q$  é também chamado de preço-sombra (*shadow price*).



Como a escolha de  $y$  é livre em  $\sup_x [U(x) - y \cdot (Ax - b)]$ , temos mais liberdades de “violiar” a restrição do problema (4.1) e, portanto, vale com  $\tilde{x}$  uma solução ótima de (4.1) que

$$\sup_x [U(x) - y \cdot (Ax - b)] \geq U(\tilde{x}) + y \cdot (A\tilde{x} - b) = U(\tilde{x}),$$

já que  $A\tilde{x} = b$  para uma solução de (4.1). O objetivo é o de determinar o menor limite superior por meio de uma minimização em  $y$ :

$$\inf_y \sup_x [U(x) - y \cdot (Ax - b)] \geq U(\tilde{x}).$$

E queremos saber quais as condições para que valha a igualdade, ou seja, que

$$\inf_y \sup_x [U(x) - y \cdot (Ax - b)] = \sup_x \inf_y [U(x) - y \cdot (Ax - b)]. \quad (4.2)$$

A partir da formulação

$$\begin{aligned} & \sup_x [U(x) - y \cdot (Ax - b)] \\ &= yb + \sup_x [U(x) - x^* \cdot x], \end{aligned}$$

com  $x^* := yA$ , introduzimos a noção da transformada de Fenchel-Legendre (ou conjugada convexa):

**Definição 4.1 (Função conjugada convexa)** *Sejam  $X, X^* \subseteq \mathbb{R}^n$ . Para uma função  $f : X \rightarrow \mathbb{R}$ , a função conjugada convexa  $f^* : X^* \rightarrow \mathbb{R}$  é dada por*

$$f^*(x^*) := \sup_{x \in X} [f(x) - x \cdot x^*].$$

No caso em que  $f$  for uma função convexa e fechada, vale que

$$f^{**}(x) = \sup_{x^* \in X^*} [f^*(x^*) - x^* \cdot x] = f(x),$$

o que provê a intuição pelo mecanismo fundamental de conexão entre o problema primal e o dual.

O problema (4.1) pode agora ser escrito como

$$\inf_y [yb + f^*(y)]$$

e observamos que  $f^*(y)$  pode ser infinito. Como estamos somente interessados no caso contrário, precisamos impor a restrição de limitação<sup>7</sup>

$$U(x) - yAx < \infty.$$

Veremos na próxima seção que essa restrição resulta na condição de os processos de preços ajustados pela incerteza serem martingales e  $y$  será interpretada como a derivada de Radon-Nykodim.

**Observação 4.2** *Isso significa que podemos identificar, a partir de um critério matemático de escolha de uma medida  $q \in \mathcal{M}(S)$  qual é implicitamente a escolha econômica de uma função de utilidade equivalente. Dizemos que procuramos, a partir do funcional de escolha da medida EMM, a função de utilidade implícita.*

O objetivo das seções 4.2.2 até 4.2.4 é o de esclarecer e formalizar com modelos de complexidade crescente a conexão entre o problema primal e dual e de estabelecer as condições para que os valores ótimos sejam iguais, i.e., que vale (4.2).

### 4.2.1 Visualização da dualidade no exemplo paradigmático

Usaremos a dualidade descrita de forma heurística para o nosso modelo básico com três estados  $\{\omega_i\}_{i=1,2,3}$ . Queremos aplicar o critério de escolher uma medida  $\mathbb{Q} \in \mathcal{M}(S)$ ,

$$\mathbb{Q} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} q(\omega_1) \\ q(\omega_2) \\ q(\omega_3) \end{pmatrix}$$

tal que ela minimize a divergência direcionada de Kullback-Leibler em relação a uma medida prévia  $\mathbb{P} = (p(\omega_1), p(\omega_2), p(\omega_3))^T$ :<sup>8</sup>

$$KL(\mathbb{Q} \parallel \mathbb{P}) := \sum_{i=1}^3 q(\omega_i) \cdot \log \left( \frac{q(\omega_i)}{p(\omega_i)} \right).$$

<sup>7</sup>Num problema com função-objetivo linear, deduz-se a restrição do problema dual como

$$U' \cdot x - yAx < \infty \Leftrightarrow U' - yA \leq 0 \Leftrightarrow U' \leq yA.$$

<sup>8</sup>Se não mencionamos explicitamente a base dos logaritmos, pressuponha-se a base natural  $e$ .

Isso significa que partimos do problema dual para chegar ao problema primal de otimização de carteiras. A medida prévia é a medida de probabilidade objetiva  $\mathbb{P} = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})^T \notin \mathcal{M}(S)$  e o problema do agente é o de determinar

$$\tilde{\mathbb{Q}} = \arg \min_{\mathbb{Q}} \{ KL(\mathbb{Q} \parallel \mathbb{P}) \mid S_{t_0} = E_{\mathbb{Q}}[S_{t_1}] \}. \quad (4.3)$$

Podemos agora determinar explicitamente  $\tilde{\mathbb{Q}} = (\tilde{q}_1, \tilde{q}_2, \tilde{q}_3)'$ :

$$\begin{aligned} \frac{\partial KL(\mathbb{Q} \parallel \mathbb{P})}{\partial \mathbb{Q}} &= \sum_{i=1}^3 \frac{\partial KL(q_i \parallel p_i)}{\partial q_i} \\ &= \sum_{i=1}^3 \left[ \log \left( \frac{\tilde{q}_i}{p_i} \right) + p_i \right] = 0. \end{aligned}$$

Com as condições de não-arbitragem determinamos o preço de Arrow-Debreu ótimo para o estado  $\omega_3$  como a solução do sistema

$$\begin{aligned} \log \left( \frac{\tilde{q}_1}{p_1} \right) + \log \left( \frac{\tilde{q}_2}{p_2} \right) + \log \left( \frac{\tilde{q}_3}{p_3} \right) &= -1 \\ 120\tilde{q}_1 + 100\tilde{q}_2 + 90\tilde{q}_3 &= 100 \\ \tilde{q}_1 + \tilde{q}_2 + \tilde{q}_3 &= 1. \end{aligned}$$

que é dada por

$$-\frac{1}{3}(1 + 4 \log(2)) = \log(\tilde{q}_3),$$

o que implica que

$$\begin{aligned} \tilde{q}_3 &\simeq 0,226, \\ \tilde{q}_1 &= \frac{1}{2}\tilde{q}_3 \simeq 0,113, \\ \tilde{q}_2 &= 1 - \frac{3}{2}\tilde{q}_3 \simeq 0,661. \end{aligned}$$

Para determinar a função de utilidade implícita, determinamos primeiro a função conjugada convexa de  $F(\mathbb{Q}) = KL(\mathbb{Q} \parallel \mathbb{P})$ . Como  $KL(\mathbb{Q} \parallel \mathbb{P})$  é separável nas probabilidades  $q(\omega_i)$ , segundo o teorema 16.4 de ROCKAFELLAR (1970), temos

$$F^*(x^*) := KL^*(y) = \sum_{i=1}^3 KL_{q,p}^*(y_i).$$

A conjugada convexa de  $KL(q(\omega_i) \parallel p(\omega_i))$ ,  $KL_{q,p}^*(y_i)$  é dada por

$$\begin{aligned} KL_{q,p}^*(y_i) &= \left( q(\omega_i) \cdot \log \left( \frac{q(\omega_i)}{p(\omega_i)} \right) \right)^* \\ &= \max_{q_i} \left\{ q_i y_i - q_i \log \left( \frac{q_i}{p_i} \right) \right\}. \end{aligned}$$

Escrevendo  $r_i := \frac{q_i}{p_i}$ , temos:

$$\begin{aligned} KL_{r_i}^*(y_i) &= \max_{r_i} \{y_i r_i p_i - r_i p_i \log(r_i)\} \\ &= p_i \max_{r_i} \{r_i y_i - r_i \log(r_i)\}. \end{aligned}$$

Para achar o máximo  $\tilde{r}_i$ , determinamos a derivada de  $g_{y_i}(r_i) := r_i y_i - r_i \log(r_i)$  em relação a  $r_i$  e igualamos o resultado a zero

$$\begin{aligned} \frac{\partial g_{y_i}(r_i)}{\partial r_i} &= y_i - \log(\tilde{r}_i) - 1 = 0 \\ \Leftrightarrow \tilde{r}_i &= \exp\{y_i - 1\}. \end{aligned}$$

Como  $\frac{\partial^2 g_{y_i}(\tilde{r}_i)}{\partial r_i^2} < 0$ , segue que  $\tilde{q}_i = p_i \cdot \exp\{y_i - 1\}$  é realmente o ponto de máximo.

Inserindo o resultado na função, obtemos

$$\begin{aligned} KL_{q,p}^*(y) &= \sum_{i=1}^3 KL_{r_i}^*(y_i) \\ &= \sum_{i=1}^3 p_i (\tilde{r}_i y_i - \tilde{r}_i \log(\tilde{r}_i)) \\ &= \sum_{i=1}^3 p_i \exp\{y_i - 1\}. \end{aligned}$$

Já que a função-objetivo do problema primal é dada pela maximização do mínimo do lagrangiano, e usando a definição da função de utilidade exponencial negativa  $U(x) := -\exp(-x)$ , temos

$$\begin{aligned} &\max_y (-KL)_{q,p}^*(-y) \\ &= \max_y \sum_{i=1}^3 -p_i \exp\{-y_i - 1\} \\ &= \frac{1}{\exp(1)} \max_{\pi \in \mathcal{A}(x)} E_{\mathbb{P}}[-\exp\{-y\}] \\ &= \max_{\pi \in \mathcal{A}(x)} E_{\mathbb{P}}\left[-\exp\left\{-V_{t_1}^{x,\pi,0,D}\right\}\right] \\ &= \max_{\pi \in \mathcal{A}(x)} E_{\mathbb{P}}\left[U\left(V_{t_1}^{x,\pi,0,-D}\right)\right], \end{aligned} \tag{4.4}$$

que é o resultado já derivado de forma geral por FRITELLI (2000a) e no caso de processos de Wiener por ROUGE; ELKAROUI (2000). DELBAEN; RHEINLÄNDER; SCHWEIZER (2000) chegaram ao mesmo resultado com uma formulação de funções de penalidade.

**Observação 4.3** A formulação (4.4) pode ser adaptada ainda para a forma de penalidade, vide ROUGE; ELKAROUI (2000), um método que será usado na seção 5.5 num contexto diferente.

Com o critério dual de escolha, podemos agora determinar uma estratégia ótima  $\tilde{\pi} = (\tilde{\pi}_{t_0}^{(0)}, \tilde{\pi}_{t_0}^{(1)})^T$  do ponto de vista de um agente que possui uma utilidade exponencial e o valor inicial da estratégia  $\tilde{\pi}_{t_0} S_{t_0}$ . Seja

$$f(\pi) = - \sum_{i=1}^3 p_i \cdot \exp \left\{ -x - \begin{pmatrix} \pi_{t_0}^{(0)} & \pi_{t_0}^{(1)} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ S_{t_1}(\omega_i) - 100 \end{pmatrix} + D_{t_0}(\omega_i) \right\},$$

e tentando resolver sob a restrição de  $V_t \geq K, \forall t$ , temos

$$\begin{aligned} \frac{\partial f(\pi)}{\partial \pi} &= \frac{\partial \left( -\frac{1}{4} \exp \{ -20\pi_{t_0}^{(1)} + 20 - x \} - \frac{1}{2} \exp \{ -x \} - \frac{1}{4} \exp \{ +10\pi_{t_0}^{(1)} - x \} \right)}{\partial \pi} \\ &= 0. \end{aligned}$$

Segue primeiramente que

$$\tilde{\pi}_{t_0}^{(1)} = \frac{1}{30} (20 + \log(2)) \approx 0,69. \quad (4.5)$$

Esse valor é a quantidade que o vendedor do derivativo compra do ativo com incerteza  $S^{(1)}$  para se proteger contra flutuações de preço. Para poder fazer essa aplicação ele tem que se financiar com uma venda de ativos sem incerteza,  $S^{(0)}$ . Com a restrição  $V_{t_0}^{x,\pi,0,D} = x \geq K$  e  $x = \begin{pmatrix} \pi_{t_0}^{(0)} & \pi_{t_0}^{(1)} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 100 \end{pmatrix} = \pi_{t_0}^{(0)} + 100\pi_{t_0}^{(1)}$ , escolhamos a aplicação em  $\pi_{t_0}^{(0)}$  de tal forma que a restrição seja atingida. Isso implica que

$$\begin{aligned} \tilde{\pi}_{t_0}^{(0)} &= K - \tilde{\pi}_{t_0}^{(1)} 100 \\ &= K - \frac{10}{3} (20 + \log(2)). \end{aligned}$$

No caso típico de  $K = 0$ , teríamos

$$\tilde{\pi}_{t_0, K=0}^{(0)} = -\frac{10}{3} (20 + \log(2)).$$

Esse valor é a quantidade do ativo subjacente  $S^{(0)}$  que o vendedor do derivativo tem que comprar (como, no caso atual, o valor é negativo, ele vende esta quantidade, ou seja ele toma emprestado) para financiar a sua aplicação em  $S^{(1)}$ .

### 4.2.2 Caso completo com $\Omega$ finito e discreto

Na análise geral do método da dualidade, começamos com o caso completo, pois não precisamos escolher entre diferentes medidas e pressupomos que  $|\Omega| = N < \infty$  e discreto.

Para um  $x \in \text{dom}(U)$ , podemos escrever o problema

$$\begin{aligned} u(x) &:= \sup_{\pi \in \mathcal{A}(x)} E_{\mathbb{P}} [U(V_T)] \\ &:= \sup_{\pi \in \mathcal{A}(x)} E_{\mathbb{P}} [U(X_T - D_T)] \\ &= \sup_{\pi \in \mathcal{A}(x)} E_{\mathbb{P}} [U(x + (\pi \cdot S)_T - D_T)] \end{aligned}$$

como

$$\begin{aligned} &\sup_{\pi \in \mathcal{A}(x)} E_{\mathbb{P}} [U(V_T)], \\ \text{tal que } &E_{\mathbb{Q}} [V_T] \leq x. \end{aligned} \tag{4.6}$$

Sendo que

$$\mathcal{G} := \{(\pi \cdot S)_T \mid \pi \in \mathcal{A}(x)\}$$

é o conjunto de processos admissíveis de ganho, definimos

$$\mathcal{X}(x) := \{X_T \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \mid \exists f \in \mathcal{G} : f \geq X_T\}$$

como o conjunto de valores finais super-replicáveis por estratégias  $\pi \in \mathcal{A}(x)$  e capital inicial  $x \in \text{dom}(U)$ .<sup>9</sup> No caso discreto finito com a notação de  $p_i := p(\omega_i)$ ,  $q_i := q(\omega_i)$  e  $V_i := V_T(\omega_i)$ , vale que

$$\begin{aligned} E_{\mathbb{P}} [U(V_T)] &= \sum_{i=1}^N p_i \cdot U(V_i), \\ E_{\mathbb{Q}} [V_T] &= \sum_{i=1}^N q_i V_i \leq x. \end{aligned}$$

Formando o lagrangiano, no qual  $y$  é o multiplicador de Lagrange, segue que

$$\begin{aligned} \mathcal{L}(V_T(\omega_1), \dots, V_T(\omega_N); y) &= \sum_{i=1}^N p_i U(V_i) - y \cdot \left( \sum_{i=1}^N q_i V_i - x \right) \\ &= \sum_{i=1}^N p_i \cdot \left( U(V_i) - y \cdot \frac{q_i}{p_i} V_i \right) + y \cdot x. \end{aligned} \tag{4.7}$$

---

<sup>9</sup>O conjunto  $\mathcal{G}$  pode também ser interpretado como o conjunto de valores terminais  $X_T$  acessíveis com o capital inicial  $x = 0$ .

No problema primal, queremos maximizar o lagrangiano sobre todos os  $\{V_T(\omega_i)\}_{i=1,\dots,N}$  com o valor minimal  $\tilde{y}$  dado. O problema dual faz o contrário, i.e., minimiza sobre todos os preços-sombra  $y$  com um conjunto máximo de valores finais da carteira  $\{\tilde{V}(\omega_i)\}$  dado. Nesse caso, queremos otimizar a partir de um  $y \geq 0$  dado:

$$KL(y) = \sup_{\{V_i\}} \mathcal{L}(V_1, \dots, V_n; y)$$

que depois tem que ser minimizado em  $y$ . Para determinar  $KL(y)$  é suficiente achar para cada  $i = 1, 2, \dots, N$

$$\sup_{V_i} \left( U(V_i) - y \cdot \frac{q_i}{p_i} \cdot V_i \right). \quad (4.8)$$

Além da vantagem de poder analisar  $N$  problemas separados, podemos também identificar o problema com a transformada de Legendre-Fenchel de  $-U(-V_i)$ :

$$U^*(\eta_i) = \sup_{V_i} (U(V_i) - \eta \cdot V_i),$$

sendo que  $0 < \eta = y \cdot \frac{q_i}{p_i}$ . Por causa das condições de regularidade de  $U$ , temos também que a função conjugada de  $U^*(\eta)$  é  $U(V_i)$  e que  $U^*$  pressupõe as seguintes regularidades:

- $U^*$  é contínuo, diferenciável, decrescente e estritamente convexa para todos os  $y \in \mathbb{R}^+$ .
- $\frac{dU^*}{dy} \Big|_{y=0} = -\infty, \frac{dU^*}{dy} \Big|_{y=\infty} = 0$ .
- $U^*(0) = U(\infty), U^*(\infty) = U(0)$ .

Inserindo a transformada de Legendre-Fenchel em (4.8), temos

$$\begin{aligned} KL(y) &= \sum_{i=1}^N p_i U^*(\eta_i) + yx \\ &= \sum_{i=1}^N p_i U^* \left( y \cdot \frac{q_i}{p_i} \right) + yx \\ &= E_{\mathbb{P}} \left[ U^* \left( y \cdot \frac{dQ}{d\mathbb{P}} \right) \right] + yx. \end{aligned}$$

$\frac{dQ}{d\mathbb{P}}$  é a derivada de Radon-Nykodim. Além disso, com as condições de regularidade de  $U^*$ , vale que

$$-(U^*(y))' = -\frac{dU^*}{dy} = I \left( \frac{dU}{dx} \right) (y) = x$$

e, portanto, podemos achar para cada  $x \in \text{dom}(U)$  um único  $\tilde{y}$  tal que

$$U^*(\tilde{y})' = -x.$$

Mas, esse  $\tilde{y}$  é também o valor mínimo de

$$KL(\tilde{y}) = \min_{y>0} E_{\mathbb{P}} \left[ U^* \left( y \cdot \frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right] + yx, x \in \text{dom}(U). \quad (4.9)$$

Tendo achado  $\tilde{y}$ , podemos inseri-lo no problema global de (4.7) para achar os valores ótimos da carteira e, implicitamente, a estratégia de investimento  $\pi \in \mathcal{A}(x)$ . Como a função  $U$  é estritamente côncava, a função  $\mathcal{L}$  pressupõe um único máximo, temos que

$$\tilde{U}_i^* \left( \tilde{y} \frac{q_i}{p_i} \right) = (U')^{-1} \left( \tilde{y} \frac{q_i}{p_i} \right) = -x.$$

Podemos finalmente deduzir a dualidade entre as funções de desempenho dos sistemas, pois

$$\begin{aligned} \inf_{y>0} \{u^*(y) + xy\} &= u^*(\tilde{y}) + x \cdot \tilde{y} \\ &= \mathcal{L}(\tilde{V}_1, \dots, \tilde{V}_n; \tilde{y}(x)) \\ &= E_{\mathbb{P}} [U(\tilde{V}_T)] \\ &= u(x). \end{aligned}$$

Se conseguirmos estabelecer essa relação para o caso geral (infinitos eventos e mercado incompleto), então temos um meio de relacionar a solução de um problema de otimização de carteiras com a solução de escolher uma medida neutralizadora de risco para estabelecer um funcional de preço linear.

### 4.2.3 Caso incompleto com $\Omega$ finito e discreto

No caso incompleto, o conjunto de medidas neutralizadoras de risco não é mais unitário e a restrição da otimização (4.6) tem que incorporar esse fato:

$$E_{\mathbb{Q} \in \mathcal{M}(S)} [V_T] \leq x.$$

O problema nesse caso é que o conjunto  $\mathcal{M}(S)$  tem infinitos elementos, o que criaria infinitas restrições. Para sair desse dilema, usamos o fato que qualquer polítopo<sup>10</sup>,  $E \subseteq$

<sup>10</sup>Um polítopo é um poliedro, ou seja, um subconjunto  $X \subseteq \mathbb{R}^N$  criado por finitas intersecções de semi-espacos fechados, que é compacto, convexo e não-vazio.



$\mathbb{R}^N$ , tem finitos pontos extremos,  $\{\bar{e}_1, \dots, \bar{e}_k\}$  (proposições 12.1.15 e 12.1.9 de BERGER, 1987). Além disso, sabemos da mesma referência que qualquer elemento do polítopo pode ser representado como uma combinação convexa desses pontos extremos, ou seja,  $\forall e \in E, \exists \lambda_1, \dots, \lambda_k \geq 0$ :

$$e = \sum_{i=1}^k \lambda_i \bar{e}_i, \quad k \leq N.$$

Com esse resultado, temos que as restrições podem ser escritas usando o conjunto de combinações convexas de medidas equivalentes martingale,  $\text{conv}(\mathcal{M}(S))$ , como

$$E_{\mathbb{Q} \in \text{conv}(\mathcal{M}(S))} [V_T] \leq x.$$

O efeito de ter várias restrições é que a minimização da função do valor dual (4.9) tem que ser feita em dois passos,

$$KL(\tilde{y}) = \inf_{y > 0} \inf_{\mathbb{Q} \in \text{conv}(\mathcal{M}(S))} KL(y, \mathbb{Q}),$$

e escreve-se com a notação de  $Y_T := y \frac{d\mathbb{Q}}{d\mathbb{P}}, y > 0$ , como

$$\begin{aligned} u^*(y) &= \inf_{\mathbb{Q} \in \mathcal{M}(S)} E_{\mathbb{P}} [V(Y_T)] \\ &= \inf_{Y \in \mathcal{Y}(y)} E_{\mathbb{P}} [V(Y_T)], \end{aligned} \quad (4.10)$$

sendo que

$$\mathcal{Y}(y) := \left\{ Y \in L^0(\Omega, \mathcal{F}, \mathbb{P}) \mid Y_T = y \frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_T, Y_0 = y \right\} \quad (4.11)$$

e  $u^*$  é a conjugada de

$$u(x) = \sup_{V \in \mathcal{X}(x)} E_{\mathbb{P}} [U(V_T)], \quad x \in \text{dom}(U),$$

i.e.,

$$\begin{aligned} u^*(y) &= \sup_{x > 0} (u(x) - xy), \quad e \\ u(x) &= \inf_{y > 0} (u^*(y) + xy). \end{aligned}$$

#### 4.2.4 Dualidade no caso geral

Ainda que o exemplo acima deixe a impressão de que a dualidade a ser estabelecida em geral nessa seção parece ser entre funções de utilidades (ou risco) e critérios de escolha entre medidas de probabilidade equivalentes, é importante ressaltar que a conexão é dada

entre o processo de valor no caso primal e processos de derivadas de Radon-Nykodim. De forma mais geral ainda, o problema pode ser visto como a conexão entre conjuntos convexos. Isso é a razão pela qual se encontra a maior parte da literatura relevante sobre o assunto em livros de análise convexa, como ROCKAFELLAR (1970).<sup>11</sup>

A análise do caso discreto finito, além de deixar alguns conceitos mais fáceis para explicar, tem a vantagem de evitar diversos problemas matemáticos. Especialmente nesse caso, vale que:

- Todas as normas são equivalentes e todas as funcionais são representáveis por matrizes.
- Todas as medidas  $\mathbb{P}_1, \mathbb{P}_2, \dots$  em  $(\Omega, \mathcal{F})$  definidas em todos os estados  $\omega \in \Omega$  são equivalentes.
- Todos os espaços  $L^p(\Omega, \mathcal{F}, \mathbb{P}), p \in [0, \infty]$  são isomorfos ao  $\mathbb{R}^N$ .
- Não é necessário diferenciar as noções de martingales entre: semi-martingale, martingale local, martingale.

Para a bipolaridade entre  $\mathcal{X}(x)$  e  $\mathcal{Y}(y)$ , temos que considerar também as duas seguintes observações:

1. É importante saber se  $\text{dom}(U) = \mathbb{R}$  ou  $\mathbb{R}^+$ .
2. Precisamos de uma condição sobre o comportamento de  $U$  no infinito que vai ser dada pela elasticidade assintótica sendo menor que 1.

A maioria dos trabalhos (vide seção 4.3 para referências) pressupõe que a função-objetivo  $U$  mapeia de  $\mathbb{R}^+$  para  $\mathbb{R}$ , enquanto o nosso problema permite valores negativos para a carteira final, i.e.,  $U : \mathbb{R} \rightarrow \mathbb{R}$ . Por causa da concavidade da função de utilidade, faz diferença incorporar o pagamento final  $D_T$  na avaliação final.

No nosso caso de preços comandados por processos de Wiener e uma utilidade atribuída ao valor final da carteira deduzindo o pagamento estocástico, i.e.,  $U(X_T - D_T)$ , podemos seguir a abordagem de ELKAROUI; QUENEZ (1995). Mesmo que esse trabalho pressuponha  $U : \mathbb{R}^+ \rightarrow \mathbb{R}$  com  $U(x) = -\infty, x \leq 0$ , podemos, sem perda de generalidade, deslocar a singularidade de  $U$  para  $K \in \mathbb{R}$ .

---

<sup>11</sup>Vide ROGERS (2001), exemplo 2 e exercício 2, para uma reformulação do problema de super-replicação com funções de utilidade.

**Observação 4.4** *O trabalho de SCHACHERMAYER (2000) é um dos poucos que trata explicitamente o caso  $U : \mathbb{R} \rightarrow \mathbb{R}$  (com os preços sendo semi-martingales), cuja generalidade não é necessária para o nosso caso. Porém, vale mencionar que a prova do caso geral de SCHACHERMAYER (2000) aplica a idéia do deslocamento da singularidade para fora de um intervalo  $[-n, n]$ ,  $n > 0$ , que define o domínio de uma seqüência de utilidades  $U^{(n)}$  convergindo para um limite  $U : \mathbb{R} \rightarrow \mathbb{R}$ .*

Com a possibilidade de aceitar  $0 > V > K$ , podemos analisar o caso sem consumo intertemporal necessário para garantir  $V_T \geq 0$ . Assim, evitamos incluir uma função de utilidade para esses fluxos.

No caso geral, precisamos de uma condição adicional sobre a elasticidade assintótica. Sem esta condição, não é possível garantir que um valor ótimo será atingido, nem para casos relativamente simples (vide KRAMKOV; SCHACHERMAYER (1999) e DELBAEN; SCHACHERMAYER (1998) para vários contra-exemplos).

Do ponto de vista econômico, a condição adicional é suficientemente intuitiva para que possamos aceitá-la sem grandes dificuldades:

**Definição 4.2 (Elasticidade assintótica razoável)** *Uma função de utilidade  $U$  tem a característica de elasticidade assintótica razoável se vale que*

$$AE_{+\infty}(U) := \lim_{x \rightarrow \infty} x \cdot \frac{U'(x)}{U(x)} < 1,$$

$$AE_{-\infty}(U) := \lim_{x \rightarrow -\infty} x \cdot \frac{U'(x)}{U(x)} > 1.$$

Observamos dois fatos importantes em relação a essa condição. Para toda função de utilidade vale, por causa de concavidade, que

$$AE_{+\infty}(U) \leq 1.$$

Sob a hipótese da existência dos limites, temos uma relação com a **aversão ao risco relativa** no sentido de PRATT (1964)

$$r_2(U)(x) := -x \cdot \frac{U''(x)}{U'(x)},$$

que facilita a aceitação da condição. Com o uso da regra de l'Hospital, segue

$$\lim_{x \rightarrow \infty} x \cdot \frac{U'(x)}{U(x)} = 1 - \lim_{x \rightarrow \infty} x \cdot \frac{U''(x)}{U'(x)}$$

e temos que<sup>12</sup>

$$AE_{+\infty}(U) < 1 \Leftrightarrow r_2(U) > 0.$$

Isso significa que um agente com aversão à incerteza relativa positiva de PRATT (1964) também tem elasticidade assintótica razoável. Além disso, a hipótese de a aversão ao risco relativa ser positiva é bem-vinda (veja também seção 6.4).

A condição de elasticidade assintótica razoável tem como objetivo garantir que as soluções dos problemas primal e dual sejam iguais, mas temos que prestar atenção a mais um fato importante: o possível uso de processos de preços gerais, como semi-martingales, impacta, no caso infinito, na caracterização do conjunto de medidas neutralizadoras de risco, vide anexo A.

Sendo  $L^1(\Omega, \mathcal{F}_T, \mathbb{P})$  o conjunto (de todas as classes de equivalência) de  $\mathcal{F}_T$ -mensuráveis variáveis aleatórias integráveis com valores em  $\mathbb{R}$ , definimos o conjunto de pagamentos acessíveis com o valor inicial de  $x > K \in \mathbb{R}$  como<sup>13</sup>

$$\mathcal{V}(x) := \mathcal{V}_K(x) := \{f \in L^1(\Omega, \mathcal{F}_T, \mathbb{P}) \mid f \leq V_T^{x, \pi}, \pi \in \mathcal{A}_K(x)\}.$$

Então, qualquer processo de valor  $V \in \mathcal{V}(x)$  tem que satisfazer a dinâmica (2.4) com  $C_t \equiv 0$  para todo  $t$ . O conjunto  $\mathcal{V}$  tem as seguintes características que são necessárias para a igualdade dos valores ótimos dos problemas duais:<sup>14</sup>

**Lema 4.5** *O conjunto  $\mathcal{V}(x) = \mathcal{V}_K(x)$  satisfaz para qualquer  $D_T \in \mathbb{R}$  e  $K \in \mathbb{R}$ :*

1.  $\mathcal{V}_K(x)$  é convexo.
2.  $\forall \lambda$  com  $\lambda \cdot x > K$ :  $\mathcal{V}_K(\lambda x) = \lambda \cdot \mathcal{V}_K(x)$ .
3.  $\forall g \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ ,  $g > K$  e  $f \in \mathcal{V}_K(x)$  com  $f \geq g$  vale que  $g \in \mathcal{V}_K(x)$ .
4. No caso de  $\text{sign}(x) = \text{sign}(K)$  e  $r_t \geq \frac{1}{t} \ln\left(\frac{K}{x}\right)$ , a função identidade  $\mathbf{1} : \omega \mapsto 1$  é contida no conjunto  $V_K := \bigcup_{x \geq k} \mathcal{V}_K(x) = \bigcup_{x \geq k} x \cdot \mathcal{V}_K(1)$ .

<sup>12</sup>Vide SCHACHERMAYER (2004b), apêndice B.

<sup>13</sup>Para poder aplicar o resultado de SCHACHERMAYER (2000), seria suficiente considerar o conjunto

$$\tilde{\mathcal{V}}(x) := \{f \in L^0(\Omega, \mathcal{F}, \mathbb{P}) \mid f \leq V_T^{x, \pi}, \pi \in \mathcal{A}_K(x), E[|U(f)|] < \infty\}.$$

<sup>14</sup>Esse resultado e a sua prova são em analogia ao desenvolvimento de ROGERS (2001, cap. 4).

**Prova.**

1. Sendo  $X^{(1)}$  e  $X^{(2)}$  processos de valor do patrimônio satisfazendo a equação (2.4) com os respectivos processos de investimento  $\pi^{(1)}$  e  $\pi^{(2)}$ . Com  $\lambda_1 = 1 - \lambda_2 \in [0, 1]$  definimos o processo de valor “médio” tal que

$$\bar{\pi} := \frac{\lambda_1 \pi^{(1)} X^{(1)} + \lambda_2 \pi^{(2)} X^{(2)}}{\lambda_1 X^{(1)} + \lambda_2 X^{(2)}}$$

e

$$\bar{X} := \lambda_1 X^{(1)} + \lambda_2 X^{(2)}.$$

A dinâmica desse valor é dada por

$$d\bar{X}_t = \bar{X}_t \cdot (r_t dt + \bar{\pi}_t ((b_t - r_t \mathbf{1}) dt + \sigma_t dW_t)).$$

Mas como  $\bar{X}_t \leq \max(X_t^{(1)}, X_t^{(2)})$  q.c. para todo  $t \in [0, T]$ , segue que  $\bar{X}_T \leq V_T^{x, \bar{\pi}}$ . Além disso, vale também que  $\bar{X}_t \geq \min(X_t^{(1)}, X_t^{(2)})$  q.c.,  $\forall t \in [0, T]$  e, portanto,  $X_T - D_T \geq K$ . Conseqentemente,  $\bar{X} - D =: \bar{V} \in \mathcal{V}_K(x)$ .

2. Para todo  $\lambda$  com  $\lambda x > K$  vale que qualquer  $V \in \mathcal{V}_K(\lambda x)$  satisfaz  $V_{t_0} = V_0 = \lambda x$ . Mas para  $V \in (\lambda \mathcal{V}_K(x))$ , vale também  $V_0 = \lambda x$ .
3. Segue a partir da parte 1 da prova.
4. Pressupõe  $\pi \equiv 0$ . Então, o processo de valor  $X^{x,0,0}$  satisfaz  $X_{t_0}^{x,0,0} = x \geq K$  e

$$\frac{dX_t^{x,0,0}}{X_t^{x,0,0}} = r_t dt.$$

A solução é dada por  $X_t^{x,0,0} = x \cdot e^{r_t t}$ . No caso de  $\text{sign}(x) = \text{sign}(K)$ ,  $\pi$  somente seria admissível se valesse  $r_t \geq \frac{1}{t} \ln\left(\frac{K}{x}\right)$ . No caso de  $\text{sign}(x) \neq \text{sign}(K)$ , como  $x \geq K$ , a condição de  $r_t \geq \frac{1}{t} \ln\left(\frac{K}{x}\right)$  não é necessária. A existência da função identidade agora é garantida usando as partes 1 e 2.

■

Resta agora formular o problema dual na sua generalidade e determinar as condições sobre o conjunto dual que finalmente eliminarão a folga entre os valores objetivos. Para chegar à formulação do problema dual, precisamos determinar a função lagrangiana que incorpora a restrição (2.4). Por razões dedutivas, pressupomos que  $\mathcal{V}(x) = L^1(\Omega, \mathcal{F}, \mathbb{P})$  é um espaço de Banach. Assim, podemos facilmente identificar o conjunto dos funcionais

$F : \mathcal{V}(x) \rightarrow \mathbb{R}^n$  lineares e contínuos,  $\mathcal{V}^*$ . Esse é também um espaço de Banach sobre  $\mathbb{R}^n$  e qualquer elemento  $F \in \mathcal{V}^*$  pode ser escrito de forma

$$F(y) = \int_0^T Y_s dV_s \quad \forall V \in \mathcal{V}(x), Y \in \mathcal{V}^* \text{ fixo com } Y_0 = y.$$

Então,  $L^\infty(\Omega, \mathcal{F}, \mathbb{P}) = \mathcal{V}^*(y)$  é o espaço dual que é identificado por processos  $Y(x)$ . Esses processos, no nosso contexto, são as densidades de Radon-Nykodim que podem ser usadas para identificar as medidas equivalentes  $\mathbb{Q} \sim \mathbb{P}$ .

Em analogia aos passos feitos no caso discreto (vide equação (4.11)), definimos agora o conjunto dual por

$$\mathcal{Y}(y) := \mathcal{V}^*(y) := \{h \in \mathbb{F} \mid h \geq 0, \forall t \in [0, T] : h_t \leq Y_t, Y_0 = y\},$$

sendo que  $Y$  é a solução da seguinte equação diferencial estocástica

$$dY_t = Y_t \cdot (\beta_t dW_t + \alpha_t dt), \quad (4.12)$$

$$Y_0 = y, \quad (4.13)$$

i.e.,

$$Y_t = y \cdot \exp \left\{ - \int_0^t r_s ds \right\} \cdot Z_t,$$

com  $Z = \{Z_t\}_{t \in [0, T]}$  um  $(\mathbb{F}, \mathbb{P})$ -martingale que ajustará os retornos dos ativos  $S$  para o retorno  $r$  do ativo livre de incerteza  $B$ .

O funcional  $F(y) = \int_0^T Y_s dV_s$  pode ser integrado por partes (vide KARATZAS; SHREVE, 1991, p.155) de forma que

$$\begin{aligned} \int_0^T Y_s dV_s &= V_T Y_T - V_0 Y_0 - \int_0^T V_s dY_s - \langle V, Y \rangle_T \\ &= Y_T (X_T - D_T) - Y_0 x - \int_0^T X_s dY_s - \langle X, Y \rangle_T \\ &= Y_T (X_T - D_T) - Y_0 x - \int_0^T X_s dY_s - \langle X, Y \rangle_T. \end{aligned}$$

sendo que  $\langle V, Y \rangle_T = \int_0^T dV_s dY_s$  é o processo compensador entre  $V$  e  $Y$ . Como queremos incluir a restrição na função objetiva que é de *valor esperado* da utilidade, usando (4.12),

segue

$$\begin{aligned}
E_{\mathbf{P}} \left[ \int_0^T Y_s dV_s \right] &= E_{\mathbf{P}} \left[ V_T Y_T - V_0 Y_0 - \int_0^T V_s dY_s - \langle V, Y \rangle_T \right] \\
&= E_{\mathbf{P}} \left[ V_T Y_T - V_0 Y_0 - \int_0^T V_s (Y_s (\beta_s dW_s + \alpha_s ds)) - \langle V, Y \rangle_T \right] \\
&= E_{\mathbf{P}} \left[ V_T Y_T - V_0 Y_0 - \int_0^T V_s Y_s \alpha_s ds - \langle V, Y \rangle_T \right] \\
&= E_{\mathbf{P}} \left[ V_T Y_T - V_0 Y_0 - \int_0^T V_s Y_s \alpha_s ds \right. \\
&\quad \left. - \int_0^T (r_s V_s ds + \pi_s \cdot [(b_s - r_s \mathbf{1}) ds + \sigma_s dW_s]) \cdot (Y_s (\beta_s dW_s + \alpha_s ds)) \right] \\
&= E_{\mathbf{P}} \left[ V_T Y_T - V_0 Y_0 - \int_0^T V_s Y_s \alpha_s ds \right. \\
&\quad \left. - \int_0^T \pi_s \sigma_s Y_s \beta_s dW_s^2 \right] \\
&= E_{\mathbf{P}} \left[ V_T Y_T - V_0 Y_0 - \int_0^T Y_s \cdot (V_s \alpha_s - \pi_s \sigma_s \beta_s) ds \right].
\end{aligned}$$

Paralelamente, vale também que

$$\begin{aligned}
E_{\mathbf{P}} \left[ \int_0^T Y_s dV_s \right] &= E_{\mathbf{P}} \left[ \int_0^T Y_s (r_s V_s ds + \pi_s \cdot [(b_s - r_s \mathbf{1}) ds + \sigma_s dW_s]) \right] \\
&= E_{\mathbf{P}} \left[ \int_0^T Y_s (r_s V_s ds + \pi_s (b_s - r_s \mathbf{1}) ds) \right] \\
&= E_{\mathbf{P}} \left[ \int_0^T Y_s (r_s X_s ds + \pi_s (b_s - r_s \mathbf{1}) ds) \right] \\
&= E_{\mathbf{P}} \left[ \int_0^T Y_s (r_s X_s + \pi_s (b_s - r_s \mathbf{1})) ds \right]
\end{aligned}$$

e como as duas expressões têm que ser iguais, deduzimos a função lagrangiana como

$$\begin{aligned}
\mathcal{L}(X, Y) &= E_{\mathbf{P}} \left[ U(V_t^{x, \pi}) + \int_0^T Y_s (r_s V_s ds + \pi_s (b_s - r_s \mathbf{1}) ds) \right. \\
&\quad \left. - \left( V_T Y_T - X_0 Y_0 - \int_0^T Y_s \cdot (X_s \alpha_s - \pi_s \sigma_s \beta_s) ds \right) \right] \\
&= E_{\mathbf{P}} [U(X_t^{x, \pi} - D_T) + X_0 Y_0 - V_T Y_T \\
&\quad + \int_0^T Y_s (X_s (r_s + \alpha_s) + \pi_s (b_s - r_s \mathbf{1} + \sigma_s \beta_s)) ds] \\
&= E_{\mathbf{P}} [U(X_t^{x, \pi} - D_T) - Y_T \cdot (X_T^{x, \pi} - D_T) + X_0 Y_0 \\
&\quad + \int_0^T Y_s (X_s (r_s + \alpha_s) + \pi_s (b_s - r_s \mathbf{1} + \sigma_s \beta_s)) ds].
\end{aligned}$$

Introduzindo a função dual convexa de  $U$ ,  $U^*(y) = \sup_x [U(x) - xy]$ , temos

$$\begin{aligned} & \sup_{\pi \in \mathcal{A}(x), X \geq K + D_T} \mathcal{L}(X, Y) \\ &= \sup_{\pi \in \mathcal{A}(x), X \geq K + D_T} E_{\mathbb{P}} \left[ U^*(Y_T) + Y_T D_T + X_0 Y_0 \right. \\ & \quad \left. + \int_0^T Y_s (X_s (r_s + \alpha_s) + \pi_s (b_s - r_s \mathbf{1} + \sigma_s \beta_s)) ds \right]. \end{aligned}$$

Para que possamos realmente calcular o ínfimo sobre  $\sup_{\pi \in \mathcal{A}(x), X \geq K + D_T} \mathcal{L}(X, Y)$ , temos que garantir que esse valor seja finito. As condições necessárias e suficientes para isso acontecer são as condições de folga, que garantem que o supremo não seja infinito. Isso implica que

$$\begin{aligned} b_s - r_s \mathbf{1} + \sigma_s \beta_s &= 0, \\ r_s + \alpha_s &\leq 0, \end{aligned}$$

sendo a primeira condição da maximização sobre  $\pi$  e segunda da maximização sobre  $X_s \geq K + D_T \mathbf{1}_{\{s=T\}}$ .<sup>15</sup>

Se o limite for atingido, temos

$$\begin{aligned} & \sup_{X \geq K + D_T} E_{\mathbb{P}} \left[ U^*(Y_T) + Y_T D_T + X_0 Y_0 + \int_0^T Y_s X_s (r_s + \alpha_s) ds \right] \\ &= E_{\mathbb{P}} \left[ U^*(Y_T) + Y_T D_T + X_0 Y_0 + \int_0^T -Y_s K (r_s + \alpha_s) ds \right] \\ &= E_{\mathbb{P}} \left[ U^*(Y_T) + V_0 Y_0 - K \int_0^T Y_s (r_s + \alpha_s) ds + D_T Y_T \alpha_T \right]. \end{aligned}$$

Definimos

$$u^*(y) := E_{\mathbb{P}} [U^*(Y_T)]$$

e

$$g(y) := D_T Y_T + K \int_0^T Y_s (r_s + \alpha_s) ds$$

para escrever

$$\mathcal{L}(X, Y) = u^*(y) + xy + g(y),$$

pois  $X_0 = x$  e  $Y_0 = y$ . Se não existir folga, teremos que o problema dual é

$$\inf_{y > 0} \mathcal{L}(X, Y) = \inf_{y > 0} u^*(y) + xy + g(y)$$

<sup>15</sup>A segunda condição vem da necessidade de  $X_s (r_s + \alpha_s) < \infty$ . No caso de  $X_s = D_T \mathbf{1}_{\{s=T\}} + K < 0$ , e então  $X_s (r_s + \alpha_s) > 0$ , pode-se pensar em criar uma estratégia tal que  $X_s \rightarrow \infty$ . Mas por causa do limite inferior, isso não seria possível.



e que

$$u(x) = \inf_{y>0} \mathcal{L}(X, Y), \quad x \in \text{dom}(U).$$

A primeira parte dessa igualdade,  $u(x) \leq u^*(y) + xy + g(y)$ , é simples. De fato, como podemos pressupor que<sup>16</sup>

$$E_{\mathbf{P}}[U(V_T)] > -\infty$$

e

$$E_{\mathbf{P}} \left[ U^* \left( y \frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right] < +\infty,$$

e sabemos a partir da bipolaridade de  $\mathcal{X}$  e  $\mathcal{Y}$  que

$$E_{\mathbf{P}}[V_T Y_T] \leq xy, \quad \forall X \in \mathcal{X}, \forall Y \in \mathcal{Y}.$$

Anotamos que mesmo sem folga vale que

$$\sup_{y>0} E_{\mathbf{P}}[V_T Y_T] \leq \inf_{x>0} xy.$$

Então, segue que

$$\begin{aligned} u(x) &= E_{\mathbf{P}}[U(V_T)] \\ &= E_{\mathbf{P}}[U(\tilde{V}_T)] \\ &= E_{\mathbf{P}}[U(\tilde{V}_T) + \tilde{V}_T Y_T - \tilde{V}_T Y_T] \\ &= E_{\mathbf{P}}[U(\tilde{V}_T) - \tilde{V}_T Y_T] + E_{\mathbf{P}}[\tilde{V}_T Y_T] \\ &\leq E_{\mathbf{P}}[U^*(Y_T)] + xy. \end{aligned}$$

Como isso vale para qualquer  $Y \in \mathcal{Y}(y)$ , temos especialmente para

$$\tilde{Y} = \arg \inf_{Y \in \mathcal{Y}(y)} E_{\mathbf{P}}[U^*(Y_T)] + xy,$$

e, portanto,  $u(x) \leq u^*(y) + xy$ .

Do outro lado, mostrar que vale

$$u(x) \geq u^*(y) + x \cdot y + g(y)$$

já é uma tarefa que precisa diversos pré-requisitos matemáticos. Para uma exposição do tratamento desse problema, vide SCHACHERMAYER (2004b).

<sup>16</sup>Senão o resultado seria trivial.

### 4.3 Algumas abordagens existentes na precificação em mercados incompletos

Os pilares centrais do presente trabalho são as novas abordagens para escolher um dos preços livres de arbitragem alternativos que serão analisadas nos capítulos 5 e 6. Por razões comparativas, fazemos agora uma compilação de algumas das propostas mais importantes existentes. Observa-se que se a abordagem implica uma possibilidade de decompor o processo de valor da carteira, podemos diretamente derivar a estratégia ótima a partir dessa decomposição. Essa possibilidade não precisa existir (ou ainda não foi provada) para todas as abordagens.

As abordagens podem ser categorizadas dependendo da forma de escolha da alternativa. Temos as abordagens que trabalham com funções de utilidade (ou de perda) no espaço primal e, via dualidade, existe uma abordagem equivalente de escolha no espaço das medidas equivalentes de martingales. No uso de funções de utilidade, existem duas possibilidades: ou aplica-se diretamente a função de utilidade à incerteza não hedgeável<sup>17</sup> ou combina-se a precificação da incerteza hedgeável com a incerteza não-hedgeável. Além disso, foram propostas abordagens que restringem os preços livres de arbitragem a um subconjunto sem escolher um preço único.

No problema primal, foram usados vários tipos de funções de utilidade, e.g. quadrática, logarítmica, ou classes de funções como CARA<sup>18</sup> e HARA<sup>19</sup> (vide MERTON, 1990; LUNENBERGER, 1998). Em alguns dos trabalhos importantes que também estabeleceram uma dualidade com um critério de escolha entre medidas, a função de utilidade exponencial foi usada, como em DELBAEN; RHEINLÄNDER; SCHWEIZER (2000), ROUGE; ELKAROUI (2000). Referimos também ao trabalho de FRITELLI (2000b), que mostra as formulações duais para uma série de funções de utilidade. No caso de querer controlar a perda esperada, a função de utilidade fica não-regular, vide FÖLLMER; LEUKERT (1999) e FÖLLMER; LEUKERT (2000). O mesmo fenômeno observamos no caso da super-replicação (ELKAROUI; QUENEZ, 1995) de uma pagamento futuro  $D_T$  no qual a

---

<sup>17</sup> Isso seria uma aplicação direta do princípio atuarial, ou seja, o uso do equivalente certo para um mercado totalmente incompleto na noção de FRITELLI (2000b).

<sup>18</sup> *Constant Absolute Risk Aversion*

<sup>19</sup> *Hyperbolic Absolute Risk Aversion*

função de utilidade  $U : \mathbb{R} \rightarrow \mathbb{R}$ ,  $(X_T - D_T) \mapsto U(X_T - D_T)$  teria a forma

$$U(x) = \begin{cases} 0, & x \geq 0, \\ -\infty, & x < 0. \end{cases}$$

No caso de mercados parcialmente completos (vide FRITELLI, 2000b, para uma definição exata), é interessante integrar as abordagens de precificação da incerteza hedgeável com a incerteza não-hedgeável. Uma das primeiras propostas nesse sentido é a do *preço justo* de DAVIS (1998). Pressupondo que o investidor tem uma função de utilidade  $U$  que ele usa para valorar a sua carteira, ele se pergunta, no momento de poder aplicar uma parte do seu patrimônio em derivativos, se a utilidade marginal de substituição da aplicação é positiva. Formalmente, define-se  $D_{t_0}$  como o preço inicial da opção,  $\delta \in \mathbb{R}^+$  o investimento (em unidades monetárias) que o investidor transfere da sua carteira de ativos para o derivativo. Assim, a utilidade esperada final em  $T$  será

$$u(x) := u(x, D_{t_0}, \delta) := \sup_{\pi \in \mathcal{A}(x)} E_{\mathbb{P}} \left[ U \left( X_T^{x, \pi - D_{t_0}, 0} + \frac{\delta}{D_{t_0}} D_T \right) \right].$$

O investidor procura o investimento  $\tilde{\delta} = \tilde{\delta}(D_{t_0})$  em derivativos tal que

$$\left. \frac{\partial u(x, D_{t_0}, \delta)}{\partial \delta} \right|_{\tilde{\delta}} = 0$$

por um preço  $D_{t_0}$  dado. Finalmente, o *preço justo* é aquele  $\tilde{D}_{t_0}$  que deixa o investidor decidir por não aplicar no derivativo, i.e.,  $\tilde{\delta}(\tilde{D}_{t_0}) = 0$ . Seria o preço  $pr_{t_0}^{\text{indif}}$  pelo qual o investidor ficaria *indiferente* entre aplicar no derivativo ou não. Essa proposta representa uma abordagem genérica de precificação de derivativos com funções de utilidade, pois leva em conta o custo alternativo dado pelas oportunidades que cria.

**Observação 4.6** *Como mencionado por DAVIS (1998, p. 217), no caso de um mercado completo, o preço de Black-Scholes é recuperado, usando-se o seguinte argumento: Se for negociado o preço  $p \neq pr_{t_0}^{\text{indif}}$ , o investidor empresta  $\delta$  unidades do numerário para comprar  $\frac{\delta}{p}$  opções e  $x - \delta$  ativos  $S$ . Nesse caso, a taxa de substituição marginal é de*

$$\left. \frac{\partial u \left( x + \delta \left( \frac{pr_{t_0}^{\text{indif}}}{p} - 1 \right), p, \delta \right)}{\partial \delta} \right|_{\delta=0} = \left( \frac{pr_{t_0}^{\text{indif}}}{p} - 1 \right) \cdot u'(x),$$

que é igual a zero se e somente se  $p = pr_{t_0}^{\text{indif}}$ .

Usando uma função de utilidade exponencial, MUSIELA; ZARIPHOUPOULOU (2004) derivaram uma fórmula explícita para o preço de indiferença. Nesse caso, sendo  $\mathbb{Q}$  a única medida martingale equivalente a  $\mathbb{P}$  na restrição à incerteza hedgeável e  $U : \mathbb{R} \rightarrow \mathbb{R}$  definida como  $x \mapsto U(x) = -e^{-\gamma x}$ ,  $\gamma > 0$ , os autores derivaram que

$$pr_{t_0}^{\text{indif}} = E_{\mathbb{Q}} \left[ \frac{1}{\gamma} \log (E_{\mathbb{Q}} [e^{\gamma D_T} | S_T]) \right].$$

Observa-se que a distorção do pagamento do derivativo

$$\tilde{D}_T := \frac{1}{\gamma} \log (E_{\mathbb{Q}} [e^{\gamma D_T} | S_T])$$

com a aversão ao risco implícita na função de utilidade é uma forma de equivalente certo, só que se usa a medida equivalente martingale e um condicionamento às informações do ativo que serve como proteção parcial. Essa distorção é não-linear, como esperamos para uma incerteza não-hedgeável. Porém, o valor esperado  $E_{\mathbb{Q}} [\tilde{D}_T]$  é a forma de precificação linear necessária, como exigido por KREPS (1981). Isso significa que após um ajuste do pagamento final à incerteza não-hedgeável, podemos aplicar uma precificação por não-arbitragem.

Sob algumas condições de regularidade aos processos de preços,<sup>20</sup> DELBAEN et al. (2002) chegaram na seguinte formulação com funções de penalidade:

$$pr_{t_0}^{\text{indif}} = \sup_{\mathbb{Q} \in \mathcal{M}(S)} E_{\mathbb{Q}} [D_T] - \frac{1}{\gamma} (KL(\mathbb{Q} \parallel \mathbb{P}) - KL(\mathbb{Q}^* \parallel \mathbb{P})), \quad (4.14)$$

sendo  $KL : (\Omega, \mathbb{F}) \times (\Omega, \mathbb{F}) \rightarrow \mathbb{R}$

$$(\mathbb{Q}, \mathbb{P}) \mapsto KL(\mathbb{Q} \parallel \mathbb{P}) := \begin{cases} \int \log \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \right) d\mathbb{P} & \mathbb{Q} \ll \mathbb{P}, \\ \infty & \text{caso contrário,} \end{cases}$$

a divergência (direcionada) de Kullback-Leibler<sup>21</sup> e  $\tilde{\mathbb{Q}} : (\Omega, \mathbb{F}) \rightarrow [0, 1]$  a medida martingale equivalente de mínima entropia relativa, i.e.,

$$\tilde{\mathbb{Q}} = \arg \min_{\mathbb{Q} \in \mathcal{M}(S)} KL(\mathbb{Q} \parallel \mathbb{P}).$$

O preço de indiferença tem algumas características interessantes mostradas por FRITELLI (2000b):<sup>22</sup>

<sup>20</sup>Essencialmente a limitação das variações.

<sup>21</sup>Alternativamente, usa-se o termo *entropia relativa*.

<sup>22</sup>Observamos que as características 1 e 2 se deduzem diretamente da função exponencial.

1. Independência do capital inicial  $x$ .
2. Monotonicidade em relação ao parâmetro de aversão à incerteza  $\gamma$ .
3. Homogeneidade no volume do derivativo:  $pr_{t_0}^{\text{Indif}}(D_T \cdot \alpha) = \alpha \cdot pr_{t_0}^{\text{Indif}}(D_T)$ ,  $\alpha > 0$ .
4. Convexidade e monotonicidade no pagamento  $D_T$ .
5. Invariância por translação determinística no pagamento.

Como mencionou BECHERER (2001), estas características servem para que possa ser definida via  $\rho : D_T \mapsto pr_{t_0}^{\text{Indif}}(-D_T)$  uma medida de incerteza convexa para um pagamento  $D_T$  limitado. A conexão com o *preço justo* de DAVIS (1998, fórmula 3) é dada no caso da aversão à incerteza convergindo para zero:<sup>23</sup>

$$\lim_{\gamma \rightarrow 0} pr_{t_0}^{\text{Indif}}(D_T) = E_{\tilde{\mathbb{Q}}}[D_T] = \frac{E_{\mathbb{P}} \left[ U' \left( x + \int_0^T \tilde{\pi} dS \right) \cdot D_T \right]}{U'(x)},$$

sendo que  $\tilde{\pi}$  é a estratégia de *hedging* ótima sob  $\tilde{\mathbb{Q}}$ .

O *coeficiente de Gini*, por ser um outro índice de concentração relativo, tem semelhança com a divergência de Kullback-Leibler e é definida por  $G : (\Omega, \mathbb{F}) \times (\Omega, \mathbb{F}) \rightarrow \mathbb{R}$  com

$$(\mathbb{Q}, \mathbb{P}) \mapsto G(\mathbb{Q} \parallel \mathbb{P}) := \begin{cases} \int \left( \frac{d\mathbb{Q}}{d\mathbb{P}} \right)^2 d\mathbb{P} & \mathbb{Q} \ll \mathbb{P} \\ \infty & \text{c.c.} \end{cases}$$

Este funcional foi usado por MACCHERONI et al. (2004) para criar um método de decisão que se aproxima ao método de uma utilidade de média-variância. Os autores se preocuparam com os defeitos teóricos das preferências implícitas numa função de utilidade quadrática, especialmente com a falta de monotonicidade. Portanto, eles criaram uma função de utilidade  $U_{\text{MMV}}$  que resulta nas mesmas aplicações como a função de utilidade  $U_{\text{MV}}$  no conjunto de monotonicidade de  $U_{\text{MV}}$ . Porém, para valores fora desse conjunto, a  $U_{\text{MMV}}$  continua sendo monótona nos valores da carteira. Mais ainda, MACCHERONI; MARINACCI; RUSTICHINI (2004) mostram que  $U_{\text{MMV}}$  é a melhor aproximação da  $U_{\text{MV}}$  na região de não-monotonicidade de  $U_{\text{MV}}$ , ou seja, para qualquer outra aproximação  $V$  de  $U_{\text{MV}}$  vale que

$$|U_{\text{MMV}}(X) - U_{\text{MV}}(X)| \leq |V(X) - U_{\text{MV}}(X)|,$$

<sup>23</sup>Vide proposição 1.3.4 e observação 1.3.6 em BECHERER (2001).

para  $X$  na região de não-monotonicidade. A função de utilidade  $U_{MMV}$  tem uma definição parecida com o preço de indiferença de 4.14:

$$U_{MMV}(X) := \min_{\mathbb{Q} \in L^2(\mathbb{P})} \left( E_{\mathbb{Q}}[X] + \frac{1}{2k} \cdot G(\mathbb{Q} \parallel \mathbb{P}) \right),$$

sendo que  $L^2(\mathbb{P})$  é o espaço das medidas probabilísticas  $\mathbb{Q}$  que têm uma derivada de Radon-Nikodym em relação a  $\mathbb{P}$  que é duas vezes integrável.

Na mesma direção da proposta de DAVIS (1998) é a análise de FRITELLI (2000a) que define o valor do pagamento futuro  $pr_{t_0}(D_T)$  como a solução da equação

$$\sup_{z: \Pi(z) \leq pr_{t_0}(D_T)} E_{\mathbb{P}}[U(z)] = \sup_{z \in \Theta(pr_{t_0}(D_T))} E_{\mathbb{P}}[U(z)],$$

sendo  $\Pi$  e  $\Theta$  definidos por:

- $\Pi(x)$  o preço livre de arbitragem de um pagamento  $x$  compatível com o funcional  $\Pi$ ,
- $\Theta(x)$  o conjunto de pagamentos futuros que satisfazem a restrição de orçamento  $x$ , ou seja,

$$\Theta(x) := \{y \in \mathcal{F}_T : \Pi(y) \leq \Pi(x)\}.$$

O autor derivou também uma equação que permite uma solução implícita para o valor do pagamento e, no caso de utilidades logarítmicas e *HARA*, conseguiu uma solução explícita (vide também a próxima seção).

COCHRANE; SAÁ-REQUEJO (2000) <sup>24</sup> em vez de escolherem uma única medida equivalente martingale, propuseram restringir as alternativas ao conjunto de medidas que satisfazem uma restrição de variância. Mesmo que todas as alternativas no intervalo de não-arbitragem sejam possíveis escolhas, pareceu-lhes razoável que, a partir de um valor do índice de Sharpe das alternativas, o investimento seria muito atraente. A escolha por índice de Sharpe é fortemente conectada com a abordagem de mínima variância, mas, apesar disso, o objetivo é diminuir o intervalo de não-arbitragem para limites de “bons negócios”. Os limites são ajustados eliminando as carteiras que oferecem um índice de Sharpe *extremamente* alto. O problema é operacionalizar o *extremamente*, pois isso depende da aversão à incerteza do agente, ou seja, da sua função de utilidade.

<sup>24</sup>CERNY; HODGES (2000) introduziram o rigor matemático para a justificação da abordagem.

Uma outra forma de restrição nas oportunidades de investimento foi proposta por BERNARDO; LEDOIT (1999, 2000). O seu conceito de limitar a razão entre ganho e perda baseia-se na idéia de que o valor esperado dos pagamentos excessivos tem que ser zero, i.e., para qualquer pagamento excessivo  $\tilde{D}_T$  vale

$$E_{\mathbf{Q}} \left[ \tilde{D}_T \right] = 0, \forall \mathbf{Q} \in \mathcal{M}(S).$$

Decompondo os pagamentos excessivos na parte positiva (ganhos)  $(\tilde{D}_T)^+$  e negativa (perdas)  $(\tilde{D}_T)^-$ , a condição é equivalente a

$$\frac{E_{\mathbf{Q}} \left[ (\tilde{D}_T)^+ \right]}{E_{\mathbf{Q}} \left[ (\tilde{D}_T)^- \right]} = 1,$$

o que define a razão ente ganho e perda. A conexão entre as duas propostas, apesar de BERNARDO; LEDOIT (1999) fazer uma análise da dualidade entre a limitação da variância do *núcleo de precificação*, i.e., das medidas martingale equivalentes e dos índices de Sharpe dos pagamentos excessivos, já foi mostrada por HANSEN; JAGANNATHAN (1991). No caso de limitar a entropia dos núcleos de precificação, a equivalência com a restrição da utilidade esperada HARA foi analisada em STUTZER (1993).

#### 4.4 Medindo distâncias entre conjuntos

Como já mencionamos antes, a escolha do preço do pagamento futuro  $D_T$  em mercados incompletos é, de um ponto de vista abstrato, equivalente à escolha de uma medida de distância. FRITELLI (2000a) usou a relação de dualidade para definir no espaço das medidas equivalentes uma função que ele chamou de *distância generalizada*. Primeiro, o autor mostra que o valor do problema primal da otimização de utilidade esperada é igual a

$$\Delta(\mathbf{Q}, \mathbb{P}; E_{\mathbf{Q}}[D_T]) := \min_{\lambda \in (0, \infty)} \left\{ \lambda x - E_{\mathbb{P}} \left[ U^* \left( \lambda \frac{d\mathbf{Q}}{d\mathbb{P}} \right) \right] \right\},$$

por um  $\mathbf{Q} \ll \mathbb{P}$  e  $x \in \mathbb{R}$  tal que

$$\sup_{z \in L^- : E_{\mathbf{Q}}[z] \leq 0} E_{\mathbb{P}}[U(D_T + z)] < \sup_{x \in \mathbb{R}} U(x).$$

Depois, no mesmo sentido como na definição do equivalente certo pela inversa da função de utilidade, FRITELLI (2000a) define o funcional

$$\delta(\mathbf{Q}, \mathbb{P}; x) := U^{-1}(\Delta(\mathbf{Q}, \mathbb{P}; x)) - x,$$

e mostra (na proposição 13 de FRITELLI (2000a)) que este funcional satisfaz as condições de uma distância generalizada entre medidas de probabilidades no sentido de CSISZAR (1995). Tendo deduzido  $\Delta$  por dualidade, mas sem especificar a função de utilidade, a medida neutralizadora de risco do investidor depende então da distância induzida pela função de utilidade subjetiva dele.

O conceito de  $\phi$ -divergência de CSISZAR (1967) foi usado por BEN-TAL; TEBOULLE (1987) para ligar a escolha de uma função de utilidade à escolha de uma distância entre medidas de probabilidade equivalentes. Sendo  $\phi \in C^1(\mathbb{R}_+, \mathbb{R})$  convexa e  $F_1, F_2$  funções distribuição com densidades  $f_1, f_2$  em relação à medida de Lebesgue  $\lambda$  em  $\mathcal{B}(\mathbb{R})$ , define-se a  $\phi$ -divergência como:<sup>25</sup>

$$I_\phi(F_1, F_2)(A) := \int_A f_2(x) \phi\left(\frac{f_1(x)}{f_2(x)}\right) \lambda(dx), \forall A \in \mathcal{B}(\mathbb{R}).$$

BEN-TAL; TEBOULLE (1987) mostraram como o problema primal de maximização da utilidade é conectado ao problema de minimizar uma  $\phi$ -divergência correspondente à utilidade escolhida.

GIBBS; SU (2002) analisaram como algumas das mais populares distâncias para medidas de probabilidade são relacionadas. Com  $\mathbb{P}, \mathbb{Q} \in (\Omega, \mathcal{F})$ ,  $f, g$  densidades com respeito a uma medida  $\mu$  finita e dominante de  $\mathbb{P}$  e  $\mathbb{Q}$ , os autores analisaram as seguintes métricas:

1. Discrepância:

$$d_{\text{Discr}}(\mathbb{P}, \mathbb{Q}) := \sup_{A \in \mathcal{F}} |\mathbb{P}(A) - \mathbb{Q}(A)|.$$

2. Distância de Hellinger:

$$d_{\text{H}}(\mathbb{P}, \mathbb{Q}) := \left( \int_{\Omega} (\sqrt{f} - \sqrt{g})^2 d\mu \right)^{\frac{1}{2}}.$$

3. Divergência (direcionada) de Kullback-Leibler, com  $s(\mathbb{P})$  o suporte de  $\mathbb{P}$ :

$$d_{\text{KL}}(\mathbb{P}, \mathbb{Q}) := \int_{s(\mathbb{P})} f \log\left(\frac{f}{g}\right) d\mu.$$

4. Métrica de Kolmogorov:

$$d_{\text{K}}(\mathbb{P}, \mathbb{Q}) := \sup_{x \in \mathbb{R}} |F(x) - G(x)|.$$

---

<sup>25</sup>A  $\phi$ -divergência chama-se também de *núcleo de informação* e foi analisada em maiores detalhes por LIESE; VAJDA (1987).



Utilidade	$U(x)$	$\delta(\mathbb{P}, \mathbb{Q})$
Exponencial	$-e^{-x}$	$d_{\text{KL}}$
Logarítmica	$\ln(x)$	$\exp\{-E_{\mathbb{P}}[\ln(\frac{d\mathbb{Q}}{d\mathbb{P}})]\} - 1$
Power HARA	$\frac{p}{1-p} x ^p$	$x \cdot \left\{ (E_{\mathbb{P}}[(\frac{d\mathbb{Q}}{d\mathbb{P}})^q])^{-\frac{1}{q}} - 1 \right\}$

Tabela 4.1: Conexão entre distâncias generalizadas e funções de utilidade

5. Métrica de Levy:

$$d_L(\mathbb{P}, \mathbb{Q}) := \inf \{ \varepsilon > 0 \mid G(x - \varepsilon) - \varepsilon \leq F(x) \leq G(x + \varepsilon) + \varepsilon, \forall x \in \mathbb{R} \}.$$

6. Métrica de Prokhorov:

$$d_P(\mathbb{P}, \mathbb{Q}) := \inf \{ \varepsilon > 0 \mid \mathbb{P}(B) \leq \mathbb{Q}(B^\varepsilon) + \varepsilon, \forall B \in \mathcal{B}(\mathbb{R}) \},$$

$$\text{com } B^\varepsilon = \{x : \inf_{y \in B} d(x, y) \leq \varepsilon\}.$$

7. Distância de separação para espaços enumeráveis:

$$d_S(\mathbb{P}, \mathbb{Q}) := \max_i \left( 1 - \frac{\mathbb{P}(i)}{\mathbb{Q}(i)} \right).$$

8. Distância de variação total:

$$d_{\text{VT}}(\mathbb{P}, \mathbb{Q}) := \sup_{A \subset \Omega} |\mathbb{P}(A) - \mathbb{Q}(A)|.$$

9. Métrica de Wasserstein:

$$d_W(\mathbb{P}, \mathbb{Q}) := \int_{-\infty}^{\infty} |F(x) - G(x)| dx.$$

10.  $\chi^2$ -distância

$$d_{\chi^2}(\mathbb{P}, \mathbb{Q}) := \int_{s(\mathbb{P}) \cup s(\mathbb{Q})} \frac{(f - g)^2}{g} d\mu.$$

Especificando algumas funções de utilidade, FRITELLI (2000a) identificou as formas das distâncias generalizadas equivalentes nos exemplos da tabela 4.1.

Observa-se que a classe *power HARA* cria alguns dos casos clássicos, e.g.,  $p = 2$ , então  $U(x) = -2x^2$  e

$$\delta(\mathbb{P}, \mathbb{Q}) = \sqrt{E_{\mathbb{P}} \left[ \left( \frac{d\mathbb{P}}{d\mathbb{Q}} \right)^2 \right]} - 1,$$

o que é igual à minimização da variância da derivada de Radon-Nikodym.

BEN-TAL; TEBoulLE (1987) fizeram também uma compilação de algumas correspondências, neste caso entre funções de utilidade e *núcleos de informação*, vide tabela 4.2.

Utilidade	$U(x)$	Distância
Exponencial	$1 - e^{-x}$	$d_{\text{KL}}$
HARA	$\frac{x}{1+x}, x > -1$	$d_{\text{H}}$
./.	$1 - \left(1 - \frac{x}{p}\right)^p, \frac{1}{p} + \frac{1}{q} = 1, p < 1$	divergência de ordem $\alpha$
./.	$\log(2 - e^{-x})$	entropia de Bose-Einstein

Tabela 4.2: Conexão entre *núcleos de informação* e funções de utilidade

A  $\alpha$ -ordem divergência é definida pela função de *núcleo*

$$\phi_{\alpha}(x) := \frac{1}{\alpha - 1} x^{\alpha} - \frac{x}{\alpha - 1} + 1, \quad \alpha \in (0, 1),$$

e a entropia de Bose-Einstein por

$$\phi_{\text{BE}}(x) := x \log x - (1 + x) \log(1 + x) + (1 + x) \log 2, \quad x \geq 0.$$

# Capítulo 5

## Alocação ótima com erros de estimação

### 5.1 Introdução

Neste capítulo, desenvolvemos uma abordagem que permite incluir e avaliar a incerteza inerente aos parâmetros numa otimização de carteiras. No problema clássico de otimização de carteiras no ambiente média (esperança)-variância de MARKOWITZ (1952), o investidor tem certeza sobre a forma da distribuição dos retornos dos ativos. Uma aplicação direta de um modelo com essas premissas pode criar graves problemas no caso em que se realizarem outros valores além dos estimados, sugerindo que o investidor está sujeito a uma otimização de erros (vide MICHAUD, 1989). O impacto de ignorar a incerteza nos momentos pode ser melhor visualizado no problema de “substituição” da figura 5.1.

Pelo fato de o ativo 1 ter um valor esperado  $\mu_1$  levemente maior de que o ativo 2 ( $\mu_2$ ) com o mesmo desvio-padrão ( $\sigma_1 = \sigma_2$ ), o otimizador de Markowitz (com restrições de não-alavancagem) atribuirá 100% do peso da carteira no primeiro ativo ( $\pi_1 = 1$ ). Na prática, os valores  $\mu_1$  e  $\mu_2$  não são tão certos para que o investidor aceitasse a proposta. Intuitivamente, preferiríamos colocar quase a metade dos recursos em cada ativo.

Na aplicação prática, pergunta-se como criar expectativas sobre os parâmetros. No uso dos dados históricos, LUENBERGER (1998, capítulo 8.5) destaca que é ainda razoavelmente possível estimar, por exemplo, a variância a partir dos dados históricos. Porém, estimativas tradicionais para a média são altamente irrealistas, pois a quantidade de dados necessários não é razoável.<sup>1</sup> No nosso exemplo de aplicações no mercado financeiro

---

<sup>1</sup>O autor LUENBERGER (1998, p.215) chama o efeito de *Blur of History*, e estima que seriam

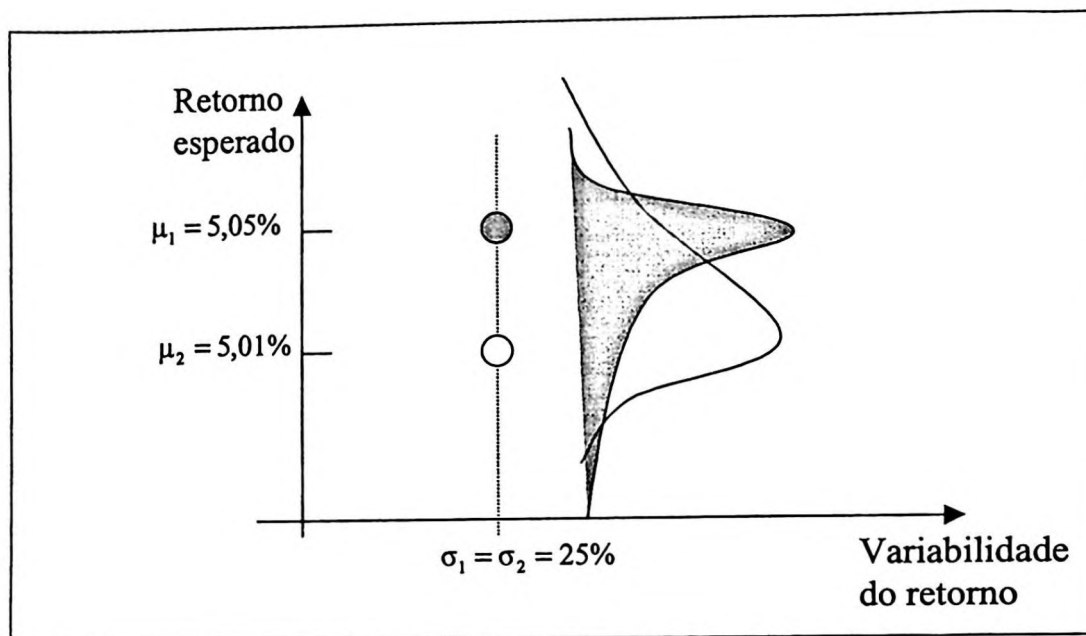


Figura 5.1: Problema de substituição

brasileiro, usaremos, portanto, estimativas subjetivas com erros subjetivos, fornecidos por analistas do mercado.

A seguir, discutiremos algumas propostas que propõem incluir a incerteza na otimização e veremos que o problema de otimização precisa de um critério de escolha para a variável de incerteza. Na seção 5.2, vamos rever algumas das propostas feitas para lidar com o problema de maximização de erros. A maior parte dos trabalhos incorpora a incerteza sobre parâmetros da distribuição dos retornos dos ativos nas suas estimações, o que implica parâmetros ajustados. Além de pressupor os parâmetros como estocásticos, podemos também pressupor isso para toda a distribuição, o que será tratado no capítulo 6. No capítulo atual, tratamos o caso em que os momentos da distribuição dos retornos são especificados por um conjunto de parâmetros  $\theta$  pertencendo ao conjunto  $\Omega_\theta$ .

A maior parte da literatura preocupa-se com a estimação dos parâmetros de valor esperado e variância. A razão para minimizar o risco de estimação é que os retornos estimados são viesados para cima, enquanto o risco é viesado para baixo (JOBSON; KORRIE, 1980). O comportamento real de gestores de patrimônio parece estar conforme com essa observação. O impacto geral de incerteza na estimação foi analisado, entre outros, por DICKENSON (1979), FRANKFURTER; PHILLIPS; SEAGLE (1971), JOBSON;

necessários por volta de 156 anos para obter uma estimativa da média com um desvio-padrão do estimador de 10% do valor esperado.

KORKIE (1980), BARRY (1974), BAWA; BROWN; KLEIN (1979), CHOPRA; ZIEMBA (1993), BROWN (1979) e KALYMON (1971).

O grau de severidade do erro na estimação varia entre as variáveis: o erro na estimação dos retornos é visto como o mais grave (vide BEST; GRAUER, 1991, por uma análise detalhada), seguido pelos erros na estimação da variância (usada como medida de risco), enquanto o erro na estimação da correlação é visto como de gravidade intermediária (vide MERTON, 1980). A má especificação da função de utilidade não é tão importante, enquanto o nível da aversão absoluta de Arrow-Pratt é especificado corretamente (vide KALLBERG; ZIEMBA, 1984).

A formulação abstrata para incorporar incerteza sobre as distribuições na construção de uma carteira eficiente tem que ser feita num espaço de produto. Chamamos de  $(\Omega, \mathbb{F}, \mathbb{P})$  o espaço probabilístico para o vetor de retornos dos ativos  $S = (S^{(1)}, \dots, S^{(n)})$ . Pressupomos que a distribuição conjunta de  $S$  pode ser parametrizada por um vetor  $\theta = (\theta^{(1)}, \dots, \theta^{(n)})$  de variáveis aleatórias (parâmetros aleatórios). Seja  $(\Omega_\theta, \mathbb{F}_\theta, \mathbb{P}_\theta)$  o espaço probabilístico para os parâmetros  $\theta \in \Omega_\theta$ . Podemos formular o problema por junção de  $\sigma$ -álgebras, ou seja, se definimos  $\mathcal{G}(t) := \mathcal{F}(t) \cup \mathcal{F}_\theta(t)$  e  $\mathbb{G} := \{\mathcal{G}(t)\}_{t \in [0, T]}$ , temos que

$$\frac{dS(t, \theta)}{S(t, \theta)} \in \mathcal{G}(t), \forall t \in [0, T].$$

Se pressupomos a situação de “desconexão”<sup>2</sup>, i.e., a existência de uma medida  $\mathbb{Q} \sim \mathbb{P}$  tal que  $\mathcal{F}(T)$  e  $\mathcal{F}_\theta(T)$  são independentes, podemos fazer uma transformação de medida para “o mundo”  $\mathbb{Q}$  e trabalhar com um modelo de espaço produto. De outra forma, poderíamos trabalhar com um espaço produto sob  $\mathbb{P}$  diretamente, se aceitássemos que as informações dos ativos não influenciam as incerteza sobre os parâmetros da distribuição conjunta  $S$  e vice versa.<sup>3</sup> Nesse caso, num modelo de Wiener com  $\theta(t) = (\mu(t), \Sigma(t))$ , os processos de

<sup>2</sup>O termo original em Inglês é “*decoupling*”. A idéia é expressar a possibilidade de separar os conjuntos  $\mathcal{F}$  e  $\mathcal{G}$  para que possamos trabalhar sob a condição de independência.

<sup>3</sup>Para mais detalhes sobre o problema matemático de aumento, vide AMENDINGER (1999).

preço, valor esperado e volatilidade seguiriam<sup>4</sup>

$$\begin{aligned}\frac{dS(t)}{S(t)} &= \mu(t) dt + \Sigma(t) dW^{(1)}(t), \\ \frac{d\mu(t)}{\mu(t)} &= m(t) dt + s(t) dW^{(2)}(t), \\ \frac{d\Sigma(t)}{\Sigma(t)} &= M(t) dt + \sigma(t) dW^{(3)}(t),\end{aligned}$$

com  $\langle dW^{(i)}, dW^{(j)} \rangle_t = 0, \forall i \neq j, t \in [0, T]$ .

Mas, no nosso caso,  $\mathcal{F}$  depende de  $\mathcal{F}_\theta$ , e podemos no máximo pressupor a situação de “desconexão”. De qualquer forma, temos que lidar com o espaço produto  $\mathbb{P}_\theta \times \mathbb{P}$ .<sup>5</sup> Assim, o problema de decisão do investidor torna-se

$$\max_{\pi \in \mathcal{A}(x)} E_{\mathbb{P}_\theta \times \mathbb{P}} \left[ U_\theta \left( V_T^{x, \pi, 0, D} \right) \right], \quad (5.1)$$

sendo que  $U_\theta$  é uma função de utilidade que leva em conta a incerteza sobre os parâmetros da distribuição. Mesmo sendo bem geral, essa formulação é difícil de ser implementada. Ela somente nos servirá para reformular algumas das abordagens existentes de uma maneira unificada.

Na análise das abordagens existentes, observamos que a maior parte delas determina, num primeiro passo, um parâmetro ótimo  $\tilde{\theta}$  que será usado na determinação da alocação ótima:  $\tilde{\pi} = \tilde{\pi}(\tilde{\theta})$ , i.e., com a ajuda de uma *função de perda*<sup>6</sup>  $L_{\tilde{\theta}} : (\Omega_\theta, \mathcal{F}_\theta, \mathbb{P}_\theta) \rightarrow \mathbb{R}$ , escolha-se o parâmetro ótimo em relação a um parâmetro  $\bar{\theta}$  inicial:<sup>7</sup>

$$\tilde{\theta} = \arg \min_{\bar{\theta}} E_{\mathbb{P}_\theta} [L_{\bar{\theta}}(\theta)].$$

Com o parâmetro ótimo determina-se a alocação ótima em pesos da carteira

$$\tilde{\pi} = \arg \max_{\pi \in \mathcal{A}(x)} E_{\mathbb{P}|\tilde{\theta}} \left[ U_{\tilde{\theta}} \left( V_T^{x, \pi, 0, D} \right) \right].$$

Como a função de perda  $L$  representa a aversão do investidor à incerteza na distribuição e  $U_\theta$  a aversão dele à incerteza do valor final, juntos, elas fazem o mesmo

<sup>4</sup>Nessa formulação sem rigor matemático, sejam  $m$  e  $M$  os processos determinísticos de médias da média e da variância (respectivamente) do processo de preços. As funções  $s$  e  $\sigma$  representam os desvio-padrões dos mesmos objetos.

<sup>5</sup>É importante lembrar que o espaço conjunto  $\mathbb{P} \times \mathbb{P}_\theta$  não é simplesmente o espaço produtivo, pois  $\mathbb{P}_\theta$  influencia  $\mathbb{P}$ . Por outro lado,  $\mathbb{P}_\theta$  é independente de  $\mathbb{P}$ .

<sup>6</sup>Essa função de perda pondera a aversão do investidor à incerteza nos parâmetros, ou seja, a aversão aos erros de estimação.

<sup>7</sup>Por exemplo, média e covariância históricas.

trabalho que  $U$  na otimização combinada (5.1). Em vez de aplicar o procedimento de dois passos, ZIMMER; NIEDERHAUSER (2004) propuseram analisar

$$\max_{\pi \in \mathcal{A}(X)} \int_{\Omega_\theta} \left( E_{\mathbb{P}|\theta} \left[ U \left( V_T^{x,\pi,0,D} \right) \right] - B_\delta(\pi, \theta) \right) d\mathbb{P}_\theta,$$

observando a conexão entre as formulações dos problemas de otimização padrão e com funções de penalidade.

## 5.2 Abordagens existentes

Nessa seção, fazemos uma revisão parcial da literatura sobre como é feita a incorporação da incerteza sobre os parâmetros. As abordagens existentes se concentram na distorção da distribuição inicial dos retornos dos ativos e pretendem escolher uma distorção conveniente. Usando a distribuição escolhida, a abordagem tradicional de MARKOWITZ (1952) é aplicada.<sup>8</sup>

Uma das primeiras propostas de como evitar os resultados irrealistas causados pela aplicação direta da otimização média-variância foi de limitar os possíveis pesos (e.g. FROST; SAVARINO, 1988; EICHHORN; GUPTA; STUBBS, 1998; GRAUER; SHEN, 2000). O efeito dessa limitação é que as restrições têm efeito, o que significa que, no pior dos casos, o investidor impôs diretamente a alocação final. Neste caso, de já conhecer a alocação ótima, ele nem precisava ter usado um otimizador. Os autores argumentam, portanto, que o uso de limites evita ficar muito exposto a erros de estimação, pois retornos altos são, muitas vezes, superestimados e retornos baixos são, muitas vezes, subestimados. Apesar de que esse erro de estimação deveria ser tratado com uma abordagem bayesiana, que pode incorporar a estrutura existente de forma teoricamente coerente, pergunta-se ainda, qual deveria ser o limite por ativo. Não existe algoritmo claro de como escolher o limite correto, pois essa proposta não consegue quantificar a incerteza nas estimações. Sem ter maiores informações sobre como determinar os limites, pode-se pensar em impor o mesmo limite para cada ativo. Se todos os limites têm efeito, chegamos à regra de pesos iguais, analisada por exemplo por CHEN; BROWN (1983). Eles propuseram resolver o seguinte programa

$$\max_{\theta \in \Omega_\theta} \left( \max_{\pi \in \mathcal{A}(x)} E_{\mathbb{P}|\theta} \left[ U \left( V_T^{x,\pi,0,D} \right) \right] - \|\pi\|_2^2 \right).$$

<sup>8</sup>Lembramos que existem abordagens alternativas como a de MICHAUD (1998) e vamos destacar as diferenças.

Essa abordagem permite calcular diretamente a carteira ótima se o parâmetro  $\tilde{\theta}$  for especificado de maneira média-variância:

$$\begin{aligned}\tilde{\pi} &= \tilde{\pi}(\tilde{\theta}) \\ &= \arg \max_{\pi \in \mathcal{A}(x)} E_{\mathbf{P}|\tilde{\theta}} \left[ U \left( V_T^{x,\pi,0,D} \right) \right] - \lambda \|\pi\|_2^2 \\ &= - \arg \min_{\pi \in \mathcal{A}(x)} \left\{ \pi' \Sigma_{\tilde{\theta}} \pi + \lambda \pi' \pi - \lambda_1 \pi' (\mu_{\tilde{\theta}} - \mu) \right. \\ &\quad \left. - \lambda_2 (\pi' \mathbf{1} - 1) \right\},\end{aligned}$$

sendo que  $\Sigma_{\tilde{\theta}}$  é a matriz de covariâncias e  $\mu_{\tilde{\theta}}$  é o vetor de médias, os dois escolhidos respeitando a realização do parâmetro (ótimo)  $\tilde{\theta}$ . Observa-se que incluímos somente as

restrições de cone com a notação de  $\mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ . Neste caso, a solução é explicitamente

calculável como

$$\tilde{\pi}(\tilde{\theta}) = \frac{\hat{\Sigma}_{\tilde{\theta}}^{-1} \mu_{\tilde{\theta}}}{\mathbf{1}' \hat{\Sigma}_{\tilde{\theta}}^{-1} \mu_{\tilde{\theta}}},$$

com a notação  $\hat{\Sigma}_{\tilde{\theta}} := (\Sigma_{\tilde{\theta}} + \lambda \mathbf{I})$ . Alguns aspectos dessa abordagem se encontram também na nossa proposta dada na seção 5.5.4.

### 5.3 O uso da função de perda e estatística de Bayes

Quando confrontamos incerteza nos parâmetros (ou até na função distribuição), uma das ferramentas mais usadas é a estatística de Bayes, que usa a *função de perda* para ponderar as diversas alternativas.<sup>9</sup>

A função de perda, já mencionada na introdução dessa seção, mede o impacto de uma má especificação ao calcular o quanto o decisor “perde” quando usa a especificação errada. Como o valor atribuído a uma perda é subjetivo, é claro que não existe função de perda única.

CHOPRA; ZIEMBA (1993) calcularam, por exemplo, a perda em equivalente certo usando uma função de utilidade exponencial para o valor da carteira e pressupondo uma distribuição multinomial. Segundo essa abordagem, quando o investidor escolhe a especificação paramétrica  $\tilde{\theta}$ , mas realiza-se um outro parâmetro  $\theta$ , a perda que lhe ocorre em

<sup>9</sup>Vide, por exemplo, ZELLNER (1987).



unidades de equivalente certo é

$$L(\theta, \bar{\theta}) = \frac{CE_{\theta} - CE_{\bar{\theta}}}{CE_{\bar{\theta}}}.$$

JORION (1986) também usa a variação relativa como função de perda, mas troca o equivalente certo pelo valor esperado

$$L(\theta, \bar{\theta}) = \frac{E_{\mathbb{P}|\theta} [U(V_T^{x,\pi,0,D})] - E_{\mathbb{P}|\bar{\theta}} [U(V_T^{x,\pi,0,D})]}{E_{\mathbb{P}|\bar{\theta}} [U(V_T^{x,\pi,0,D})]}.$$

Uma alternativa na função de perda é usada por HORST; ROON; WERKER (2001), que não analisam a variação relativa, mas uma função quadrática

$$L(\theta, \bar{\theta}) = \mu'_{\theta} (\bar{\pi}(\theta) - \bar{\pi}(\bar{\theta})) - \frac{1}{2} \gamma (\bar{\pi}(\theta) - \bar{\pi}(\bar{\theta})) \Sigma_{\theta} (\bar{\pi}(\theta) - \bar{\pi}(\bar{\theta})),$$

$\gamma \in \mathbb{R}^+$ .

Com a função de perda especificada, o decisor escolhe uma distribuição alternativa  $\mathbb{P}_{\bar{\theta}}$  a partir do parâmetro  $\bar{\theta}$  que foi obtido da minimização de uma perda esperada (vide ZELLNER, 1987)

$$\bar{\theta} = \arg \min_{\bar{\theta}} \left\{ \int L(\theta, \bar{\theta}) d\mathbb{P}_{\theta} = E_{\mathbb{P}_{\theta}} [L(\theta, \bar{\theta})] \right\},$$

lembrando que  $\mathbb{P}_{\theta}$  foi definida como a medida de probabilidade no espaço  $(\Omega_{\theta}, \mathcal{F}_{\theta})$  para o parâmetro  $\theta \in \Omega_{\theta}$ .

Estamos agora confrontando com um problema parecido com o original, de achar uma carteira  $\pi$  que maximiza para um nível de risco dado o retorno do investidor, somente num outro espaço. Usa-se uma função de “utilidade”  $L$  para os valores finais das “carteiras” criadas a partir das estratégias  $\theta$ .

Tendo otimizado o problema da incerteza, os parâmetros escolhidos serão usados como variáveis de entrada para o problema original dos retornos, sendo variáveis aleatórias.

Se a estratégia é parametrizada por um parâmetro de aversão à incerteza  $\gamma$ , pode ser deduzido o parâmetro implícito a partir de uma escolha do agente. No caso de HORST; ROON; WERKER (2001), começa-se com

$$\pi = \pi(\gamma) = (\gamma \Sigma)^{-1} (\mu - \Sigma \mathbf{1}),$$

e eles deduzem que o parâmetro  $\gamma$  de aversão à incerteza após consideração da incerteza sempre tem que ser maior ou igual ao parâmetro sem incerteza  $\tilde{\gamma}$ . Quanto maior o número de ativos  $K$  e quanto menor as observações históricas  $T$ , maior o novo índice

$$\gamma = \tilde{\gamma} \cdot \left(1 + c \frac{K-1}{T}\right),$$

e a existência da incerteza somente aumenta a aversão por pressupor incerteza. Intuitivamente, esse resultado era esperado (vide seção 6.4) e é coerente com as conclusões de outros autores: incerteza de estimação aumenta a incerteza da carteira (vide FROST; SAVARINO, 1986), mas o impacto diminui quando o tamanho da amostra usada para estimação aumenta (vide KEMPF; KREUZBERG; MEMMEL, 2001).

O mesmo resultado foi deduzido por LAUPRETE (2001), mas sem o foco no valor ótimo do parâmetro de aversão ao risco. A perda ótima (assintótica) com o parâmetro ajustado de aversão à incerteza,  $\gamma$ , pode ser calculada como

$$L(\pi(\gamma)) = \pi(\tilde{\gamma})' \Sigma \cdot \frac{K-1}{T}.$$

Resultados equivalentes foram deduzidos para o caso de desvio médio absoluto como medida de variabilidade e para a distribuição  $t$  de Student multivariada por LAUPRETE (2001) e LAUPRETE; SAMAROV; WELSCH (2003). O parâmetro alternativo de aversão ao risco pode ser interpretado como o parâmetro original que o investidor teria se tivesse incorporado a incerteza na distribuição.

A idéia principal nas abordagens bayesianas é a de incorporar todas as informações existentes na estimação do retorno esperado. No caso em que existem vários ativos no universo de aplicação do investidor, os retornos passados de **todos os ativos** devem ser usados na estimação do retorno esperado de cada um desses ativos. O valor esperado estimado dessa forma é diferente do valor esperado como se fossem somente usados os retornos históricos do ativo sob consideração. O estimador mais famoso é o de James-Stein que, por causa da sua forma, é chamado um estimador de *compressão*<sup>10</sup>

$$\hat{\mu}_{\text{shrinkage}} = \alpha \cdot \mu + (1 - \alpha) \bar{\mu},$$

sendo que  $\bar{\mu}$  é uma estimativa alternativa. STEIN (1962) propôs um  $\alpha$  de tal forma que o desempenho da estimação em relação a uma função de perda quadrática aumente. BERGER (1978), por sua vez, usou uma função de perda polinomial para determinar o

<sup>10</sup>*Shrinkage* em Inglês.

$\alpha$ . No âmbito de otimização de carteiras, uma das primeiras aplicações do estimador de *compressão* foi feita por JORION (1986).

Usando a idéia de incluir o máximo de informação na estimação podem ser usados não somente retornos históricos de outros ativos, mas também previsões do investidor, por exemplo. Assim, uma distribuição prévia dos retornos pode ser atualizada pelas estimativas novas do investidor. Essa abordagem foi usada por BLACK; LITTERMAN (1990) e BLACK; LITTERMAN (1992) na escolha de composições de carteiras de renda fixa internacional. Apesar das críticas práticas dessa proposta,<sup>11</sup> é importante ressaltar que a mistura de distribuições é uma forma geral do estimador de *compressão*.

Na abordagem de *compressão*, a função de perda foi usada para decidir qual **parâmetro** alternativo da distribuição deve ser escolhido. Com a divergência de Kullback-Leibler, temos um critério de escolha entre distribuições que é não-paramétrico. Ela foi usada por ANDERSON; HANSEN; SARGENT (2000) e MAENHOUT (2001) no ambiente dinâmico como função de perda. Teoricamente, qualquer medida de distância pode ser usada como função de perda, como já deveria ser claro a partir da discussão na seção 4.4.

O exemplo do pior caso (i.e., super-replicação) foi aplicado na determinação dos parâmetros alternativos por GOLDFARB; IYENGAR (2002), IYENGAR (2002), e COSTA; PAIVA (2002), sendo o critério de escolher a especificação para o pior caso.<sup>12</sup>

## 5.4 O método de simulação (*reamostragem*)

Diferente das abordagens em dois passos da seção anterior, MICHAUD (1998) propôs uma forma de criar uma fronteira eficiente que ele chamou de fronteira "*reamostrada*".<sup>13</sup> Como a fronteira é criada por várias carteiras, é suficiente analisar como é criada uma carteira "*reamostrada*". Para cada especificação alternativa dos parâmetros  $\theta \in \Omega_\theta$ , uma carteira ótima  $\tilde{\pi}_\theta$  é determinada, usando o otimizador média-variância típico. Depois, calcula-se

---

<sup>11</sup>Por exemplo, pergunta-se porque o mercado não tem incerteza, mas somente a estimativa do investidor, ou como se determina a carteira do mercado etc.

<sup>12</sup>Observa-se que a especificação do pior caso é chamada de "*robusta*", mas como esta nomenclatura é usada também para o uso de funções de perda em geral na Estatística, preferimos não aplicar o termo para a abordagem de pior caso.

<sup>13</sup>OLIVEIRA; FERRAZ (2001) usaram o método de *bootstrap* para simular as distribuições dos ativos sob consideração que, nesse caso, eram fundos.

a média dos pesos de cada carteira ótima criada a partir do conjunto dos parâmetros:

$$\begin{aligned}\tilde{\pi} &= E_{\mathbf{P}_\theta} [\tilde{\pi}_\theta] \\ &= \int_{\Omega_\theta} \arg \max_{\pi \in \mathcal{A}(x)} E_{\mathbf{P}|\theta} \left[ U \left( V_T^{x,\pi,0,D} \right) \right] d\mathbb{P}_\theta.\end{aligned}$$

Como a média das composições é calculada e não uma única composição a partir de um conjunto de parâmetros ótimos, a *reamostragem* é diferente das abordagens de função de perda descritas na seção anterior. Mesmo assim, pode-se interpretar a *reamostragem* como uma abordagem bayesiana, pois ela considera o espaço dos parâmetros alternativos para escolher a especificação final.

Um ponto interessante é que na determinação da carteira  $\tilde{\pi}$  foi utilizado nenhum critério de aversão a resultados não desejados, mas simplesmente a média. Já que os pesos se tornam variáveis aleatórias, esse aspecto deveria ser considerado. Mais ainda, observando que a distribuição dos pesos pode, em muitos casos, ter uma curtose grande, a média torna-se uma estatística que explica a distribuição, vide a figura 5.2.

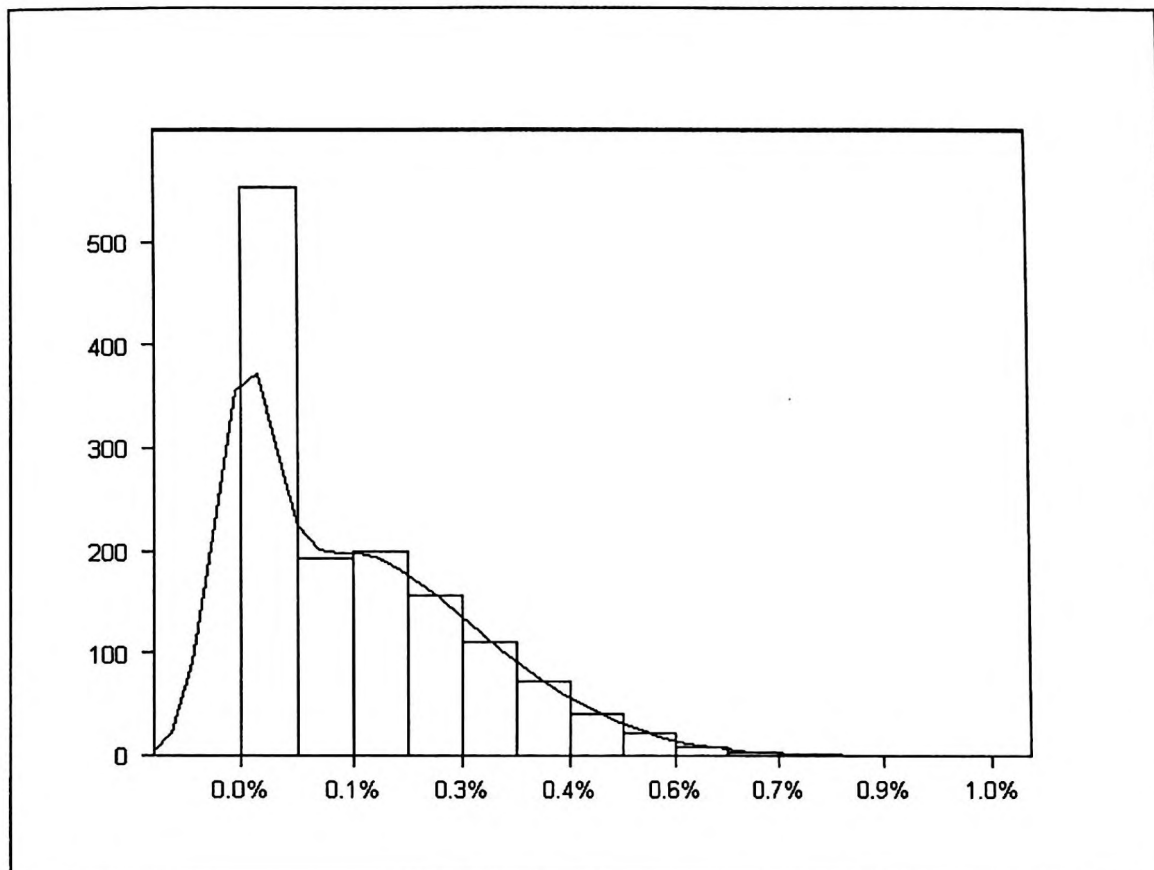


Figura 5.2: Gráfico das frequências absolutas (ordenadas) dos pesos do ativo IBOVESPA (abscissas) na carteira numa simulação com 1500 amostras.

Uma alternativa intuitiva seria escolher os pesos que tenham maior probabilidade, i.e., de aplicar o critério de máxima verossimilhança:

$$\tilde{\pi}_{\text{ML}} = \arg \max_x \mathbb{P}_\theta (\omega \in \Omega_\theta | x - \varepsilon < \tilde{\pi} \leq x).$$

O problema é que temos que garantir que a soma de todos os pesos seja 1:

$$\tilde{\pi}_{\text{ML}} = \sum_{i=1}^n \tilde{\pi}_{\text{ML}}^{(i)} = 1.$$

## 5.5 Abordagens alternativas usando funções de penalidade

Tendo em vista que MICHAUD (1998) determinou a composição média de todas as carteiras criadas com as diferentes distribuições de retornos, queremos focalizar a atenção

ao fato que isso pode ser interpretada como uma abordagem de mínima distância. De fato, a carteira média pode ser vista como a carteira que tem a mínima distância linear com todas as outras carteiras. De forma equivalente, não existe outra carteira que, na média, é mais próxima de todas as carteiras alternativas conjuntas.

Quando queremos levar em conta não somente a distância média, mas outros aspectos distribucionais do conjunto de carteiras alternativas, surgem outras medidas de distância que podem ser usadas (vide seção 4.4). Também é possível formular a função-objetivo usando funções de penalidade. Nesse caso, usamos uma função que penaliza desvios da média de uma certa forma funcional, por exemplo,

$$\max_{\pi \in \mathcal{A}(x)} E_{\mathbf{P}} [V^{x,\pi,0,D}] - B(V^{x,\pi,0,D}),$$

sendo que  $B(x)$  pode ser uma norma ou outra função de distância escolhida. O objetivo dessa seção é propor duas funções de penalidade que levam em conta o possível desejo do investidor de não mudar muito a composição da carteira. Obviamente, isso implica abordagens de otimização de carteiras como problema primal e não de medidas neutralizadoras de risco. Começamos com a formulação explícita dos casos de uma aversão quadrática à incerteza e do mínimo desvio da utilidade esperada.

### 5.5.1 Aversão quadrática à incerteza

Uma das formas mais usadas para calcular o desvio de uma variável aleatória da sua média é o uso da variância que penaliza de forma simétrica valores muito diferentes da média. A penalidade como função, dependendo da carteira escolhida e dos parâmetros, pode ser escrita como

$$B(\pi, \theta) := \left( E_{\mathbf{P}|\theta} \left[ U \left( V_T^{x,\pi,0,D} \right) \right] \right)^2 - \left( \int_{\Omega_\theta} E_{\mathbf{P}|\theta} \left[ U \left( V_T^{x,\pi,0,D} \right) \right] d\mathbb{P}_\theta \right)^2.$$

### 5.5.2 Minimização do desvio da melhor escolha

Na escolha de uma carteira média, temos o problema de desviar das carteiras ótimas de cada caso. Podemos querer minimizar a distância da carteira ótima em cada realização. A função de penalidade, em unidades de utilidade esperada, torna-se

$$B(\pi, \theta) := \left\| E_{\mathbf{P}|\theta} \left[ U \left( V_T^{x,\pi,0,D} \right) \right] - \max_{\pi_\theta \in \mathcal{A}(x)} E_{\mathbf{P}|\theta} \left[ U \left( V_T^{x,\pi_\theta,0,D} \right) \right] \right\|_2^2. \quad (5.2)$$

Essa abordagem segue o mesmo espírito de JORION (1986) em medir o desvio do melhor resultado por estado. Ela obriga escolher uma carteira perto das carteiras que são ótimas em estados com utilidades esperadas altas.

Podemos agora calcular diretamente a partir de (5.2) a composição ótima

$$\tilde{\pi} = \arg \min_{\pi} \int_{\theta \in \Omega_{\theta}} \left\{ E_{\mathbb{P}|\theta} [U(\langle X, \pi \rangle)] - \max_{\pi \in \mathcal{A}(X|\theta)} E_{\mathbb{P}|\theta} [U(\langle X, \pi_{\theta} \rangle)] \right\}^2 d\mathbb{P}_{\theta}. \quad (5.3)$$

Observe-se que, no caso de uma utilidade quadrática com

$$E_{\mathbb{P}|\theta} [U(\langle X, \pi \rangle)] = \langle \mu_{\theta}, \pi \rangle - \frac{\gamma}{2} \langle \pi' \Sigma_{\theta}, \pi \rangle,$$

a fórmula pode ser escrita como<sup>14</sup>

$$\begin{aligned} \tilde{\pi} &= \arg \min_{\pi} \int_{\theta \in \Omega_{\theta}} \left\{ \left( \langle \mu_{\theta}, \pi \rangle - \frac{\gamma}{2} \langle \pi' \Sigma_{\theta}, \pi \rangle \right) \right. \\ &\quad \left. - \max_{\pi_{\theta}} \left( \langle \mu_{\theta}, \pi_{\theta} \rangle - \frac{\gamma}{2} \langle \pi'_{\theta} \Sigma_{\theta}, \pi_{\theta} \rangle \right) \right\}^2 d\mathbb{P}_{\theta} \\ &= \arg \min_{\pi} \int_{\theta \in \Omega_{\theta}} \left\{ \langle \mu_{\theta}, \pi - \pi_{\theta}^* \rangle - \frac{\gamma}{2} \langle (\pi - \pi_{\theta}^*)' \Sigma_{\theta}, (\pi - \pi_{\theta}^*) \rangle \right\}^2 d\mathbb{P}_{\theta}, \end{aligned}$$

o que é similar à função-objetivo de HORST; ROON; WERKER (2001)

$$\begin{aligned} &\arg \min_{\bar{\theta}} \int_{\theta \in \Omega_{\theta}} L(\theta, \bar{\theta}) d\mathbb{P}_{\theta} \\ &= \arg \min_{\bar{\theta}} \int_{\theta \in \Omega_{\theta}} \left\{ \langle \mu_{\theta}, \pi_{\theta} - \pi_{\bar{\theta}} \rangle - \frac{\gamma}{2} \langle (\pi_{\theta} - \pi_{\bar{\theta}})' \Sigma_{\theta}, (\pi_{\theta} - \pi_{\bar{\theta}}) \rangle \right\} d\mathbb{P}_{\theta}. \end{aligned}$$

Observamos que a diferença é a variável de otimização que é, no primeiro caso, a composição da carteira  $\pi$ , e no segundo caso, o conjunto de parâmetros  $\bar{\theta}$ .

### 5.5.3 Minimização do desvio da escolha ótima

Uma outra forma de determinar o desvio de uma carteira especificada da carteira ótima com o parâmetro  $\theta$  é de considerar a distância entre os pesos. Nesse caso, definimos a função de penalidade como

$$B_{\delta}(\pi, \theta) := \delta \cdot \|\pi_{\theta} - \pi\|_2^2. \quad (5.4)$$

Mais uma vez, calculamos a carteira ótima diretamente da otimização dessa função

$$\tilde{\pi} = \arg \min_{\pi} \int_{\theta \in \Omega_{\theta}} \|\pi_{\theta} - \pi\|_2^2 d\mathbb{P}_{\theta}. \quad (5.5)$$

<sup>14</sup>Usamos a notação  $\pi_{\theta}^*$  para a carteira ótima do problema da maximização da utilidade esperada sob  $\mathbb{P}_{\theta}$ .

### 5.5.4 A característica da mínima sensibilidade

A sensibilidade mínima seria um critério interessante de ser escolhido quando o investidor não está contente com muitas mudanças na composição da carteira. Na abordagem de pesos iguais para todos os ativos, grandes pesos em alguns ativos são penalizados, pois a norma  $L_2$  aumenta nesse caso: quanto mais equilibrada a carteira, menor a penalidade. Assim, a carteira ótima está situada no meio das possíveis carteiras. Ela pode ser vista também como a que é insensível a mudanças nas distribuições dos retornos. Então, podemos formular a função-objetivo com penalidade como

$$\tilde{\pi} = \arg \max_{\pi \in \mathcal{A}} \int_{\theta \in \Omega_\theta} E_{\mathbb{P}|\theta} [U(\langle X, \pi \rangle)] - \|\pi\|_2^2 d\mathbb{P}_\theta. \quad (5.6)$$

Como mencionamos durante a formulação geral das abordagens de penalidade, a otimização é aplicada depois que tiramos a média, o que resulta em uma formulação diferente da maioria das propostas. Além disso, o uso da norma dos pesos da carteira é um caminho intuitivo de incluir a aversão aos pesos individuais altos na carteira, mesmo que não seja o método típico de medir sensibilidade. Pergunta-se também o porquê da nossa proposta refletir a idéia de uma estabilidade aumentada. Uma forma natural para penalizar variações grandes na utilidade esperada por razões de pequenas mudanças nos parâmetros pode se basear no uso do gradiente dos pesos em relação aos parâmetros, por exemplo

$$B_\delta(\pi, \theta) := \delta \cdot \|\nabla_\theta E_{\mathbb{P}|\theta} [U(\langle X, \pi_\theta(\rho) \rangle)]\|_2^2.$$

No caso especial de uma média com distribuição normal,  $\theta \equiv \mu, X \sim \mathcal{N}(\mu, \Sigma)$ , podemos escrever

$$B_\delta(\pi, \theta) = 2\delta (\Sigma^{-1})^2 \cdot E_{\mathbb{P}_\theta} [E_{\mathbb{P}|\theta} [U(\langle X, \pi_\theta(\rho) \rangle)]^2],$$

que resulta numa função-objetivo similar a da penalidade pela variância

$$f = \max_{\pi \in \mathcal{A}(X)} \int_{\Omega_\theta} E_{\mathbb{P}|\theta} [U(\langle X, \pi_\theta(\rho) \rangle)] d\mathbb{P}_\theta - 2\delta (\Sigma^{-1})^2 \int_{\Omega_\theta} E_{\mathbb{P}|\theta} [U(\langle X, \pi_\theta(\rho) \rangle)]^2 d\mathbb{P}_\theta.$$

Penalizando por grandes variações na composição da carteira, teremos

$$B_\delta(\pi, \theta) := \delta \cdot \|\nabla_\theta \pi(\theta)\|_2^2.$$



O problema em medir a sensibilidade dessa forma é que estamos confrontados com composições que podem mudar abruptamente nos pontos em que um parâmetro começa a dominar o outro, pois o máximo não é função contínua. Para entender esse argumento, visualizamos o comportamento no problema de substituição: pressupondo  $\sigma_1 = \sigma_2$ , a carteira ótima é dada por

$$\tilde{\pi} = \arg \min_{\pi} \int_{\Omega_{\theta}} \|\nabla_{\theta} \pi(\theta)\|_2^2 d\mathbb{P}_{\theta}.$$

Com  $\mu = (\mu_1, \mu_2)$  fixo, temos  $\pi_1 = \mathbf{1}_{\{\mu_1 \geq \mu_2\}} = 1 - \pi_2$ , e o gradiente satisfaz

$$\nabla_{\theta} \pi(\theta) = 0, \mathbb{P}_{\theta}\text{-q.c.}$$

o que significa que o gradiente, mesmo nesse caso simples, não é uma boa medida da sensibilidade e deve ser substituído por uma medida de distância (com normalização, se for necessário)  $\|\nabla_{\theta} \pi(\theta)\| \rightsquigarrow \|\pi_{\theta} - \pi\|$ , com  $\pi_{\theta}$  sendo uma carteira fixa, no nosso caso, a ótima com respeito ao parâmetro  $\theta$ . Observamos que não especificamos explicitamente a norma, pois a escolha dela depende do investidor. Por exemplo, pode ser usada a norma máxima. Isso significa que o uso de

$$\tilde{\pi} = \arg \min_{\pi} \int_{\Omega_{\theta}} \|\pi_{\theta} - \pi\|_{\infty} d\mathbb{P}_{\theta}$$

cria uma carteira  $\tilde{\pi}$  que tem mínima distância máxima entre quaisquer dois pesos na carteira e segue assim a idéia de uma composição com pesos iguais, eliminando grandes diferenças de pesos na carteira. A norma  $L_2$  foi aplicada na nossa segunda proposta da seção anterior. Quando a sensibilidade não é analisada em termos da composição da carteira, mas em termos da utilidade esperada, teremos o caso geral da nossa proposta de minimizar o desvio da melhor escolha em unidades de utilidade esperada.

Essas observações sobre as possibilidades de descrever a sensibilidade da composição da carteira classificam as nossas propostas como tendo uma característica de estabilidade, pois elas minimizam a sensibilidade.

## 5.6 O problema de determinar valores esperados equivalentes no espaço dos valores médios

Quando a fronteira eficiente é visualizada no  $\mathbb{R}^2$ , plotando as carteiras com incerteza (desvio-padrão dos retornos<sup>15</sup>) crescentes contra os retornos (valor esperado dos retornos) das carteiras, na abordagem clássica de MARKOWITZ (1952), a curva tem uma forma côncava. Tentando plotar as médias contra os desvio-padrões de qualquer carteira usando os valores de entrada no otimizador de média-variância de MARKOWITZ (1952), não é possível ficar acima da fronteira eficiente de MARKOWITZ (1952). Pior ainda, plotado no gráfico de médias e variâncias, a curva de *reamostragem* pode pressupor uma forma não-côncava. Mesmo que esse fato surpreenda alguns autores no primeiro momento (vide SCHERER, 2002), a explicação é simples: a curva *reamostrada* foi criada a partir de carteiras cujas composições são as médias das composições de várias otimizações. Como a distribuição dos pesos não precisa obedecer às condições de Lindeberg para garantir a validade do teorema central do limite<sup>16</sup> (vide BILLINGSLEY, 1995), as médias dos pesos não precisam ser os pesos da otimização determinística (i.e., os pesos resultando das médias e variâncias esperadas). Mais do que isso, a figura 5.2 expressa graficamente o fato de que a média dos pesos não precisa ser uma boa variável descritiva da distribuição dos pesos.

Fica aberto para futuras análises como se comportam as outras propostas discutidas na seção anterior.

**Observação 5.1** *Observou-se que na prática os resultados das abordagens bayesianas (ou, nesse caso, da reamostragem) têm melhor desempenho fora da amostra que os da otimização determinística, vide MARKOWITZ; USMEN (2003) para uma análise empírica no mercado financeiro.*

Mesmo que exista uma certa incompatibilidade entre a curva determinística de MARKOWITZ (1952) e as curvas estocásticas, pode se usar uma heurística que ajuda a analisar a sensibilidade das carteiras de MARKOWITZ (1952) em relação aos erros da estimação.

<sup>15</sup>Para ser mais exato, deve ser usado o termo *taxa de retorno composto continuamente*, mas, por razões de legibilidade, usamos simplesmente o termo *retornos*.

<sup>16</sup>Ao menos não existe ainda resultado formal que prove isso.

A idéia básica é determinar para cada carteira da fronteira *reamostrada* sob a hipótese de falta de restrições na otimização qual seria o vetor de retornos no mundo MARKOWITZ (1952) que teria criado a composição. Essa “inversão” é possível no caso de uma otimização na qual as únicas restrições são limites individuais de alocação e desfruta do fato que uma composição ótima pode ser determinada de forma analítica (vide LUENBERGER, 1998). Assim, teremos um vetor de médias dos retornos que pode ser comparado ao vetor de médias originais determinísticas.

Na seção 5.7.2, analisamos quais seriam os retornos implicitamente necessários para que possam ser criadas as alocações ao longo das fronteiras eficientes. Observamos que as diferenças dos retornos determinísticos implícitos ao longo das fronteiras eficientes são bastante pequenas de um ponto de vista prático, i.e., numa aplicação real é pouco provável que um agente financeiro distinga entre os retornos esperados das diversas abordagens. Esse fato confirma que os resultados das abordagens, por serem derivados a partir de otimizações, são altamente sensíveis às variações nos dados de entrada. Portanto, o método usado nas aplicações reais deveria levar esse ponto crítico em conta e tentar aumentar a estabilidade. Este objetivo é atingido pelas abordagens propostas.

## 5.7 O problema de erro de estimação no mercado financeiro brasileiro

### 5.7.1 Os dados do mercado financeiro

Analisamos os desempenhos de algumas das abordagens no mercado financeiro brasileiro. Os ativos com as suas características populacionais são dados na tabel 5.1. Pressupomos como distribuições marginais a distribuição normal e usamos a volatilidade (desvio-padrão) como medida de incerteza.<sup>17</sup>

---

<sup>17</sup>Os dados são do dia 05.11.2002. A média e a variância são estimadas de forma clássica de dados com periodicidade mensal. Os ativos são vértices artificiais criados a partir de títulos públicos e contratos da Bolsa de Mercadorias e Futuros de São Paulo (*BM&F*). As representações são: SELIC=taxa Selic; RF curto=Taxa de juros pré-fixada de curto prazo (3 meses); RF longo=Taxa de juros pré-fixada de longo prazo (6 meses); IBOVESPA=Índice da Bolsa de Valores de São Paulo; T. Camb.=Título cambial de 6 meses (DDI); IGP-M=Título lastreado na variação do índice geral dos preços do mercado (12 meses); deságio da LFT=Letra Financeira Tesouro Nacional (12 meses).

Deixando a pressuposição de médias fixas e determinísticas e pressupondo que elas seguem uma distribuição normal com as médias populacionais como valores esperados e com a matriz de covariância dos retornos como matriz de covariância das médias, satisfaremos as condições do arcabouço teórico de HORST; ROON; WERKER (2001). O conjunto de fronteiras *reamostradas*, como propostas por MICHAUD (1998), é visualizado na figura 5.3.

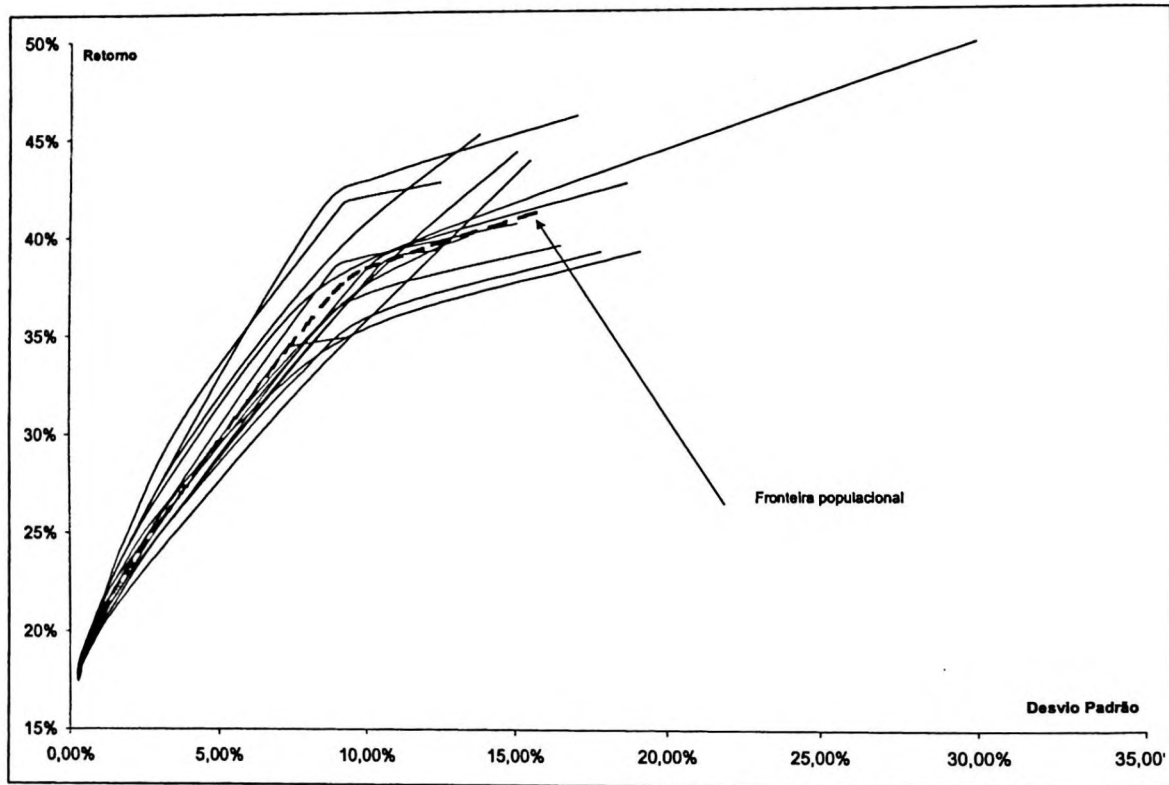


Figura 5.3: Conjunto de fronteiras reamostradas.

Na nossa simulação, com os retornos seguindo uma distribuição multinomial, simulamos também a estrutura de dependência dada pela matriz de correlações. Assim, a matriz de covariância segue uma distribuição de Wishart, que é definida como a soma de produtos independentes de vetores de normais (vide JOHNSON; WICHERN, 1992).

### 5.7.2 Os efeitos de diferentes abordagens

Resultado da otimização usando o método de MARKOWITZ (1952)

Aplicando a otimização de média-variância tradicional ao conjunto de informações originais, chegamos aos resultados da tabela 5.2.

O problema com essa abordagem fica claro quando observamos os pesos ótimos após uma mudança gradual no valor esperado do retorno da taxa Selic e do título LFT. Nas aplicações reais, uma mudança no retorno esperado de 0,6%, como foi introduzida no nosso caso, seria, provavelmente, classificada pela maioria dos agentes financeiros como ruído. Infelizmente, a composição da carteira nesse caso sofre mudanças altas, ou seja, a composição é altamente sensível às mudanças nos valores esperados. O problema é que os custos de transação causados pelos ajustes na carteira compensariam facilmente os retornos adicionais. De qualquer forma, usaremos essa abordagem como modelo de referência para comparar o desempenho das outras propostas. Além do retorno criado pela carteira escolhida, é a estabilidade da composição que nos preocupa e que será analisada.

### **Restrição dos pesos dos ativos**

Uma das alternativas mais simples para evitar alocações extremas é a de restringir pesos altos, o que é feito por limites de alavancagem ou de venda à descoberto. Obviamente, altos pesos em ativos individuais são eliminados, mas observa-se nas aplicações que os multiplicadores de Lagrange das restrições são, muitas vezes, positivos.

### **Distorção bayesiana nos parâmetros da distribuição: a proposta de BLACK; LITTELMAN (1992)**

Quando queremos usar a idéia bayesiana de atualizar uma distribuição normal prévia, parametrizada pela média e variância, precisamos especificar uma distribuição alternativa. BLACK; LITTELMAN (1992) pressupõem que existem especialistas na firma que criaram esses parâmetros alternativos e não que foram estimados a partir das observações usando, por exemplo, um estimador de James-Stein. Além disso, na prática, é pouco provável obter uma distribuição prévia explícita. Por isso, usaremos os dados originais como parâmetros prévios e interpretamos os parâmetros alternativos como os parâmetros da distribuição de atualização. No cálculo da atualização usamos as variâncias originais como incerteza sobre as médias alternativas, pressupondo independência entre as estimações das alternativas. Assim, criamos uma distribuição atualizada que fornece os dados alternativos na tabela 5.4.

Observa-se que uma mudança no retorno da taxa Selic e do título LFT causou uma mudança em **todas** as médias da distribuição atualizada. As mudanças nos pesos são

menores que no primeiro exemplo.

### Simulação de diferentes fronteiras eficientes

Para uma análise do método de MICHAUD (1998) de criar diferentes fronteiras eficientes a partir de conjuntos de parâmetros alternativos, analisamos as fronteiras médias criadas por uma variação do coeficiente de aversão ao risco relativo de Arrow-Pratt. Tabela 5.5 mostra a composição ao longo da fronteira usando os dados originais, enquanto a tabela 5.6 mostra a fronteira média com os dados alternativos.<sup>18</sup>

Observamos que as composições das carteiras nos pontos extremos da fronteira (mínima variância e máxima variância) são relativamente estáveis. Porém, no meio da fronteira, constatamos uma forte realocação dos pesos entre os ativos com uma mudança máxima de 19% nos pesos do ativo Selic. A última linha nas tabelas mostra a norma  $L_2$  da composição, que é um indicador da concentração das alocações. O problema das mudanças altas no meio da curva é causado pela relativa alta concentração dos pesos nos ativos Selic e pré curto. Usamos a restrição de não-alavancagem.

Mesmo que não seja o objetivo principal da abordagem, observamos uma estabilidade maior nas composições. Este fenômeno é causado pela escolha final da carteira como uma carteira média das alternativas simuladas.

### Abordagens de penalidade

Calculamos as carteiras ótimas com as funções de penalidade (5.3) e (5.5). As seguintes tabelas mostram as composições resultantes e as normas  $L_2$  para os casos de dados iniciais e alternativos. Tabelas 5.7 e 5.8 mostram as fronteiras usando a penalidade por (5.3), e as tabelas 5.9 e 5.10 mostram os resultados com a penalidade por grandes desvios nas composições.

Comparando os resultados das abordagens de penalidade propostas com os resultados da *reamostragem*, observamos que as duas abordagens propostas criam carteiras com composições mais estáveis. Este efeito não pode somente ser observado na norma  $L_2$  das diferenças das composições entre as alternativas, mas também no máximo da mudança dos pesos (norma  $L_\infty$ ). Observamos que a soma da norma  $L_2$  ao longo das fronteiras das

---

<sup>18</sup>Os números na primeira linha são os valores do coeficiente de aversão ao risco ao longo da fronteira. Foi usada a condição de não-alavancagem.

diferenças das mudanças das composições no caso da abordagem de MICHAUD (1998) é quase três vezes maior que na abordagem de penalidade por desvios na utilidade esperada.

#### Diferenças entre retornos determinísticos ao longo da curva eficiente

Calculamos no exemplo do mercado financeiro quais seriam os retornos implícitos para que possam surgir as alocações propostas por *reamostragem*. O cálculo foi feito para as alocações criadas com a abordagem de MARKOWITZ (1952) (tabela 5.11), pelo o método bayesiano de BLACK; LITTERMAN (1992) (tabela 5.12) e a partir da *reamostragem* (tabela 5.13).

Ativo	Médias	Desvio-padrão	1	2	3	4	5	6	7
SELIC	17,6%	0,3%	1						
RF curto	19,4%	1,2%	0,15	1					
RF longo	19,1%	2,9%	0,17	0,96	1				
IBOVESPA	24,2%	35,3%	-0,03	0,60	0,67	1			
T. Camb.	41,3%	16,0%	-0,15	-0,49	-0,60	-0,68	1		
IGP-M.	23,5%	3,9%	0,02	-0,08	-0,12	-0,07	0,37	1	
LFT	16,5%	0,6%	0,29	0,31	0,43	0,21	-0,23	-0,13	1

Tabela 5.1: Ativos usados na otimização.



Ativo	Dados originais		Markowitz	
	Médias	Pesos	Médias	Pesos
SELIC	17,6%	40%	17,0%	-134%
RF curto	19,4%	182%	19,4%	284%
RF longo	19,1%	-59%	19,1%	-110%
IBOVESPA	24,2%	1%	24,2%	1%
T. Camb.	41,3%	3%	41,3%	1%
IGP-M	23,5%	3%	23,5%	7%
LFT	16,5%	-71%	17,1%	51%

Tabela 5.2: Carteira segundo a otimização média-variância simples.

Ativo	Dados originais		Dados alternativos	
	Médias	Pesos	Médias	Pesos
SELIC	17,6%	33%	17,0%	16%
RF curto	19,4%	57%	19,4%	74%
RF longo	19,1%	0%	19,1%	0%
IBOVESPA	24,2%	1%	24,2%	0%
T. Camb.	41,3%	5%	41,3%	5%
IGP-M	23,5%	4%	23,5%	5%
LFT	16,5%	0%	17,1%	0%

Tabela 5.3: Pesos dos ativos na carteira com restrições de aplicação.

Ativo	Dados originais		Dados Alternativos	
	Médias	Pesos	Médias	Pesos
SELIC	17,6%	40%	<i>17,3%</i>	28%
RF curto	19,4%	182%	<i>19,4%</i>	62%
RF longo	19,1%	-59%	<i>19,2%</i>	0%
IBOVESPA	24,2%	1%	<i>27,3%</i>	1%
T. Camb.	41,3%	3%	<i>42,1%</i>	6%
IGP-M	23,5%	3%	<i>23,2%</i>	3%
LFT	16,5%	-71%	<i>16,7%</i>	0%

Tabela 5.4: Pesos dos ativos usando a abordagem de Black-Litterman com restrições de alavancagem.

Ativo	50000	14881,8	4429,3	1318,3	392,4	172,5	63,7	19	5,6	1,7
SELIC	82%	82%	83%	74%	30%	2%	0%	0%	0%	0%
RF curto	4%	5%	9%	21%	59%	76%	29%	0%	0%	0%
RF longo	0%	0%	0%	0%	0%	2%	22%	3%	0%	0%
IBOVESPA	0%	0%	0%	0%	1%	2%	7%	21%	16%	11%
T. Camb.	1%	1%	1%	2%	5%	11%	27%	70%	84%	89%
IGP-M	0%	0%	1%	1%	5%	8%	15%	5%	0%	0%
LFT	13%	11%	6%	0%	0%	0%	0%	0%	0%	0%
$L_2$ -Norma	0,687	0,694	0,707	0,601	0,444	0,592	0,234	0,541	0,735	0,804

Tabela 5.5: Fronteira reamostrada com dados originais.

<b>Ativo</b>	50000	14881,8	4429,3	1318,3	392,4	172,5	63,7	19	5,6	1,7
SELIC	82%	81%	77%	61%	11%	0%	0%	0%	0%	0%
RF curto	4%	5%	9%	25%	70%	77%	28%	0%	0%	0%
RF longo	0%	0%	0%	0%	0%	2%	21%	2%	0%	0%
IBOVESPA	0%	0%	0%	0%	1%	2%	8%	22%	16%	13%
T. Camb.	1%	1%	1%	2%	6%	11%	29%	71%	84%	87%
IGP-M	0%	1%	1%	2%	5%	8%	13%	5%	0%	0%
LFT	13%	13%	12%	10%	8%	0%	0%	0%	0%	0%
<i>L<sub>2</sub>-Norma</i>	0,696	0,680	0,619	0,442	0,510	0,606	0,234	0,551	0,725	0,779

Tabela 5.6: Fronteira reamostrada com dados alternativos.

Ativo	50000	14881,8	4429,3	1318,3	392,4	172,5	63,7	19	5,6	1,7
SELIC	82%	86%	74%	58%	23%	19%	15%	0%	0%	0%
RF curto	11%	6%	0%	7%	25%	19%	15%	6%	0%	0%
RF longo	0%	0%	24%	0%	29%	18%	15%	5%	0%	0%
IBOVESPA	0%	0%	1%	3%	1%	3%	5%	16%	8%	0%
T. Camb.	0%	0%	0%	0%	0%	8%	20%	57%	78%	99%
IGP-M	0%	0%	0%	0%	0%	15%	16%	15%	14%	1%
LFT	6%	8%	0%	33%	22%	18%	15%	0%	0%	0%
$L_2$ -Norma	0,688	0,742	0,610	0,444	0,248	0,167	0,156	0,383	0,640	0,978

Tabela 5.7: Fronteira eficiente usando (5.3) com dados originais.

<b>Ativo</b>	50000	14881,8	4429,3	1318,3	392,4	172,5	63,7	19	5,6	1,7
SELIC	89%	82%	75%	53%	22%	18%	15%	0%	0%	0%
RF curto	5%	7%	0%	7%	25%	19%	15%	7%	0%	0%
RF longo	0%	0%	24%	0%	29%	18%	15%	7%	0%	0%
Equity	0%	0%	2%	3%	1%	3%	5%	12%	8%	0%
T. Camb.	0%	0%	0%	0%	0%	8%	20%	58%	79%	99%
IGP-M	0%	0%	0%	0%	0%	15%	16%	16%	14%	1%
LFT	6%	12%	0%	37%	22%	18%	15%	0%	0%	0%
$L_2$ -Norm	0,798	0,687	0,613	0,424	0,247	0,167	0,156	0,387	0,643	0,978

Tabela 5.8: Fronteira eficiente usando (5.3) com dados alternativos.

Ativo	50000	14881,8	4429,3	1318,3	392,4	172,5	63,7	19	5,6	1,7
SELIC	79%	79%	79%	67%	26%	5%	0%	0%	0%	0%
RF curto	5%	6%	10%	25%	55%	56%	22%	1%	0%	0%
RF longo	0%	0%	0%	0%	3%	10%	19%	4%	0%	0%
IBOVESPA	0%	0%	0%	0%	1%	3%	9%	20%	16%	9%
T. Camb.	1%	1%	1%	3%	7%	14%	32%	68%	84%	91%
IGP-M	1%	1%	1%	3%	7%	13%	18%	7%	0%	0%
LFT	14%	12%	8%	2%	0%	0%	0%	0%	0%	0%
$L_2$ -Norma	0,641	0,645	0,640	0,509	0,385	0,357	0,225	0,505	0,730	0,831

Tabela 5.9: Fronteira eficiente usando (5.5) com dados originais.

Ativo	50000	14881,8	4429,3	1318,3	392,4	172,5	63,7	19	5,6	1,7
SELLIC	78%	77%	71%	50%	12%	1%	0%	0%	0%	0%
RF curto	5%	6%	11%	29%	62%	60%	23%	1%	0%	0%
RF longo	0%	0%	0%	0%	2%	9%	19%	4%	0%	0%
IBOVESPA	0%	0%	0%	0%	1%	3%	9%	20%	16%	10%
T. Camb.	1%	1%	1%	3%	7%	13%	31%	67%	84%	90%
IGP-M	1%	1%	1%	3%	8%	12%	18%	7%	0%	0%
LFT	15%	15%	16%	15%	8%	1%	0%	0%	0%	0%
$L_2$ -Norma	0,633	0,612	0,538	0,357	0,421	0,398	0,225	0,499	0,736	0,826

Tabela 5.10: Fronteira eficiente usando (5.5) com dados alternativos.



Ativo	50000	14881,8	4429,3	1318,3	392,4	172,5	63,7	19	5,6	1,7
Taxa SELIC	-5,66	-1,76	-0,60	-0,26	-0,16	-0,13	-0,11	-0,08	-0,02	-0,01
RF curto	46,64	14,52	4,95	2,15	1,32	1,05	0,87	0,68	0,18	0,05
RF longo	48,56	15,11	5,16	2,24	1,37	1,10	0,91	0,71	0,19	0,05
IBOVESPA	72,76	22,64	7,73	3,35	2,06	1,64	1,36	1,06	0,28	0,07
T. Camb.	-24,17	-7,52	-2,57	-1,11	-0,68	-0,55	-0,45	-0,35	-0,09	-0,02
IGP-M	18,88	5,87	2,00	0,87	0,53	0,43	0,35	0,27	0,07	0,02
LFT	14,80	4,61	1,57	0,68	0,42	0,33	0,28	0,22	0,06	0,01

Tabela 5.11: Retornos determinísticos implícitos da fronteira eficiente de MARKOWITZ (1952) com restrições de não-alavancagem

Ativo	50000	14881,8	4429,3	1318,3	392,4	172,5	63,7	19	5,6	1,7
Taxa SELIC	-5,67	-1,78	-0,62	-0,27	-0,17	-0,14	-0,12	-0,08	-0,02	-0,01
RF curto	46,76	14,63	5,07	2,24	1,41	1,13	0,97	0,68	0,19	0,05
RF longo	48,68	15,23	5,28	2,33	1,47	1,17	1,01	0,71	0,19	0,05
IBOVESPA	72,94	22,82	7,91	3,50	2,20	1,76	1,51	1,06	0,29	0,07
T. Camb.	-24,23	-7,58	-2,63	-1,16	-0,73	-0,58	-0,50	-0,35	-0,10	-0,02
IGP-M	18,92	5,92	2,05	0,91	0,57	0,46	0,39	0,28	0,08	0,02
LFT	14,84	4,64	1,61	0,71	0,45	0,36	0,31	0,22	0,06	0,01

Tabela 5.12: Retornos determinísticos implícitos da fronteira eficiente de BLACK; LITTEPMAN (1992) com restrições de não-alavancagem

Ativo	50000	14881,8	4429,3	1318,3	392,4	172,5	63,7	19	5,6	1,7
Taxa SELIC	-5,89	-1,83	-0,63	-0,28	-0,17	-0,14	-0,12	-0,08	-0,02	-0,01
RF curto	48,55	15,08	5,17	2,28	1,44	1,13	0,95	0,63	0,18	0,05
RF longo	50,54	15,70	5,38	2,37	1,50	1,17	0,99	0,66	0,19	0,06
IBOVESPA	75,73	23,52	8,07	3,56	2,24	1,76	1,48	0,99	0,29	0,08
T. Camb.	-25,16	-7,81	-2,68	-1,18	-0,74	-0,58	-0,49	-0,33	-0,10	-0,03
IGP-M	19,65	6,10	2,09	0,92	0,58	0,46	0,39	0,26	0,07	0,02
LFT	15,40	4,79	1,64	0,72	0,46	0,36	0,30	0,20	0,06	0,02

Tabela 5.13: Retornos determinísticos implícitos da fronteira eficiente de reamostragem com restrições de não-alavancagem.

## Capítulo 6

# Alocação ótima com erros de modelo

### 6.1 Introdução: o problema de comparação de modelos

O erro de modelo expressa-se pela incerteza sobre o tipo de distribuição dos ativos do mercado, contrariamente ao erro de estimação que pressupõe uma certa forma de distribuição, mas tem incerteza sobre os seus parâmetros. Um possível efeito da diferença é que, por exemplo, valores extremos sob hipótese de distribuição normal teriam impacto no aumento da variância. No caso de incerteza da distribuição, será escolhida uma distribuição que incorpora as observações extremas na cauda.

Quando confrontamos erros na estimação, vimos que podem ser usadas funções de perda que atribuem valores a possíveis mudanças da utilidade. Mesmo que o critério de decisão seja subjetivo, ao menos existe um sistema de referência em relação ao qual se medem os desvios. Quando não existe certeza sobre o modelo, perde-se também o sistema de referência, pois não sabemos qual dos modelos é correto. Infelizmente, nem os ajustes bayesianos são robustos aos erros de modelo, pois dependem das hipóteses da distribuição inicial. Porém, existe a possibilidade de usar estatística robusta que incorpora pequenos desvios locais de uma distribuição de referência.<sup>1</sup> Para podermos usar esta abordagem, temos que ter uma hipótese sobre a distribuição referencial dos desvios do modelo de referência. Os trabalhos de PÁSTOR; STAMBAUGH (1999), PÁSTOR; STAMBAUGH

---

<sup>1</sup>Para um exemplo de aplicação dessa abordagem, vide CAVADINI; SBUELZ; TROJANI (2001).

(2000) e AVRAMOV (2002) usam essa abordagem bayesiana com a hipótese de que os erros entre o modelo de referência e as alternativas seguem uma distribuição normal. KACPERCZYK (2003), por sua vez, trata os erros de forma não-paramétrica.

Resumindo, na literatura sobre incerteza de modelo dominam as abordagens bayesianas e a maioria segue o mesmo espírito que no âmbito da incerteza sobre os parâmetros: mede-se a distância entre um modelo *a priori* e as alternativas.

O objetivo dessa seção é introduzir três conceitos originais que contribuem para uma melhor decisão do investidor quando confrontado com erros de modelo. Nenhuma dessas propostas tenta medir distâncias entre modelos alternativos no estrito senso. Porém, propomos fórmulas de precificação que são sensíveis às diferentes características das distribuições, permitindo uma comparação e ordenação entre elas. Obviamente, essas propostas servem também para comparar especificações diferentes de parâmetros de uma mesma distribuição.

Primeiramente, com a medida *entrópica de dispersão relativa* (ERD), propomos uma nova forma de comparar distribuições. O ERD permite ordenar de forma total todas as variáveis aleatórias, mesmo com domínios diferentes. A possibilidade de ordenar distribuições ou, equivalentemente, carteiras de ativos financeiros, é importante na escolha entre modelos alternativos.

Depois, propomos um novo método de avaliação, chamado de *precificação consistente* por causa da sua forma especial de determinar o valor justo de um ativo. Ele se baseia na nossa interpretação de risco como estabelecido na seção 3.1, e na idéia que um agente econômico racional mantém o mesmo método de avaliação para todas as variáveis aleatórias.

Por último, introduzimos um conceito de precificação que usa tanto uma função de utilidade dada quanto uma distribuição estabelecida. Mesmo parecendo que não seja apta para analisar erros de modelo, a abordagem do *equivalente monetário determinístico* (EMD) é útil, pois ela ajuda a entender como um investidor “racional” deve se comportar. Essa abordagem normativa serve para poder restringir funções de utilidade e distribuições de retornos de forma que possamos excluir aquelas que, segundo o EMD, não fazem sentido econômico.

## 6.2 A medida entrópica de dispersão relativa

### 6.2.1 Introdução e definição da nova medida

O objetivo dessa seção é propôr uma medida de dispersão que combina variabilidade e entropia de uma forma que permite comparar distribuições em domínios quaisquer, i.e., contínuo, discreto, finito e infinito. Pretendemos derivar uma medida que resulte numa ordenação total.<sup>2</sup>

A razão de ter incluída esta medida no capítulo de incerteza sobre as distribuições é que ela, além de focar na variância, considera a distribuição inteira pelo uso da entropia. Assim, pode ser tomada uma decisão sobre qual distribuição alternativa escolher. A variância, além de não satisfazer alguns critérios que podem ser considerados importantes,<sup>3</sup> cria incoerências na ordenação de distribuições porque é uma medida absoluta que não considera a localização da distribuição. Para resolver esse problema, pode-se pensar em usar o coeficiente de variação de Pearson para uma variável aleatória  $X$  com  $E[X] \neq 0$ :

$$\rho_{\text{Pearson}}(X) := \frac{\sigma(X)}{E[X]}.$$

Mesmo que incorpore a localização, esse coeficiente tem a desvantagem de não ser invariante sob transformação linear, i.e., para qualquer  $a, b \in \mathbb{R}, X \in L^2(\Omega, \mathcal{F}, \mathbb{R})$  e  $Y := aX + b$  vale que<sup>4</sup>

$$\begin{aligned} \rho_{\text{Pearson}}(Y) &= \rho_{\text{Pearson}}(aX + b) \\ &= \frac{a \cdot \sigma(X)}{a \cdot E[X] + b} \\ &\neq \rho_{\text{Pearson}}(X). \end{aligned}$$

Uma alternativa que é invariante sob transformação linear pode ser derivada em analogia à proposta de EISENHAUER (1993), normalizando-se o desvio-padrão com a *amplitude* da variável  $X$ ,  $\text{rg}(X)$ :<sup>5</sup>

$$\rho_{\text{EH}}(X) := \frac{\sigma(X)}{\text{rg}(X)}.$$

<sup>2</sup>Nessa seção, pressupomos a existência de, ao menos, o segundo momento das distribuições.

<sup>3</sup>Como de não ser uma medida de risco coerente, vide seção anterior.

<sup>4</sup>Exclui-se o caso de  $a = b = 0$ .

<sup>5</sup>Observa-se que a proposta original de EISENHAUER (1993) é feita para criar uma estatística a partir de dados amostrais, enquanto nós propomos usar uma função aleatória geral.

Essa medida é indiferente à localização de  $X$ , porque somente o tamanho do amplitude é de interesse. A normalização permite comparar diferentes distribuições, pois a medida  $\rho_{EH}$  tem limites finitos se a distribuição for finita. De fato, estimando  $\sigma^2(X)$  por<sup>6</sup>

$$\hat{\sigma}^2(X) := \frac{1}{n-1} \sum_{i=1}^n \left( X_i - \frac{1}{n} \sum_{i=1}^n X_i \right)^2$$

e  $\text{rg}(X)$  com

$$\widehat{\text{rg}}(X) := \max \left( \{X_i\}_{i=1,2,\dots,n} \right) - \min \left( \{X_i\}_{i=1,2,\dots,n} \right),$$

temos que

$$\frac{1}{\sqrt{2(n-1)}} \leq \frac{\hat{\sigma}(X)}{\widehat{\text{rg}}(X)} \leq \frac{1}{2} \sqrt{\frac{n}{n-1}},$$

e com  $n \rightarrow \infty$  temos  $0 \leq \frac{2\hat{\sigma}(X)}{\widehat{\text{rg}}(X)} \leq 1$ .

Como a variância calcula a dispersão de  $X$  em torno da sua esperança, perde-se, nas duas abordagens citadas, informação sobre a distribuição. Isso acontece especialmente quando outros momentos centrados de ordem superior têm importância, a distribuição é multi-modal etc.

Baseando-se na variância como a função fundamental para calcular a dispersão, procuramos uma outra forma de obter uma medida que

- seja invariante sob transformação linear, e
- considera toda a informação contida na distribuição.

Para incorporar a forma geral da distribuição, usaremos o volume do (conjunto de) suporte da distribuição, um indicador da uniformidade, que nos permite levar em conta a concentração, em concordância com a (meta-) definição 3.5. Identificamos com

$$\xi(X) := \{A \in \mathcal{F} \mid \mathbb{P}(X \in A) = 1\}$$

o suporte da variável aleatória  $X$ . COVER; THOMAS (1991) interpretaram o volume do suporte de  $X$  da seguinte forma:<sup>7</sup>

<sup>6</sup>Identificamos com  $\{X_i\}_{i=1,2,\dots,n}$  as realizações da variável  $X$ .

<sup>7</sup>Essa noção do volume não deve ser confundida com a noção clássica do volume de um conjunto  $A^{\text{cont}} \in \mathcal{B}(\mathbb{R})$  que é definido como

$$\text{vol}(A^{\text{cont}}) := \int_{A^{\text{cont}}} dx.$$

para domínio contínuos e com  $\text{vol}(A^{\text{disc}}) := |A^{\text{disc}}|$  no caso discreto.

**Definição 6.1 (Volume do suporte)** *O volume do suporte de  $X$  é o menor conjunto possível do suporte que contém a maior parte da distribuição:*

$$\text{vol}(\xi(X)) := \exp(H(\mathbb{P}_X)),$$

sendo que

$$H(\mathbb{Q}) := \int_{\Omega} \log \mathbb{Q} d\mathbb{Q}$$

é a entropia de uma medida  $\mathbb{Q}$ .

Essa expressão está conforme a seguinte intuição: uma distribuição contendo muita incerteza e que, portanto, tem uma entropia alta, apresenta um volume grande. Combinando a medida de dispersão  $\sigma(X)$  com a menor região da maior parte da distribuição  $\text{vol}(\xi(X))$ , temos que:

**Definição 6.2 (Dispersão entrópica relativa)** *Para  $X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ ,  $p \geq 2$ , chame-se*

$$\text{ERD}(X) := \frac{\sigma(X)}{\text{vol}(\xi(X))} \quad (6.1)$$

a dispersão entrópica relativa (*Entropic Relative Dispersion*).

Como o volume leva em conta, via entropia, a distribuição toda, temos que verificar somente a invariância sob transformação linear.

**Lema 6.1** *Seja  $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$  e contínuo e  $a, b \in \mathbb{R}$ ,  $Y := aX + b$ , então vale que*

$$\text{ERD}(Y) = \text{ERD}(X).$$

**Prova.** Seguindo REZA (1961, p. 274), para  $X$  contínuo vale que  $H(Y) = H(X) + \ln |a|$ .

Por isso

$$\begin{aligned} \text{ERD}(Y) &= \text{ERD}(aX + b) \\ &= \frac{\sigma(aX + b)}{\exp(H(aX + b))} \\ &= \frac{a\sigma(X)}{\exp(H(X) + \ln |a|)} \\ &= \frac{\sigma(X)}{\exp(H(X))} \\ &= \text{ERD}(X). \end{aligned}$$



■

Com a introdução da seguinte definição, usaremos a medida ERD para ordenar variáveis aleatórias.

**Definição 6.3 (Ordem estocástica com respeito a ERD)** Para quaisquer  $X_1, X_2 \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ ,  $p \geq 2$ , dizemos que  $X_2$  domina  $X_1$  com respeito a ERD, ou seja,  $X_1 \stackrel{\text{ERD}}{>} X_2$  ou  $\mathbb{P}_{X_1} \stackrel{\text{ERD}}{>} \mathbb{P}_{X_2}$ , se  $\text{ERD}(X_1) < \text{ERD}(X_2)$ .

Uma vantagem de ERD é que esta permite comparar variáveis aleatórias discretas  $X_d$  com contínuas  $X_c$ , desde que tenham o mesmo espaço de imagem (campo numérico):

$$X_d : (\Omega_d, \mathcal{F}_d, \mathbb{P}_d) \rightarrow E,$$

$$X_c : (\Omega_c, \mathcal{F}_c, \mathbb{P}_c) \rightarrow E.$$

**Exemplo 6.1 (Comparação de uniforme discreta com uniforme contínua)** Sejam  $a, b \in \mathbb{R}$  e  $X_c \sim U([a; b])$ , com  $f_{a,b}(x) = \frac{1}{b-a}$ ,  $x \in \mathbb{R}$ . Então,

$$H(\mathbb{P}_{X_c}) = \log(b-a),$$

$$\text{vol}(\xi(X_c)) = b-a, \text{ e}$$

$$\sigma^2(X_c) = \frac{(b-a)^2}{12}.$$

Por outro lado, para uma variável discreta com distribuição uniforme  $X_d \sim U(n)$ ,  $n \in \mathbb{N}$ , temos

$$H(\mathbb{P}_{X_d}) = \log(n),$$

$$\text{vol}(\xi(X_d)) = n, \text{ e}$$

$$\sigma^2(X_d) = \frac{n^2-1}{12}.$$

Segue então que

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{ERD}(X_d) &= \lim_{n \rightarrow \infty} \frac{\sigma(X_d)}{\text{vol}(\xi(X_d))} \\ &= \lim_{n \rightarrow \infty} \sqrt{\frac{1}{12} - \frac{1}{12n^2}} = \sqrt{\frac{1}{12}} \\ &= \frac{\sqrt{\frac{(b-a)^2}{12}}}{b-a} = \text{ERD}(X_c). \end{aligned}$$

Um ponto interessante a lembrar é que não usamos uma aproximação da distribuição, pois a desigualdade informacional teria criado um “viés” de tal forma que o ERD da variável contínua não teria o mesmo valor que o limite do ERD da variável discretizada (vide COVER; THOMAS, 1992).

Usaremos a medida ERD para ordenar algumas das distribuições mais comuns que podem ser agrupadas nas seguintes famílias:<sup>8</sup>

1. família de distribuições  $f$  contínuas de tipo escala-localização:<sup>9</sup> uniforme, Gumbel, Laplace, logística, e gaussiana com

$$f_{\alpha,\beta}(x) = \frac{1}{\beta} \cdot f\left(\frac{x-\alpha}{\beta}\right).$$

2. família de distribuições contínuas  $f$  de tipo escala-formato<sup>10</sup>: lognormal, gama, triangular com

$$f_{\beta,\gamma}(x) = \frac{1}{\beta} \cdot f_{\gamma}\left(\frac{x}{\beta}\right)$$

3. outras:  $t$  de Student, beta.

A partir dos cálculos explícitos relatados no anexo B, concluímos que

$$\text{uniforme} \stackrel{\text{ERD}}{<} \text{Gumbel} \stackrel{\text{ERD}}{<} \text{Laplace} \stackrel{\text{ERD}}{<} \text{logística} \stackrel{\text{ERD}}{<} \text{normal}.$$

Esse resultado é interessante tendo em vista os seguintes conhecimentos heurísticos:

1. A distribuição normal é definida no intervalo  $[-\infty, \infty]$  e o “volume” com a medida de Lebesgue seria  $\infty$ . Em comparação a uma distribuição uniforme num intervalo restrito, cria-se a intuição de que a distribuição uniforme deveria ser preferida à normal. O *ERD* mostra que isso não é o caso.
2. Observando visualmente as funções de densidade das distribuições uniforme e normal, a maior concentração da normal induz uma preferência por ela, pressupondo, implicitamente, um corte da distribuição normal para que as duas distribuições sejam comparáveis. Essa intuição é confirmada pelos resultados de cálculo com o *ERD*.

<sup>8</sup>As densidades das funções de densidade são resumidas no anexo B.

<sup>9</sup>Essas distribuições dependem de um parâmetro de localização  $\alpha$  e de um parâmetro de escala  $\beta$ .

<sup>10</sup>Essas distribuições dependem de um parâmetro de formato  $\gamma$  e de um parâmetro de escala  $\beta$ .

## 6.2.2 Aplicação da medida ERD na precificação em mercados incompletos

### Exemplo paradigmático

Segundo a tabela dos volumes no anexo B, a função ERD é convexa em  $\pi$  para distribuições usuais. Então, podemos escrever o problema primal de otimização de utilidade esperada com o novo critério de decisão no mercado incompleto geral como

$$\begin{aligned} & \min_{\pi \in \mathcal{A}(x)} \text{ERD} \left( V_{t_1}^{x, \pi, 0, -D} \right) \\ & = \max_{\pi \in \mathcal{A}(x)} -\text{ERD} \left( V_{t_1}^{x, \pi, 0, -D} \right) \\ & = \max_{\pi \in \mathcal{A}(x)} -\frac{\sigma \left( V_{t_1}^{x, \pi, 0, -D} \right)}{\text{vol} \left( V_{t_1}^{x, \pi, 0, -D} \right)}. \end{aligned}$$

No exemplo paradigmático, é também fácil verificar a convexidade. Observamos que, sem a normalização pelo volume da distribuição, temos um problema parecido com o de mínima variância de SCHWEIZER (1992) se usarmos  $\sigma$  em vez de  $\sigma^2$  e a condição  $E \left[ V_{t_1}^{x, \pi, 0, -D} \right] = 0$ . Explicitando o valor da carteira em  $t_1$  nos diferentes estados, temos

$$\begin{aligned} V_{t_1}(\omega) &= x + \pi_{t_0}^{(1)} \left( S_{t_1}^{(1)}(\omega) - S_{t_0}^{(1)} \right) + \pi_{t_0}^{(0)} \cdot 0 - D_{t_1}(\omega) \\ &= \begin{cases} x + 20 \cdot \left( \pi_{t_0}^{(1)} - 1 \right), & \omega = \omega_1, \\ x, & \omega = \omega_2, \\ x - 10 \cdot \pi_{t_0}^{(1)}, & \omega = \omega_3, \end{cases} \end{aligned}$$

o que implica que

$$E \left[ V_{t_1}^{x, \pi, 0, -D} \right] = x + \frac{10}{4} \pi_{t_0}^{(1)} - 5.$$

Calcula-se o desvio-padrão do processo de valor da carteira em  $t_1$  como

$$\sigma \left( V_{t_1}^{x, \pi, 0, -D} \right) = \left( 75 + \frac{25}{4} \cdot \pi_{t_0}^{(1)} \left( 19\pi_{t_0}^{(1)} - 28 \right) \right)^{\frac{1}{2}}.$$

Para calcular o volume, observamos primeiro que

$$H \left( V_{t_1}^{x, \pi, 0, -D} \right) = - \sum_{i=1}^3 \mathbb{P}_{V_{t_1}}(\omega_i) \cdot \log \left( \mathbb{P}_{V_{t_1}}(\omega_i) \right)$$

e que

$$\mathbb{P}_{V_{t_1}}(\omega) = \begin{cases} \frac{1}{4}, & V_{t_1}^{x, \pi, 0, -D} = x + 20 \cdot \left( \pi_{t_0}^{(1)} - 1 \right), \\ \frac{1}{2}, & V_{t_1}^{x, \pi, 0, -D} = x, \\ \frac{1}{4}, & V_{t_1}^{x, \pi, 0, -D} = x - 10 \cdot \pi_{t_0}^{(1)}. \end{cases}$$

Nos casos especiais de  $\pi_{t_0}^{(1)} \in \{0, 1\}$ , temos que

$$\begin{aligned} H\left(V_{t_1}^{x, \{0,1\}, 0, -D}\right) &= -\left(\frac{3}{4} \cdot \log\left(\frac{3}{4}\right) + \frac{1}{4} \cdot \log\left(\frac{1}{4}\right)\right) \\ &= 2 - \frac{3}{4} \log(3) \\ &\approx 0,81. \end{aligned}$$

No outro caso especial de  $\pi_{t_0}^{(1)} = \frac{2}{3}$  temos dois possíveis estados de valores diferentes, cada um com a mesma probabilidade de  $\frac{1}{2}$ . Isso implica uma entropia de

$$H\left(V_T^{x, \frac{2}{3}, 0, -D}\right) = 1.$$

Portanto, para o caso  $\pi_{t_0}^{(1)} \notin \{0, 1, \frac{2}{3}\}$  segue

$$H\left(V_T^{x, \{0,1,\frac{2}{3}\}^c, 0, -D}\right) = \frac{3}{2}.$$

Observa-se que, no último caso de  $\pi_{t_0}^{(1)} \notin \{0, 1, \frac{2}{3}\}$ , a variação do valor da carteira acontece em três estados distintos. Os outros casos reduzem-se a dois e um estado contendo, assim, menos incerteza. Nesse exemplo discreto, vemos também que qualquer decisão de  $\pi_{t_0}^{(1)} \notin \{0, 1, \frac{2}{3}\}$  não teria influência no volume da distribuição do valor da carteira, mas ressaltamos que isso não valerá mais no caso contínuo. Para o caso  $\pi_{t_0}^{(1)} \notin \{0, 1, \frac{2}{3}\}$ , determina-se a estratégia ótima como  $\pi_{t_0}^{(1)} = \frac{14}{19}$ .

Podemos agora determinar o valor do ERD:

$$\text{ERD} = \begin{cases} \frac{\sqrt{\frac{75}{4}}}{\exp\left\{2 - \frac{3}{4} \log(3)\right\}}, & \pi_{t_0}^{(1)} = 1 \\ \frac{\sqrt{75}}{\exp\left\{2 - \frac{3}{4} \log(3)\right\}}, & \pi_{t_0}^{(1)} = 0 \\ \frac{\sqrt{\frac{100}{9}}}{e^1}, & \pi_{t_0}^{(1)} = \frac{2}{3} \\ \frac{\sqrt{\frac{200}{19}}}{\exp\left(\frac{3}{2}\right)}, & \pi_{t_0}^{(1)} = \frac{14}{19}, \end{cases}$$

ou, em aproximação numérica:<sup>11</sup>

$$\text{ERD} \approx \begin{cases} \frac{\sqrt{75}}{2 \cdot 2,25}, & \pi_{t_0}^{(1)} = 1 \\ \frac{\sqrt{75}}{2,25}, & \pi_{t_0}^{(1)} = 0 \\ \frac{10}{3 \cdot 2,72}, & \pi_{t_0}^{(1)} = \frac{2}{3} \\ \frac{10\sqrt{\frac{2}{19}}}{4,48}, & \pi_{t_0}^{(1)} = \frac{14}{19}, \end{cases}$$

$$\approx \begin{cases} 1,92 & \pi_{t_0}^{(1)} = 1 \\ 3,85 & \pi_{t_0}^{(1)} = 0 \\ 1,23 & \pi_{t_0}^{(1)} = \frac{2}{3} \approx 0,67 \\ 0,72 & \pi_{t_0}^{(1)} = \frac{14}{19} \approx 0,74. \end{cases}$$

Segue que devemos escolher a estratégia de  $\tilde{\pi}_{t_0}^{(1)} = \frac{14}{19}$ , o que seria a mesma no caso de minimizar somente a variância (ou o desvio-padrão) do valor da carteira. A normalização pelo volume desfavorece essa estratégia em relação às outras, comparando com o critério de variância pura. Mas, o ajuste pela normalização não é suficiente para mudar a decisão. Se tivéssemos o caso de  $\varepsilon \in \left(0, \sqrt{\frac{200}{19}} \cdot (\exp(\frac{1}{2}) - 1)\right)$  e

$$\sigma \left( V_{t_1}^{x, \frac{2}{3}, 0, -D} \right) = \sqrt{\frac{200}{19 + \varepsilon}} =: \tilde{\sigma},$$

poderíamos observar uma mudança da escolha entre o critério de variância e *ERD*. De fato,

$$\sqrt{\frac{200}{19 + \varepsilon}} = \tilde{\sigma} = \sigma \left( \pi_{t_0}^{(1)} = \frac{2}{3} \right) < \sigma \left( \pi_{t_0}^{(1)} = \frac{14}{19} \right) = \sqrt{\frac{200}{19}},$$

mas

$$\text{ERD} \left( \pi_{t_0}^{(1)} = \frac{14}{19} \right) = \sqrt{\frac{200}{19}} \exp \left( -\frac{3}{2} \right) < \sqrt{\frac{200}{19 + \varepsilon}} \exp(-1).$$

Finalmente, temos que verificar se as restrições de admissibilidade são respeitadas. Observamos que

$$V_{t_1}(\omega) \geq K, \quad K \in \mathbb{R},$$

se

$$\frac{K - x}{20} + 1 \leq \pi_{t_0}^{(1)} \leq \frac{K - x}{10}.$$

<sup>11</sup>Observe o uso do logaritmo na base 2.

Obviamente, tem que valer também  $x \geq K$ . Com a notação  $x_i := V_{t_1}^{x, \pi, 0, -D}(\omega_i)$ , o conjunto primal escreve-se como

$$\mathcal{X}(x) := \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_3 = 2x + 10\pi_{t_0}^{(1)} - 20, x_2 = x \right\}.$$

### O problema dual

Como a função da raiz é monótona nos reais positivos, teríamos a mesma estratégia ótima se usássemos a função-objetivo

$$\text{ERD}_{\sigma^2} \left( V_{t_1}^{x, \pi, 0, -D} \right) := \frac{\sigma^2(V_{t_1}^{x, \pi, 0, -D})}{e^{H(V_{t_1}^{x, \pi, 0, -D})}}. \quad (6.2)$$

Porém, o valor da função-objetivo seria diferente, ou seja,

$$\text{ERD}_{\sigma^2} \left( V_{t_1}^{x, \pi, 0, -D} \right) \approx 2, 3.$$

Infelizmente, no presente caso, o desvio-padrão não tem expressão analítica. Esse fato complica a dedução de uma solução analítica quando se analisa a condição de primeira ordem na otimização do problema dual. Portanto, decidimos mudar a nossa função-objetivo para (6.2) para poder melhor exemplificar o uso da teoria dual. O uso da dispersão entrópica relativa na sua versão original de (6.1), já foi mostrado na comparação das distribuições da seção anterior.

Usando a notação

$$\mathcal{L}(\pi, y) := -\text{ERD}_{\sigma^2} \left( V_{t_1}^{x, \pi, 0, -D} \right) - y \cdot (x + (\pi \cdot S)_{t_1} - D_{t_1} - K),$$

$y \in \mathbb{R}^+$ , temos para o modelo básico que

$$\begin{aligned} \sup_{\pi} \mathcal{L}(\pi, y) &= \sup_{\pi} \left\{ -y \cdot x - \text{ERD}_{\sigma^2} \left( V_{t_1}^{x, \pi, 0, -D} \right) - y \cdot ((\pi \cdot S)_{t_1} - D_{t_1} - K) \right\} \\ &= f^*(x^*) - y \cdot ((\tilde{\pi} \cdot S)_{t_1} - D_{t_1} - K), \end{aligned}$$

sendo que  $f^*(x^*) := \text{ERD}_{\sigma^2}(-y)$ . O problema nessa formulação é que  $\mathbb{P}_{V_{t_1}}$  não é contínua em  $\pi = \pi^{(1)}$  e, portanto, a entropia também não é:

$$H(\mathbb{P}_{V_{t_1}}) = \begin{cases} 2 - \frac{3}{4} \log(3), & \pi^{(1)} \in \{0, 1\}, \\ 1, & \pi^{(1)} = \frac{2}{3}, \\ \frac{3}{2}, & \pi^{(1)} \notin \{0, 1, \frac{2}{3}\}. \end{cases}$$

Deveríamos analisar os casos separadamente, mas como sabemos da seção anterior que  $\pi_{t_0}^{(1)} = \frac{14}{19}$  é a estratégia ótima em ações, vamos nos concentrar ao caso  $\pi^{(1)} \notin \{0, 1, \frac{2}{3}\}$ .

Com a notação,  $\Delta S = (S_{t_1} - S_{t_0})$  determinamos a estratégia ótima analisando a condição de primeira ordem:

$$\frac{\partial \mathcal{L}(\pi, y)}{\partial \pi} \Big|_{\pi^{(1)} \notin \{0, 1, \frac{2}{3}\}} = -\frac{\partial \text{ERD}_{\sigma^2} \left( V_{t_1}^{x, \pi, 0, -D} \right)}{\partial \pi} - y \cdot dS = 0.$$

Portanto,

$$\tilde{\pi}_y = \frac{14}{19} - e^{-\frac{3}{2}} \cdot y \cdot dS.$$

Inserindo essa estratégia ótima em  $\mathcal{L}(\pi, y)$ , sob a hipótese de  $\frac{14}{19} - e^{-\frac{3}{2}} \cdot y \cdot dS \notin \{0, 1, \frac{2}{3}\}$ , temos

$$\begin{aligned} \mathcal{L}(\tilde{\pi}_y, y) &= -\text{ERD}_{\sigma^2} \left( V_{t_1}^{x, \tilde{\pi}_y, 0, -D} \right) - y \cdot \left( x + (\tilde{\pi}_y \cdot S)_{t_1} - D_{t_1} - K \right) \\ &= -e^{-\frac{3}{2}} \left\{ 75 + \frac{25}{4} \left( \frac{14}{19} - e^{-\frac{3}{2}} \cdot y \cdot dS \right) \left( -19 \cdot e^{-\frac{3}{2}} \cdot y \cdot dS - 14 \right) \right\} \\ &\quad - y \left( x + \left( \frac{14}{19} - e^{-\frac{3}{2}} \cdot y \cdot dS \right) dS - D_{t_1} - K \right) \\ &= -25e^{-\frac{3}{2}} \left( \frac{8}{19} + \frac{19}{4} e^{-3} \cdot (y \cdot dS)^2 \right) \\ &\quad + e^{-\frac{3}{2}} \cdot (y \cdot dS)^2 - y \left( x + \frac{14}{19} dS - D_{t_1} - K \right). \end{aligned}$$

Derivando por  $y$ , segue

$$\begin{aligned} \frac{\partial \mathcal{L}(\tilde{\pi}_y, y)}{\partial y} &= -\frac{25 \cdot 19}{2} e^{-\frac{3}{2}} e^{-3} y \cdot dS^2 + 2 \cdot e^{-\frac{3}{2}} \cdot y \cdot dS^2 \\ &\quad - \left( x + \frac{14}{19} dS - D_{t_1} - K \right) \end{aligned}$$

e o valor ótimo é dado a partir da condição de primeira ordem por

$$\tilde{y} = \frac{x - D_{t_1} - K + \frac{14}{19} dS}{(dS)^2 \cdot 2 \cdot e^{-\frac{3}{2}} \left( 1 - \frac{19 \cdot 25}{4} e^{-3} \right)}.$$

Excluindo o caso  $\omega = \omega_2$ , para o qual vale que  $dS = 0$ , temos

$$\begin{aligned} \tilde{y}(\omega_1) &= \frac{x - 20 - K + \frac{14}{19} 20}{2(20)^2 \cdot e^{-\frac{3}{2}} \left( 1 - \frac{19 \cdot 25}{4} e^{-3} \right)} \\ &\approx \frac{x - K - 5,26}{-876,85}. \\ \tilde{y}(\omega_3) &= \frac{x - K + \frac{14}{19} \cdot (-10)}{2(-10)^2 \cdot e^{-\frac{3}{2}} \left( 1 - \frac{19 \cdot 25}{4} e^{-3} \right)} \\ &\approx \frac{x - K - 7,37}{-219,21}. \end{aligned}$$

Falta determinar o capital inicial  $x$  e o limite de crédito  $K$  tal que  $\tilde{y} = (\tilde{y}(\omega_1), \tilde{y}(\omega_2), \tilde{y}(\omega_3))$  seja uma probabilidade, i.e., para  $i = 1, 2, 3$ :

$$\begin{aligned}\tilde{y}(\omega_i) &\in [0, 1], \\ \sum_{i=1}^3 \tilde{y}(\omega_i) &= 1.\end{aligned}$$

A partir de  $\tilde{y}(\omega_1) \in [0, 1]$ , temos  $x \leq K + 5,26$  e  $x \geq K - 871,59$ . O caso de  $\tilde{y}(\omega_3) \in [0, 1]$  cria as condições de  $x \leq K + 7,37$  e  $x \geq K - 211,84$ . Segue que tem que valer  $x \in [K - 211,84, K + 5,26]$ . Finalmente, tem que valer também que

$$\tilde{y}(\omega_2) = 1 + \frac{x - K - 5,26}{876,85} + \frac{x - K - 7,37}{219,21} \in [0, 1].$$

Essa condição será satisfeita se (aproximadamente)

$$x \in [K - 168,42, K + 6,95].$$

Finalmente, tem que valer que

$$x \in [K - 168,42, K + 5,26]. \quad (6.3)$$

O conjunto dual é dado por

$$\mathcal{Y}(y) = \left\{ (y_1, y_2, y_3) \in \mathbb{R}^+ : \begin{aligned} y_1 &\approx \frac{x - K - 5,26}{-876,85}, \\ y_3 &\approx \frac{x - K - 7,37}{-219,21}, \\ y_2 &\approx 1 + \frac{x - K - 5,26}{876,85} + \frac{x - K - 7,37}{219,21} \end{aligned} \right\}.$$

Se quisermos achar o valor inicial  $x$  que seria negociado no mercado para o derivativo, temos que usar as estratégias de mínima *ERD* para o vendedor e o comprador do derivativo, respetivamente, no cálculo do valor esperado de cada um deles. Assim, obtemos o preço máximo de compra e o preço mínimo de venda. A sua interseção fornece o preço de mercado.<sup>12</sup>

<sup>12</sup>Na seção 6.3.4 calcularemos no exemplo paradigmático o preço do mercado e o comparamos com os preços de outras abordagens.



### Caso geral

Analizamos agora o comportamento da nossa regra de decisão ERD no mercado incompleto geral descrito na seção 2.2.

Primeiramente, queremos saber se o problema realmente tem uma solução nesse mercado:

**Lema 6.2** *No mercado da seção 2.2, vale que para todo  $K \in \mathbb{R}$  existe um capital inicial  $x \in \mathbb{R}$  e uma estratégia admissível  $\tilde{\pi} \in \mathcal{A}_K(x)$  tal que*

$$\tilde{\pi} = \arg \max_{\pi \in \mathcal{A}_K(x)} \left( -\text{ERD} \left( V_T^{x, \pi, 0, D} \right) \right). \quad (6.4)$$

**Prova.** Pressupomos  $K = x = 0$ . ERD é limitado, pois  $\pi$  tem que ser admissível. Portanto, junto com a convexidade, existe um ponto ótimo  $\tilde{\pi}$  de ERD em  $\mathcal{A}_K(x)$ . Como  $K$  e  $x$  são valores reais, para qualquer  $K \in \mathbb{R}$  pode ser escolhida  $x = -K$  e aplicamos o caso inicial. ■

Sabendo que existe uma solução, queremos analisar o problema dual correspondente. Para isso, precisamos deduzir a conjugada convexa.

**Lema 6.3** *A conjugada convexa de ERD é dada por*

$$\text{ERD}^*(x^*) = \max_{\pi} \left\{ -\text{ERD}(x) - \frac{dQ}{dP} \cdot (\pi S)_T \right\}.$$

Finalmente, podemos formular o problema dual e estabelecer a seguinte relação com o problema primal (6.4):

**Teorema 6.4** *O problema*

$$\begin{aligned} & \min_{Q \in \mathcal{Y}(y)} \text{ERD}(Q), \\ & \text{tal que } S \text{ é } Q\text{-martingale} \end{aligned}$$

é o problema dual de (6.4) e vale

$$\inf_{Q \in \mathcal{Y}(y)} \sup_{\pi \in \mathcal{A}_K(x)} \mathcal{L}(\pi, Q) = \sup_{\pi \in \mathcal{A}_K(x)} \inf_{Q \in \mathcal{Y}(y)} \mathcal{L}(\pi, Q),$$

com

$$\mathcal{L}(\pi, Q) = -\text{ERD} \left( V_T^{x, \pi, 0, D} \right) - \frac{dQ}{dP} \cdot (x - (\pi, dS)_T - D_T - K).$$

**Prova.** As características do conjunto de pagamentos acessíveis  $\mathcal{V}(x)$  e do conjunto dual dos processos de derivadas de Radon-Nikodym são garantidas para os modelos das seções 2 e 2.2. Falta verificar se  $U(x) = -ERD(x) = -\frac{\sigma(x)}{\exp(H(x))}$  satisfaz a condição de elasticidade assintótica razoável (vide definição 4.2). Para verificar essa condição é necessário lembrar que a utilidade quadrática<sup>13</sup> não representa uma utilidade em todo seu domínio, pois a sua derivada muda de sinal. Para contornar esse problema, pode ser pressuposta uma utilidade constante a partir do ponto de máximo. Mas, isso significa que a elasticidade assintótica é igual a 1, um resultado não desejado. Porém, no momento do uso de uma utilidade quadrática, o investidor deve ter consciência do fato que o seu critério de decisão não é mais racional a partir de um certo valor  $x_0$ . Para poder gerar valores  $x > x_0$ , é preciso definir alguma forma de utilidade e, já que a escolha é livre, usa-se uma para qual vale  $AE_{+\infty}(U_{x>x_0}) < 1$ . ■

## 6.3 A precificação consistente

### 6.3.1 Introdução

Introduzimos neste capítulo uma nova forma para atribuir um valor a uma variável aleatória no caso de um investidor averso ao risco.<sup>14</sup> O núcleo da nossa nova proposta é a seguinte suposição:

**Suposição (Consistência)** O agente é consistente na aplicação da função de valoração de uma variável aleatória de tal forma que, para qualquer  $X \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ , usa-se a mesma função  $v : L^\infty \rightarrow \mathbb{R}$ .

A suposição significa que um investidor consistente usa a mesma fórmula para avaliar todos os ativos. Tendo uma função  $r : L^\infty \rightarrow \mathbb{R}$  para avaliar o risco de  $X^{(0)}$ , escreve-se a fórmula de avaliação de  $X^{(0)}$  sob a hipótese da aversão ao risco como

$$v(X^{(0)}) = E[X^{(0)}] - r(X^{(0)}).$$

<sup>13</sup>No nosso caso, nem temos o termo constante que, no caso da utilidade esperada, daria o valor esperado.

<sup>14</sup>Usamos a noção de risco estabelecida na seção 3.1. *Aversão ao risco* significa então, em analogia à definição clássica de ROTHSCILD; STIGLITZ (1970), que o investidor prefere mais a menos.

É importante como identificamos o risco de uma variável  $X^{(i)}$ ,  $i \geq 0$ . Seguindo a discussão sobre risco na seção 3.2, sabemos que o risco para o investidor é a parte não desejável da distribuição em relação a um valor de referência.<sup>15</sup> Analisaremos, portanto, duas alternativas sendo que o risco de qualquer variável  $X^{(i)}$  é interpretado como

1. o valor de  $X^{(i)}$  no lado não desejado do  $i$ -ésimo *benchmark*  $BM^{(i)}$  :<sup>16</sup>

$$X^{(i+1)} := X^{(i)} \cdot \mathbf{1}_{\{X^{(i)} \leq BM^{(i)}\}}.$$

2. O desvio não desejado de  $X^{(i)}$  do *benchmark*  $BM^{(i)}$  :<sup>17</sup>

$$X^{(i+1)} := (BM^{(i)} - X^{(i)}) \cdot \mathbf{1}_{\{X^{(i)} \leq BM^{(i)}\}}.$$

A nova variável  $X^{(i+1)}$  representa o risco de  $X^{(i)}$  e tem que ser avaliada usando a mesma forma de avaliação de  $X^{(i)}$  para o investidor ser consistente.

*Observação 6.5* Teoricamente, podemos usar qualquer medida de incerteza e aplicar a hipótese da consistência. Por exemplo, podemos empregar a variância na análise desse capítulo, em vez das medidas propostas anteriormente.

Num primeiro passo, vamos analisar a alternativa 1, pois as formulações matemáticas que surgem são mais compactas. Assim, será mais claro o conceito de recursão envolvido. Usaremos também uma formulação alternativa que incorpora partes da distribuição da variável original  $X^{(0)}$ . Nessa formulação, ficará mais claro como serão adicionadas e subtraídas partes da distribuição para chegar ao valor final. Porém, como vamos mostrar numa comparação dos valores de uma distribuição uniforme com uma distribuição beta, a escolha entre variáveis não é a esperada economicamente, pois cria-se preferência para concentração menor (compara com a discussão na seção 3.1.2). Na observação 1 na seção 3.2.1, já mencionamos o resultado de DANIELSSON et al. (2005) que consta que a medida de perda esperada não está em coerência com a dominância estocástica de primeira e segunda ordem e mostraremos que *perda esperada* também não é consistente com a

<sup>15</sup>Sem perda de generalidade, pressupomos que a parte não-desejável de uma variável aleatória  $X$  em relação a um *benchmark*  $BM$  é dada pelos valores “menores” do *benchmark*, ou seja, pelo conjunto  $\{\omega \in \Omega : X(\omega) \leq BM\}$ . Dessa forma,  $X$  pode modelar, por exemplo, um retorno, mas não uma perda.

<sup>16</sup>Essa abordagem é em coerência com a medida de risco “perda esperada”.

<sup>17</sup>Essa abordagem é em coerência com a medida de risco “*expected regret*”.

preferência por momentos. Essa deficiência será resolvida no uso da segunda alternativa. Nesse caso, serão punidos os desvios do valor de referência que são maiores para a distribuição uniforme de que para a beta.

### 6.3.2 Risco como valores não desejados

Pressupondo que o investidor avalia o risco de uma variável aleatória com o valor da parte não desejada em relação ao seu *benchmark*, trabalhamos nesta seção com a seguinte hipótese:

**Conjectura 6.1** *O (valor do) risco de uma variável aleatória  $X^{(i)}$  é dado por*

$$r(X^{(i)}) := v(X^{(i+1)}) = v\left(X^{(i)} \cdot \mathbf{1}_{\{X^{(i)} \leq BM^{(i)}\}}\right). \quad (6.5)$$

Para facilitar a notação, pressupomos que  $X^{(0)} \geq 0$  q.c., o que evita o uso de valores absolutos. Por exemplo, sendo  $X^{(1)} = X^{(0)} \cdot \mathbf{1}_{\{X^{(0)} \leq 0\}}$  significa que  $X^{(1)} \leq 0$  q.c. e  $E[X^{(1)}] \leq 0$ , o que implica que devemos usar a notação

$$v(X^{(0)}) = E[X^{(0)}] - |r(X^{(0)})|.$$

Além disso, para evitar degenerações, pressupomos também que  $BM^{(i)} \in \text{supp}(X^{(i)}) \subseteq \mathbb{R}^+, \forall i = 1, 2, \dots$ , e, por razões de simplicidade, que o agente não atribui valor para possíveis desvios favoráveis do benchmark<sup>18</sup>.

**Notação 6.6** *Como a distribuição da nova variável  $X^{(i)}$  é dada pelo condicionamento da*

<sup>18</sup>Se o investidor atribuisse valor aos desvios desejados, a fórmula de avaliação seria escrita como

$$\begin{aligned} v(X^{(i)}) &= E[X^{(i)}] - r(X^{(i)}) + s(X^{(i)}), \\ s &: L^\infty \rightarrow \mathbb{R}. \end{aligned}$$

A suposição de  $s \equiv 0$  pode ser contestada por não ser realista, mas ela simplifica a notação e evita condições sobre as relações entre  $r(X^{(i)})$  e  $s(X^{(i)})$ .

variável  $X^{(i)}$  ao conjunto  $\{X^{(i)} \leq BM^{(i)}\}$ , ou seja, vale para  $A \in \mathcal{F}$  que<sup>19</sup>

$$\begin{aligned} \mathbb{P}_{X^{(i)}}(A) &= \mathbb{P}(\{\omega \in \Omega : X^{(i)}(\omega) \in A\}) \\ &= \mathbb{P}(\{\omega \in \Omega : X^{(i-1)}(\omega) \in A\} \mid X^{(i-1)}(\omega) \leq BM^{(i-1)}) \\ &= \frac{\mathbb{P}(\{\omega \in \Omega : X^{(i-1)}(\omega) \cdot \mathbf{1}_{\{X^{(i-1)} \leq BM^{(i-1)}\}}(\omega) \in A\})}{\mathbb{P}(\{\omega \in \Omega : X^{(i-1)}(\omega) \leq BM^{(i-1)}\})} \\ &= \frac{\mathbb{P}_{X^{(i-1)} \mathbf{1}_{\{X^{(i-1)} \leq BM^{(i-1)}\}}}(A)}{\mathbb{P}_{X^{(i-1)}}(X^{(i-1)} \leq BM^{(i-1)})}. \end{aligned}$$

Baseado nessa notação, devemos avaliar o valor esperado condicionado como

$$\begin{aligned} E_{\mathbb{P}_{X^{(i)}}}[X^{(i)} \mid X^{(i)} \leq BM^{(i)}] &= \frac{E_{\mathbb{P}_{X^{(i)}}}[X^{(i)} \cdot \mathbf{1}_{\{X^{(i)} \leq BM^{(i)}\}}]}{\mathbb{P}_{X^{(i)}}(X^{(i)} \leq BM^{(i)})} \\ &= E_{\mathbb{P}_{X^{(i)} \mid X^{(i)} \leq BM^{(i)}}}[X^{(i)} \cdot \mathbf{1}_{\{X^{(i)} \leq BM^{(i)}\}}]. \end{aligned}$$

Como essa notação precisa complicar a leitura, usaremos nos casos óbvios a notação

$$\begin{aligned} E_{X^{(i)}}[X^{(i)} \mid X^{(i)} \leq BM^{(i)}] &:= E_{\mathbb{P}_{X^{(i)}}}[X^{(i)} \mid X^{(i)} \leq BM^{(i)}] \\ &= E_{\mathbb{P}_{X^{(i+1)}}}[X^{(i)} \cdot \mathbf{1}_{\{X^{(i)} \leq BM^{(i)}\}}] =: E_{X^{(i+1)}}[X^{(i)} \cdot \mathbf{1}_{\{X^{(i)} \leq BM^{(i)}\}}] \end{aligned}$$

Se o investidor for consistente, então a fórmula de avaliação do risco  $r(X^{(0)})$  será

$$\begin{aligned} r(X^{(0)}) &= v(X^{(0)} \cdot \mathbf{1}_{\{X^{(0)} \leq BM^{(0)}\}}) \\ &= E[X^{(0)} \mid X^{(0)} \leq BM^{(0)}] - r(X^{(0)} \cdot \mathbf{1}_{\{X^{(0)} \leq BM^{(0)}\}}) \\ &= E[X^{(1)}] - r(X^{(1)}), \end{aligned}$$

e ele continua a aplicar a mesma avaliação para  $r(X^{(1)})$ .

### A aplicação recursiva de avaliação

Na seção anterior, vimos que um investidor consistente deve aplicar a mesma função de avaliação não somente para o ativo  $X^{(0)}$ , mas também para o seu risco, para  $X^{(1)}$ , o risco de  $X^{(1)}$  etc. Expandindo a recursão até o nível  $n > 1$ , obtemos a seguinte formulação:

$$\begin{aligned} v(X^{(0)}) &= E[X^{(0)}] - E[X^{(1)}] + r(X^{(1)}) \\ &= E[X^{(0)}] - E[X^{(1)}] + E[X^{(2)}] - r(X^{(2)}) \\ &= \dots \\ &= E[X^{(0)}] + \sum_{i=0}^{n-1} (-1)^{i+1} E[X^{(i)} \mid X^{(i)} \leq BM^{(i)}] + (-1)^{n-1} r(X^{(n)}). \end{aligned}$$

<sup>19</sup> Assim, garantimos que  $F_{\mathbb{P}_{X^{(i)}}}(BM^{(i)}) = 1$ .

Se o investidor não quer levar em conta o risco no nível  $n$ , define-se  $r_n := r(X^n) = 0$ . Podemos também introduzir formalmente o valor pela abordagem de consistência com a seguinte definição:<sup>20</sup>

**Definição 6.4 (Precificação consistente)** Dizemos que um investidor averso ao risco usa uma abordagem de precificação consistente (ao nível  $n$ ) se ele determina o valor de uma variável aleatória  $X$  com benchmark de nível- $i$   $BM^{(i)}$ ,  $i < n$ , como

$$v_n(X^{(0)}) := E[X^{(0)}] + \sum_{i=0}^{n-1} (-1)^{i+1} E[X^{(i)} | X^{(i)} \leq BM^{(i)}]. \quad (6.6)$$

A importância dessa abordagem vem do seguinte resultado, que determina que o valor da variável  $X^{(0)}$  converge por  $n \rightarrow \infty$ .

**Lema 6.7 (Valor limite com precificação consistente)** A seqüência

$$S_n := \sum_{i=0}^{n-1} (-1)^{i+1} E[X^{(i)} | X^{(i)} \leq BM^{(i)}]$$

converge quase certamente por  $n \rightarrow \infty$  para um valor finito se  $X$  for limitado inferiormente e a seqüência  $\{BM^{(i)}\}_{i=1,2,\dots}$  for decrescente.

**Prova.** Seja sem limitação de generalidade  $\inf_{\omega \in \Omega} X(\omega) = 0$ . Como

$$\mathbb{P}\left(\bigcap_{i=1}^{\infty} \{\omega \in \Omega : X^{(i)}(\omega) \leq BM^{(i)}\}\right) = 0,$$

temos que  $\lim_{n \rightarrow \infty} X^{(n)} = 0$  quase certamente. Isso implica que

$$E[X^{(i)}] \geq E[X^{(i+1)}]$$

e que, para ao menos um  $i$ , a desigualdade é estrita porque  $\{BM^{(i)}\}_{i=1,2,\dots}$  é decrescente. Como  $S_n$  é alternando, segue o resultado por uma aplicação do critério de Leibniz. ■

Observamos que o risco de  $X^{(0)}$  é dado por

$$\begin{aligned} -r(X^{(0)}) &= \lim_{n \rightarrow \infty} \left\{ \sum_{i=0}^{n-1} (-1)^{i+1} E[X^{(i)} | X^{(i)} \leq BM^{(i)}] + (-1)^{n-1} \cdot r(X^{(n)}) \right\} \\ &= \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} (-1)^{i+1} E[X^{(i)} | X^{(i)} \leq BM^{(i)}], \end{aligned}$$

<sup>20</sup>Pressupomos que preço é igual ao valor.

como  $\lim_{n \rightarrow \infty} (-1)^{n+1} \cdot r(X^{(n)}) = 0$  porque  $\lim_{n \rightarrow \infty} X^{(n)} = 0$ , q.c. Mas, nem sempre será possível analisar a seqüência até o infinito e um nível de corte de expansão, e.g.  $n \in \mathbb{N}$ , tem que ser escolhido. A avaliação do risco restante  $r(X^{(n)})$  tem um papel importante na comparação da nossa abordagem com outras abordagens existentes, como veremos na seção 6.3.2. Para formalizar a nomenclatura, introduzimos as seguintes definições adicionais:

**Definição 6.5 (Aversão ao risco e indiferença ao risco)** *Se um investidor escolhe o grau de expansão  $n \in \mathbb{N}$ , dizemos que ele é*

- *Averso ao risco até o grau  $n$ .*
- *Indiferente ao risco de grau  $n$ , se  $r(X^{(n)}) = 0$ .*

Se um investidor for **avverso ao risco ao menos de primeiro grau**, ele aplica a fórmula de avaliação ao menos duas vezes e vale que

$$v(X^{(0)}) \in [E[X^{(0)}] - E[X^{(0)} | X^{(0)} \leq BM^{(0)}], E[X^{(0)}].$$

A segunda parte da definição significa que o investidor aplica a fórmula de avaliação  $n$  vezes, mas na próxima aplicação, ele fica indiferente ao risco restante, ou seja:

$$v(X^{(n)}) = E[X^{(n)}].$$

Segue diretamente a seguinte conexão com o critério de *perda esperada*:

**Proposição 6.8 (Consistência de *perda esperada*)** *Um investidor que usa o critério de “perda esperada” com valor de referência  $BM^{(0)}$  é consistente e é indiferente ao risco de segundo grau:*

$$\begin{aligned} ES(X^{(0)}) &= E[X^{(0)}] - r(X^{(0)}) \\ &= E[X^{(0)}] - E[X^{(1)}] \\ &= v_1(X^{(0)}). \end{aligned}$$

Vamos ilustrar o funcionamento da abordagem num exemplo simples (vide figura 6.1 para uma visualização).

**Exemplo 6.2 (Distribuição uniforme contínua)** Seja  $X^{(0)} \sim U([0, 1])$  - distribuída uniformemente no intervalo fechado  $[0, 1] \subseteq \mathbb{R}^+$  - então  $E[X^{(0)}] = \frac{1}{2}$ ,  $X^{(1)} = X^{(0)} \cdot \mathbf{1}_{\{X^{(0)} \leq \frac{1}{2}\}}$  com  $X^{(1)} \sim U([0, \frac{1}{2}])$ , e, para  $n > 0$ :

$$X^{(n)} = X^{(n-1)} \cdot \mathbf{1}_{\{X^{(n-1)} \leq E[X^{(n-1)}]\}} = X^{(0)} \cdot \mathbf{1}_{\{X^{(0)} \leq (\frac{1}{2})^{n-1}\}}.$$

A partir disso, junto com (6.6), segue que

$$\begin{aligned} v(X^{(0)}) &= \frac{1}{2} - \frac{1}{4} + \sum_{i=1}^{\infty} (-1)^{i+1} \cdot E\left[X^{(0)} \mid X^{(0)} \leq \left(\frac{1}{2}\right)^{i+1}\right] \\ &= \frac{1}{2} - \frac{1}{4} - \sum_{i=1}^{\infty} \left(-\frac{1}{2}\right)^{i+2} \\ &= \frac{1}{2} - \sum_{i=2}^{\infty} \left(-\frac{1}{2}\right)^i \\ &= 1 - \sum_{i=0}^{\infty} \left(-\frac{1}{2}\right)^i = 1 - \frac{1}{1 + \frac{1}{2}} = \frac{1}{3}. \end{aligned}$$

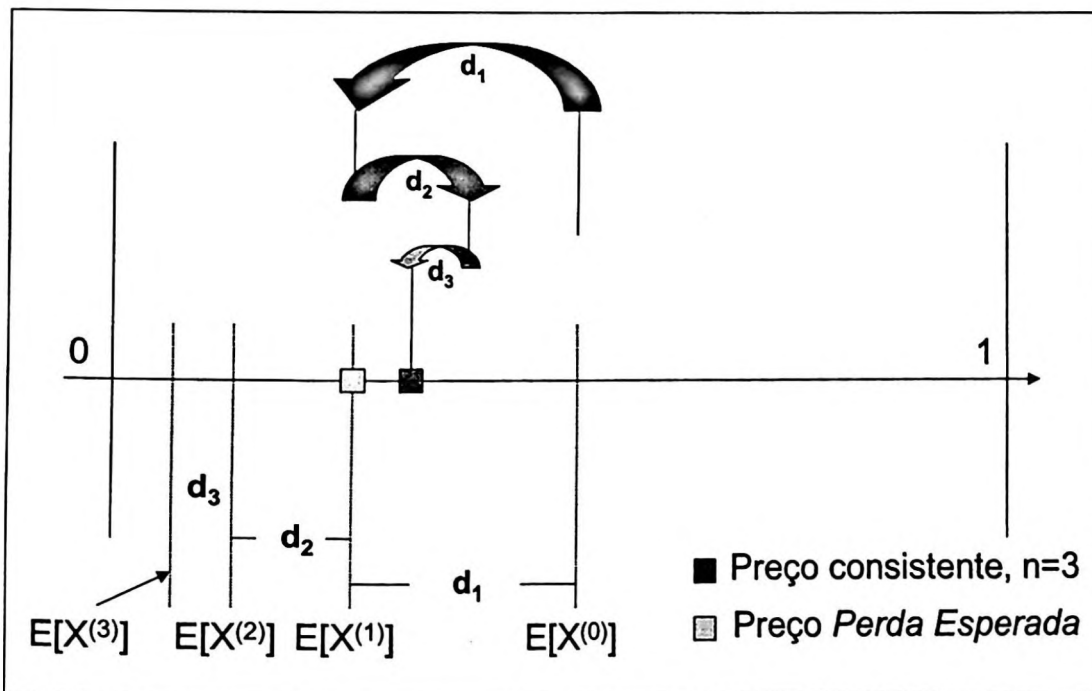


Figura 6.1: A abordagem de consistência no caso de distribuição uniforme contínua.

O valor de precificação consistente para uma distribuição uniforme diverge então de forma significativa do valor atribuído à distribuição se usarmos a medida de incerteza desvio-padrão, cujo valor seria

$$v_{DP}(X^{(0)}) = 0,5 - \frac{1}{\sqrt{12}} \approx 0,21 < 0,3.$$



Essa diferença pode ser interpretada como uma aversão ao risco exagerado no uso do desvio-padrão, ou seja, um investidor consistente aceitaria um preço maior por um ativo com retornos uniformes em  $[0, 1]$ .

### O “*fatiamento*” da distribuição

Na aplicação recursiva, observamos que a aproximação do valor limite é feita “em saltos”: em cada dois passos é subtraído um valor e adicionado outro. Convém juntar esses dois passos e subtrair somente a diferença restante. Usaremos agora a idéia de abrir a recursão da seção anterior para escrever o valor final da variável aleatória  $X^{(0)}$  de uma maneira que soma “*fatias*” da distribuição (interquantis). A seqüência de interquantis a serem somados depende do *benchmark*. Para obter um resultado explícito, pressupomos que a seqüência de *benchmarks* é não somente decrescente, mas também determinística. Observamos que

$$\begin{aligned} X^{(i)} &= X^{(i-1)} \cdot \mathbf{1}_{\{X^{(i-1)} \leq BM^{(i-1)}\}} \\ &= X^{(i-2)} \cdot \mathbf{1}_{\{X^{(i-2)} \leq BM^{(i-2)}\}} \cdot \mathbf{1}_{\{X^{(i-1)} \leq BM^{(i-1)}\}} \\ &= \dots \\ &= X^{(0)} \cdot \prod_{j=0}^{i-1} \mathbf{1}_{\{X^{(j)} \leq BM^{(j)}\}}. \end{aligned}$$

Para avaliar essa expressão, usaremos o seguinte resultado.

**Lema 6.9** Para qualquer  $n \geq 0$  vale que:

$$E_{\mathbf{P}} \left[ X^{(0)} \cdot \prod_{j=0}^n \mathbf{1}_{\{X^{(j)} \leq BM^{(j)}\}} \right] = E_{\mathbf{P}_{X^{(0)}}} [X^{(0)} | X^{(0)} \leq BM^{(n)}].$$

**Prova.** Escreve-se primeiro

$$E_{\mathbf{P}} \left[ X^{(0)} \cdot \prod_{j=0}^n \mathbf{1}_{\{X^{(j)} \leq BM^{(j)}\}} \right] = E_{\mathbf{P}_{X^{(0)}}} \left[ X^{(0)} \cdot \prod_{j=0}^n \mathbf{1}_{\{X^{(j)} \leq BM^{(j)}\}} \right]$$

e observa-se que

$$\begin{aligned}
 \bigcap_{j=0}^n \{X^{(j)} \leq BM^{(j)}\} &= \{X^{(0)} \leq BM^{(0)}\} \cap \{X^{(1)} \leq BM^{(1)}\} \cap \dots \\
 &\quad \cap \{X^{(n)} \leq BM^{(n)}\} \\
 &= \{X^{(0)} \leq BM^{(0)}\} \cap \left\{ X^{(0)} \mathbf{1}_{\{X^{(0)} \leq BM^{(0)}\}} \leq BM^{(1)} \right\} \cap \dots \\
 &\quad \cap \left\{ X^{(n-1)} \mathbf{1}_{\{X^{(n-1)} \leq BM^{(n-1)}\}} \leq BM^{(n)} \right\} \\
 &= \{X^{(0)} \leq BM^{(0)}\} \cap \{X^{(0)} \leq BM^{(1)}\} \cap \dots \\
 &\quad \cap \{X^{(0)} \leq BM^{(n)}\}.
 \end{aligned}$$

Mas, como  $\{BM^{(i)}\}_{i \geq 0}$  é uma seqüência decrescente, segue que

$$\bigcap_{j=0}^n \{X^{(j)} \leq BM^{(j)}\} = \{X^{(0)} \leq BM^{(n)}\}.$$

■

Portanto, temos que o valor do ativo  $X^{(0)}$  pode ser representado por

$$v(X^{(0)}) = E[X^{(0)}] + \sum_{i=0}^{n-1} (-1)^{i+1} \cdot E_{\mathbb{P}_{X^{(0)}}}[X^{(0)} | X^{(0)} \leq BM^{(i)}] + (-1)^{n+1} r(X^{(n)}). \quad (6.7)$$

Pressupomos agora que a distribuição de  $X^{(0)}$  seja equivalente à medida de Lebesgue  $\lambda$  em  $([a, b], \mathcal{B}([a, b]))$ ,  $0 \leq a \leq b$ , ou seja  $\mathbb{P}_{X^{(0)}} \sim \lambda$ , e que existe uma densidade  $f$  tal que

$$\mathbb{P}_{X^{(0)}}(A) = F(A) = \int_a^b f(x) \lambda(x) = \int_a^b f(x) d(x).$$

Então, o termo da soma em (6.7) pode ser expresso por integrais:

$$\sum_{i=0}^{n-1} (-1)^{i+1} \cdot E_{\mathbb{P}_{X^{(0)}}}[X^{(0)} | X^{(0)} \leq BM^{(i)}] \quad (6.8)$$

$$\begin{aligned}
 &= - \int_a^{BM^{(0)}} x \frac{f(x)}{F_{X^{(0)}}(BM^{(0)})} dx + \int_a^{BM^{(1)}} x \frac{f(x)}{F_{X^{(0)}}(BM^{(1)})} dx \\
 &\quad - \int_a^{BM^{(2)}} x \frac{f(x)}{F_{X^{(0)}}(BM^{(2)})} dx + \dots \quad (6.9)
 \end{aligned}$$

$$+ (-1)^n \int_a^{BM^{(n-1)}} x \frac{f(x)}{F_{X^{(0)}}(BM^{(n-1)})} dx. \quad (6.10)$$

Podemos agora escrever essas integrais de forma que todas tenham como limite superior de integração o menor *benchmark*. Exceto na última integral, sobrarão parcelas

adicionais nas outras:

$$\begin{aligned} & \sum_{i=0}^{n-1} (-1)^{i+1} \cdot E_{\mathbb{P}_{X^{(0)}}} [X^{(0)} | X^{(0)} \leq BM^{(i)}] \\ = & - \int_a^{BM^{(n-1)}} x \frac{f(x)}{F_{X^{(0)}}(BM^{(0)})} dx - \int_{BM^{(n-1)}}^{BM^{(0)}} x \frac{f(x)}{F_{X^{(0)}}(BM^{(0)})} dx \end{aligned} \quad (6.11)$$

$$+ \int_a^{BM^{(n-1)}} x \frac{f(x)}{F_{X^{(0)}}(BM^{(1)})} dx + \int_{BM^{(n-1)}}^{BM^{(1)}} x \frac{f(x)}{F_{X^{(0)}}(BM^{(1)})} dx + \dots \quad (6.12)$$

$$\begin{aligned} & + (-1)^{n-1} \int_a^{BM^{(n-1)}} x \frac{f(x)}{F_{X^{(0)}}(BM^{(n-2)})} dx \\ & + (-1)^{n-1} \int_{BM^{(n-1)}}^{BM^{(n-2)}} x \frac{f(x)}{F_{X^{(0)}}(BM^{(n-2)})} dx \end{aligned} \quad (6.13)$$

$$+ (-1)^n \int_a^{BM^{(n-1)}} x \frac{f(x)}{F_{X^{(0)}}(BM^{(n-1)})} dx \quad (6.14)$$

$$\begin{aligned} = & \sum_{i=0}^{n-1} \frac{(-1)^{i+1}}{F_{X^{(0)}}(BM^{(i)})} \int_a^{BM^{(n-1)}} x f(x) dx - \int_{BM^{(n-1)}}^{BM^{(0)}} x \frac{f(x)}{F_{X^{(0)}}(BM^{(0)})} dx \\ & + \int_{BM^{(n-1)}}^{BM^{(1)}} x \frac{f(x)}{F_{X^{(0)}}(BM^{(1)})} dx + \dots \\ & + (-1)^{n-1} \int_{BM^{(n-1)}}^{BM^{(n-2)}} x \frac{f(x)}{F_{X^{(0)}}(BM^{(n-2)})} dx. \end{aligned} \quad (6.15)$$

Continuando esse “fatiamento” das integrais, obtemos

$$\begin{aligned} & \sum_{i=0}^{n-1} (-1)^{i+1} \cdot E [X^{(0)} | X^{(0)} \leq BM^{(i)}] \\ = & \sum_{i=0}^{n-1} \frac{(-1)^{i+1}}{F_{X^{(0)}}(BM^{(i)})} \int_a^{BM^{(n-1)}} x f(x) dx \\ & + \sum_{i=0}^{n-2} \frac{(-1)^{i+1}}{F_{X^{(0)}}(BM^{(i)})} \int_{BM^{(n-1)}}^{BM^{(n-2)}} x f(x) dx \\ & + \sum_{i=0}^{n-3} \frac{(-1)^{i+1}}{F_{X^{(0)}}(BM^{(i)})} \int_{BM^{(n-2)}}^{BM^{(n-3)}} x f(x) dx + \dots \\ & + \sum_{i=0}^0 \frac{(-1)^{i+1}}{F_{X^{(0)}}(BM^{(i)})} \int_{BM^{(1)}}^{BM^{(0)}} x f(x) dx. \end{aligned}$$

Provamos assim o seguinte lema:

**Lema 6.10 (Valor consistente por fatiamento)** *O valor consistente de um ativo  $X^{(0)}$  pode ser escrito para qualquer  $n \geq 1$  como*

$$\begin{aligned} v(X^{(0)}) &= E[X^{(0)}] \\ &+ \sum_{i=0}^{n-1} \frac{(-1)^{i+1}}{F_{X^{(0)}}(BM^{(i)})} \int_a^{BM^{(n-1)}} x f(x) dx \\ &+ \sum_{j=0}^{n-1} \sum_{i=0}^j \frac{(-1)^{i+1}}{F_{X^{(i)}}(BM^{(i)})} \int_{BM^{(i)}}^{BM^{(j-1)}} x f(x) dx. \end{aligned}$$

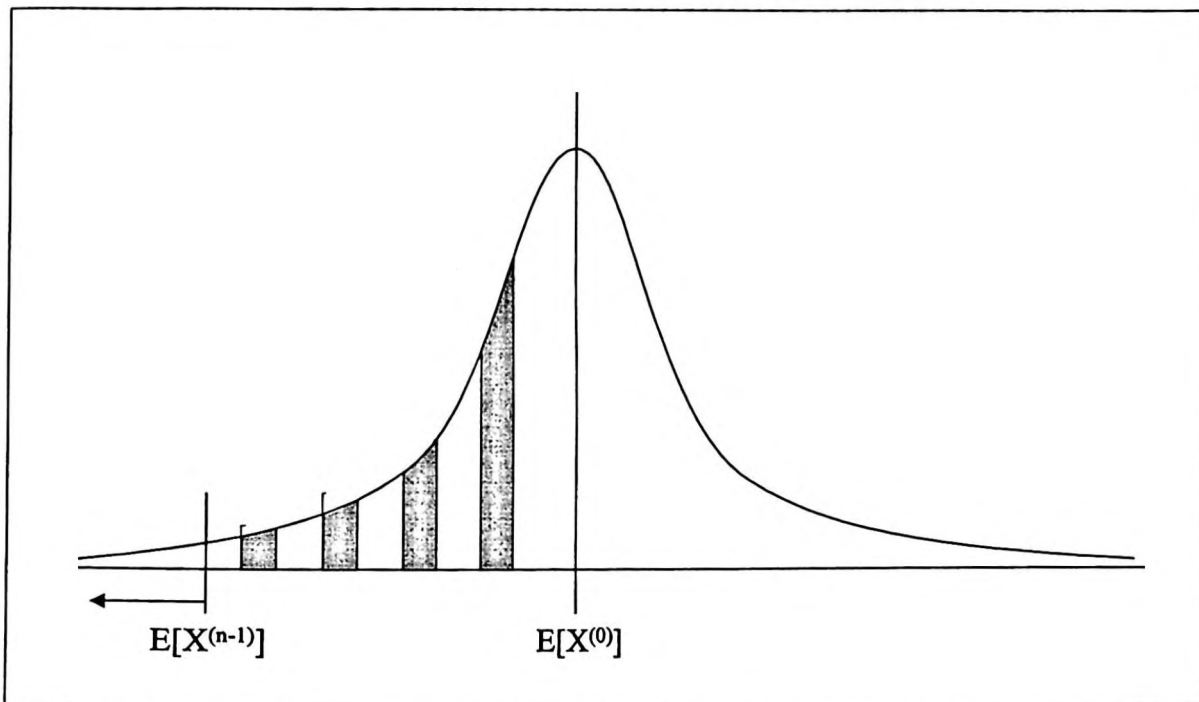


Figura 6.2: O *fatiamento* da distribuição na precificação por consistência.

Calcularemos agora o valor consistente de um ativo no caso específico de  $X^{(0)} \sim \text{Beta}(4, 4)$ :<sup>21</sup>

**Lema 6.11 (Precificação consistente para uma Beta em  $(0, 1)$ )** *Seja  $X^{(0)} \sim \text{Beta}(4, 4)$  e pressuponha que o investidor é indiferente ao risco ao nível  $n = 5$ . Seja também  $BM^{(i)} = E[X^{(i)}]$ . Então, vale que o valor de  $X^{(0)}$  no nível  $n = 5$  é dado por*

$$v_{n=5}(X^{(0)}) \approx 0,365$$

<sup>21</sup>Usamos uma distribuição *Beta* com parâmetros  $a = 4$  e  $b = 4$ , pois o formato da densidade se parece com a de uma normal. Não usaremos uma densidade normal, pois ela é definida também nos valores negativos e tem domínio infinito, o que complicaria a notação.

**Prova.** Com a definição de  $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ , e, no nosso caso,  $a = b = 4$  temos que

$$f_{\text{Beta}(4,4)}(x) = \frac{1}{B(4,4)} x^3 (1-x)^3 \cdot \mathbf{1}_{(0,1)}(x)$$

e

$$F_{\text{Beta}(4,4)}(t) = \frac{1}{B(4,4)} \int_0^t x^3 (1-x)^3 dx$$

e portanto, com  $c \in (0, 1)$

$$\begin{aligned} f_{\text{Beta}(4,4)}(x|x \leq c) &= \frac{f_{\text{Beta}(4,4)}(x)}{F_{\text{Beta}(4,4)}(c)} \\ &= \frac{x^3 (1-x)^3 \cdot \mathbf{1}_{(0,1)}(x)}{\int_0^c x^3 (1-x)^3 dx}. \end{aligned}$$

Sabendo que  $E[X^{(0)}] = \frac{a}{a+b} = \frac{1}{2}$ , escreve-se a fórmula (6.7) como

$$\begin{aligned} v_n(X^{(0)}) &= \frac{1}{2} + \sum_{i=0}^{n-1} (-1)^{i+1} \cdot E[X^{(0)} | X^{(0)} \leq E[X^{(i)}]] \\ &= \frac{1}{2} + \sum_{i=0}^{n-1} (-1)^{i+1} \cdot \int_0^{E[X^{(i)}]} x \cdot f_{\text{Beta}(4,4)}(x|x \leq E[X^{(i)}]) dx \\ &= \frac{1}{2} + \sum_{i=0}^{n-1} (-1)^{i+1} \cdot \int_0^{E[X^{(i)}]} \frac{x^4 (1-x)^3}{\int_0^{E[X^{(i)}]} x^3 (1-x)^3 dx} dx \\ &= \frac{1}{2} + \sum_{i=0}^{n-1} (-1)^{i+1} \cdot \frac{\int_0^{E[X^{(i)}]} x^4 (1-x)^3 dx}{\int_0^{E[X^{(i)}]} x^3 (1-x)^3 dx}. \end{aligned}$$

Já que

$$\begin{aligned} \int_a^b x^3 (1-x)^3 dx &= \left[ -\frac{1}{7}x^7 + \frac{1}{2}x^6 - \frac{3}{5}x^5 + \frac{1}{4}x^4 \right]_a^b, \\ \int_a^b x^4 (1-x)^3 dx &= \left[ -\frac{1}{8}x^8 + \frac{3}{7}x^7 - \frac{1}{2}x^6 + \frac{1}{5}x^5 \right]_a^b, \end{aligned}$$

temos que

$$v_n(X^{(0)}) = \frac{1}{2} + \sum_{i=0}^{n-1} (-1)^{i+1} \cdot \frac{-\frac{1}{8} \cdot E[X^{(i)}]^4 + \frac{3}{7}E[X^{(i)}]^3 - \frac{1}{2}E[X^{(i)}]^2 + \frac{1}{5}E[X^{(i)}]}{-\frac{1}{7}E[X^{(i)}]^3 + \frac{1}{2}E[X^{(i)}]^2 - \frac{3}{5}E[X^{(i)}] + \frac{1}{4}}.$$

Agora, calculam-se os passos da recursão para  $i = 1, \dots, 4$  como

$$E[X^{(1)}] \approx 0,363,$$

$$E[X^{(2)}] \approx 0,275,$$

$$E[X^{(3)}] \approx 0,212,$$

$$E[X^{(4)}] \approx 0,165,$$

e, portanto,

$$\begin{aligned} v_{n=5}(X^{(0)}) &= \frac{1}{2} - 0,363 + 0,275 - 0,212 + 0,165 - r(X^{(5)}) \\ &\approx 0,365 - r(X^{(5)}). \end{aligned}$$

Desprezando o risco a partir do nível  $n = 5$ , temos  $r(X^{(5)}) = 0$  e o investidor avaliará o ativo  $X^{(0)}$  com 0,365. ■

Quando comparamos esse resultado com o cálculo do valor consistente da distribuição uniforme em  $[0, 1]$ , observamos que o valor atribuído à distribuição beta é ligeiramente maior, ou seja

$$v_{n=5}(X_{\text{Beta}}^{(0)}) \approx 0,365 > 0,3\bar{3} = v(X_{\text{Uniforme}}^{(0)}).$$

Devido à maior incerteza da uniforme, é economicamente esperado obter um valor maior para a distribuição beta. Porém, é importante ressaltar que usamos um corte da recursão ao nível 5, o que permite ainda variações no valor exato de  $X^{(0)}$ . Na próxima seção, veremos que a diferença dos riscos restante é significativa e, por exemplo, no caso de  $n = 14$ , mostramos que

$$v_{n=14}(X_{\text{Beta}}^{(0)}) \approx 0,299 < 0,3\bar{3} = v(X_{\text{Uniforme}}^{(0)}).$$

Esse é um resultado contra-intuitivo, pois no caso de preferência por momentos,  $X_{\text{Beta}}^{(0)}$  deveria ser preferido. Mas, apesar de que DANIELSSON et al. (2005) mostrarem que a medida de risco *perda esperada* não é coerente com a dominância estocástica, não podemos afirmar que  $X_{\text{Beta}}^{(0)}$  domina  $X_{\text{Uniforme}}^{(0)}$  estocasticamente de segundo grau. Sabemos somente que

$$\begin{aligned} \mu(X_{\text{Beta}}^{(0)}) &= \mu(X_{\text{Uniforme}}^{(0)}) \text{ e} \\ \sigma^2(X_{\text{Beta}}^{(0)}) &< \sigma^2(X_{\text{Uniforme}}^{(0)}), \end{aligned}$$

mas isso não significa que  $X_{\text{Beta}}^{(0)}$  domina  $X_{\text{Uniforme}}^{(0)}$  estocasticamente de segundo grau. Porém, o resultado significa que *perda esperada* não é coerente com a preferência por momentos. Do outro lado, isso não precisa implicar que *perda esperada* seja incoerente com o conceito de utilidade esperada. De fato, como analisaremos na seção 6.4.3, uma preferência por momentos não precisa ser incompatível com a existência de uma função de utilidade “razoável” (a ser definida).

Mas, por outro lado, a partir da nossa discussão sobre risco na seção 3.1, espera-se que a maior dispersão dos valores da distribuição uniforme seja punida de tal forma

que a distribuição beta seja preferida. Usando a medida consistente, seria preferida a distribuição uniforme à beta, pois ela tem valor maior. Esse resultado é causado pelo fato de que a maior concentração da beta cria, no primeiro passo, um valor esperado maior em termos absolutos que a uniforme. Subtraindo esse valor do valor esperado de  $X^{(0)}$ , obtém-se que  $v_{n=1} \left( X_{\text{Beta}}^{(0)} \right) < v_{n=1} \left( X_{\text{Uniforme}}^{(0)} \right)$ .

O dilema está causado pelo fato de avaliar o valor de  $X^{(0)}$  abaixo do *benchmark* em vez de considerar o desvio  $BM^{(0)} - X^{(0)}$ . Somente com a incorporação da distância entre o resultado não-desejado e o benchmark que se avalia corretamente o risco. Usaremos a forma do desvio  $(BM^{(i)} - X^{(i)})$  como valor de um ativo na próxima seção e veremos que a formulação ficará mais complicada. Mas, antes de entrar nesse estudo, analisamos ainda como se comporta a recursão no limite e como pode ser avaliado o uso de uma medida de risco alternativa.

### Convergência da recursão

A análise da diferença entre dois passos pode ser feita de forma bastante simples, pois

$$r(X^{(n)}) = E[X^{(0)} | X^{(0)} \leq BM^{(n)}] - r(X^{(n+1)}).$$

Assim, vale para todos  $k \geq 1$  e  $n \geq 1$

$$r(X^{(n)}) + (-1)^{k+1} r(X^{(n+k)}) = \sum_{j=0}^{k-1} (-1)^j E[X^{(0)} | X^{(0)} \leq BM^{(n+j)}].$$

Vale especialmente que  $\forall n \geq 0 \exists K_n$  tal que

$$r(X^{(n)}) + r(X^{(n+1)}) \leq K_n$$

e a seqüência  $\{v_n(X^{(0)})\}_{n=1,2,\dots}$  é de Cauchy.<sup>22</sup>

---

<sup>22</sup>Esse resultado segue também a partir dos fatos que

1. o espaço probabilístico com a distância de “quase certamente” é um espaço métrico,
2. a seqüência  $\{v_n(X^{(0)})\}_{n=1,2,\dots}$  converge,

e que num espaço métrico qualquer seqüência convergente é de Cauchy.

A velocidade de convergência é dada pelo seguinte raciocínio:

$$\begin{aligned}
& v_{n+1}(X^{(0)}) - v_n(X^{(0)}) \\
&= \sum_{i=0}^n (-1)^{i+1} E[X^{(0)} | X^{(0)} \leq BM^{(i)}] \\
&\quad - \sum_{i=0}^{n-1} (-1)^{i+1} E[X^{(0)} | X^{(0)} \leq BM^{(i)}] \\
&\quad + (-1)^n r(X^{(n+1)}) \\
&\quad + (-1)^{n-1} r(X^{(n)}) \\
&= (-1)^n E[X^{(0)} | X^{(0)} \leq BM^{(n)}] \\
&\quad + (-1)^n (r(X^{(n+1)}) + r(X^{(n)})).
\end{aligned}$$

Já que

$$r(X^{(n)}) = v(X^{(n+1)}) = E[X^{(n)} | X^{(n)} \leq BM^{(n)}] - r(X^{(n+1)}), \quad (6.16)$$

temos

$$\begin{aligned}
& v_{n+1}(X^{(0)}) - v_n(X^{(0)}) \\
&= (-1)^n E[X^{(0)} | X^{(0)} \leq BM^{(n)}] \\
&\quad + (-1)^n (r(X^{(n+1)}) + E[X^{(n)} | X^{(n)} \leq BM^{(n)}] - r(X^{(n+1)})) \\
&= (-1)^n \cdot 2 \cdot E[X^{(0)} | X^{(0)} \leq BM^{(n)}]
\end{aligned}$$

e

$$|v_{n+1}(X^{(0)}) - v_n(X^{(0)})| = 2 \cdot E[X^{(0)} | X^{(0)} \leq BM^{(n)}].$$

Pode ser interessante calcular o tamanho do erro restante após a interrupção da recursão ao nível  $n$ :  $r(X^{(n)})$ . Mas, como  $r(X^{(n)}) = v(X^{(n+1)})$ , esse valor já foi mencionado na discussão em seguida da definição 6.5 e, de forma geral, temos

$$\begin{aligned}
r(X^{(n)}) \in [E[X^{(n)} | X^{(n)} \leq BM^{(n)}] - E[X^{(n+1)} | X^{(n+1)} \leq BM^{(n+1)}], \\
E[X^{(n)} | X^{(n)} \leq BM^{(n)}].
\end{aligned}$$

**Exemplo 6.3** (Precificação consistente para uma Beta em  $(0, 1)$ , cont.) *No lema 6.11, interrompemos a recursão no nível  $n = 5$  o que implica um erro de  $r(X^{(5)})$ . Se*



interrompemos no nível  $n = 10$ , chegamos à diferença

$$\begin{aligned}
 & r(X^{(5)}) - r(X^{(10)}) \\
 = & \int_0^{E[X^{(5)}]} x f_{\text{Beta}(4,4)}(x | x \leq E[X^{(5)}]) dx \\
 & - \int_0^{E[X^{(6)}]} x f_{\text{Beta}(4,4)}(x | x \leq E[X^{(6)}]) dx \\
 & + \int_0^{E[X^{(7)}]} x f_{\text{Beta}(4,4)}(x | x \leq E[X^{(7)}]) dx \\
 & - \int_0^{E[X^{(8)}]} x f_{\text{Beta}(4,4)}(x | x \leq E[X^{(8)}]) dx \\
 & + \int_0^{E[X^{(9)}]} x f_{\text{Beta}(4,4)}(x | x \leq E[X^{(9)}]) dx \\
 \approx & 0,094819.
 \end{aligned}$$

Percebe-se que a diferença entre o quinto e o décimo passo é ainda bastante grande. Se cortássemos a recursão após o passo  $n = 9$ , chegaríamos a uma distância menor:

$$r(X^{(5)}) - r(X^{(9)}) \approx 0,0440512.$$

Parece que o erro está aumentando, mas o que acontece é que a cada dois passos, um valor é adicionado ao valor anterior e um outro é subtraído. Portanto, quando comparamos dois cortes no momento de diminuir o valor, obtemos uma diferença maior de que no caso do passo que soma um valor. De preferência, deveria ter usado um corte com  $n$  par. É importante analisar as diferenças entre os passos para saber se o valor final sofre ainda alterações grandes, e para isso, deve-se usar a fórmula (6.16).

### Relação da precificação consistente com outras abordagens

Quando analisamos a precificação por consistência usando um corte na recursão, desprezamos uma parte do valor por causa da interrupção da recursão. Além de calcular o grau dessa negligência de uma medida de risco consistente, pode-se pensar também em analisar a diferença com uma abordagem não consistente. No caso do uso do desvio-padrão como medida de risco, define-se o grau de inconsistência como a diferença dos valores atribuídos ao risco de  $X^{(0)}$ .

**Definição 6.6 (Grau de inconsistência)** *O grau de inconsistência de uma medida de*

risco  $R$  para a variável  $X^{(0)}$  é dado por

$$z_R(X^{(0)}) := \left| \frac{R(X^{(0)}) - r(X^{(0)})}{r(X^{(0)})} \right|.$$

No caso de  $R$  consistente, mas com um grau  $n$  de indiferença ao risco, falamos de nível de risco negligenciado em vez de nível de inconsistência. Além da medida  $ES$  (e suas equivalências), todas as outras medidas citadas na seção 5.2 são inconsistentes. No caso da *perda esperada*, temos que

$$ES(X^{(0)}) - r(X^{(0)}) = r(X^{(1)})$$

e esse valor é dado por (6.8) sem o primeiro termo:

$$\begin{aligned} r(X^{(1)}) &= \sum_{i=0}^{n-1} \frac{(-1)^{i+1}}{F_{X^{(0)}}(BM^{(0)})} \int_a^{BM^{(n-1)}} x f(x) dx \\ &+ \sum_{j=1}^{n-1} \sum_{i=0}^j \frac{(-1)^{i+1}}{F_{X^{(i)}}(BM^{(i)})} \int_{BM^{(i)}}^{BM^{(j-1)}} x f(x) dx. \end{aligned}$$

A diferença está em começar a somar as fatias somente a partir do índice  $j = 2$ , ou seja, a negligência de risco pelo *perda esperada* é dada pela soma das fatias da distribuição até o benchmark  $BM^{(2)}$ .

**Exemplo 6.4** (Nível de inconsistência de  $\sigma$  no caso da Beta(4,4) em (0, 1)) *O desvio-padrão de  $X^{(0)} \sim \text{Beta}(4, 4)$  é dado por*

$$\sigma(X^{(0)}) = 0,1\bar{6}.$$

*O valor de  $X^{(0)}$  usando  $\sigma$  para avaliar a incerteza de  $X^{(0)}$  seria*

$$v_\sigma(X^{(0)}) = \frac{1}{2} - 0,1\bar{6} \approx 0,3\bar{3}$$

*Em comparação, temos*

$$v_{n=14}(X^{(0)}) \approx 0,299$$

*e o grau de inconsistência é aproximadamente*

$$\begin{aligned} z_\sigma(X^{(0)}) &\approx \left| \frac{0,1\bar{6} - 0,201}{0,201} \right| \\ &\approx 0,1687. \end{aligned}$$

O valor atribuído a uma distribuição uniforme usando a medida desvio-padrão de quase 0,17 é menor que o valor atribuído à distribuição beta. Esse resultado está em coerência com a intuição de maior concentração da beta em torno do valor esperado e da preferência por momentos. Observamos que a igualdade dos valores de uma distribuição beta usando desvio-padrão e o de uma uniforme usando consistência é uma coincidência numérica e não oferecemos interpretação econômica para esse acaso.

### 6.3.3 Risco como valor do desvio do *benchmark*

Mencionamos que o risco de um ativo não pode ser visto simplesmente como o valor das realizações não-desejadas. Mas, deve-se avaliar o valor do desvio do seu *benchmark*. Essa idéia reflete-se nas abordagens de *downside risk* e *expected regret*. No nosso caso, definimos o valor de uma variável  $X^{(i)}$  como<sup>23</sup>

$$X^{(i)} := (BM^{(i-1)} - X^{(i-1)}) \cdot \mathbf{1}_{\{X^{(i-1)} \leq BM^{(i-1)}\}} \quad (6.17)$$

e consideramos, em analogia à *perda esperada*, a sua distribuição na avaliação do seu valor esperado. Observamos aqui uma diferença sutil entre as abordagens *perda esperada* (*ES*) e *expected regret* (*ER*) causada pela normalização do *ES* :

$$\begin{aligned} ES_{BM=0}(X) &= E[X | X \leq 0] = \frac{E[X \cdot \mathbf{1}_{\{X \leq 0\}}]}{\mathbb{P}(X \leq 0)} \\ &\neq E[X \cdot \mathbf{1}_{\{X \leq 0\}}] = ER_0(X). \end{aligned}$$

Isso significa que, no *ES*, usa-se o valor esperado da nova variável  $X^{(i)}$  calculado pela sua distribuição, enquanto o *ER* calcula o valor esperado da variável  $X^{(i)}$  com a distribuição de  $X^{(i-1)}$ . Dessa forma, o risco da variável  $X^{(i-1)}$  é avaliado como se nela tivesse implícita uma opção de venda (*put*):

$$v_{ER}(X^{(i-1)}) = E_{X^{(i-1)}}[X^{(i-1)}] - E_{X^{(i-1)}}[(BM^{(i-1)} - X^{(i-1)})^+],$$

sendo que

$$P(X^{(i-1)}) := E_{X^{(i-1)}}[(BM^{(i-1)} - X^{(i-1)})^+]$$

<sup>23</sup>Quando o risco é visto como desvio do *benchmark*, não precisamos nos preocupar com o sinal da variável, pois consideramos o valor em relação a um valor de referência e a diferença sempre será não positivo.

é o valor esperado de um pagamento de uma opção de venda com preço de exercício de  $BM^{(i-1)}$ . No caso de  $ES$ , temos

$$v_{ES}(X^{(i-1)}) = E_{X^{(i-1)}}[X^{(i-1)}] - E_{X^{(i)}}[(BM^{(i-1)} - X^{(i-1)})^+]$$

e o segundo termo não tem interpretação de um valor esperado de uma opção de venda.

Problemático no uso da definição (6.17) é o caso  $i = 0$ . Mesmo que pressupuséssemos que  $BM^{(-1)} = \infty$ , não conseguiríamos uma formulação correta. Tendo o problema de definir consistentemente as variáveis  $X^{(i)}$ , pode se questionar o uso da palavra *consistente*. Porém, como o caso  $i = 0$  é a única exceção, e a consistência está refletida no uso da mesma forma de avaliação, i.e., ainda vale que

$$v(X) = E[X] - r(X).$$

Mantemos a nomenclatura e chamamos a precificação de *consistente*.

### A formulação recursiva do valor consistente

Aplicando o raciocínio da consistência uma vez, segue que

$$\begin{aligned} v^{DBM}(X^{(0)}) &= E[X^{(0)}] - r(X^{(0)}) \\ &= E[X^{(0)}] - E_{X^{(0)}}[(BM^{(0)} - X^{(0)}) | X^{(0)} \leq BM^{(0)}] + r(X^{(1)}). \end{aligned}$$

Expandindo, por exemplo, até  $n = 3$ , obtemos<sup>24</sup>

$$\begin{aligned} &v_{n=3}^{DBM}(X^{(0)}) \\ &= E[X^{(0)}] - (E[X^{(1)}] - E[X^{(2)}] + E[X^{(3)}]) \\ &= E[X^{(0)}] \\ &\quad - E_{X^{(0)}}[(B^{(0)} - X^{(0)}) | X^{(0)} \leq BM^{(0)}] \\ &\quad + E_{X^{(1)}}[(B^{(1)} - X^{(1)}) | X^{(1)} \leq BM^{(1)}] \\ &\quad - E_{X^{(2)}}[(B^{(2)} - X^{(2)}) | X^{(2)} \leq BM^{(2)}] \\ &= E[X^{(0)}] \\ &\quad + E_{X^{(0)}}[X^{(0)} | X^{(0)} \leq BM^{(0)}] - B^{(0)} \\ &\quad - E_{X^{(1)}}[X^{(1)} | X^{(1)} \leq BM^{(1)}] + B^{(1)} \\ &\quad + E_{X^{(2)}}[X^{(2)} | X^{(2)} \leq BM^{(2)}] - B^{(2)}. \end{aligned}$$

<sup>24</sup>Sob a hipótese que  $BM^{(i)}$ ,  $i = 0, 1, 2, \dots$  sejam constantes.

Continuando a expansão para o termo  $E[X^{(1)} | X^{(1)} \leq BM^{(1)}]$ , obtemos

$$\begin{aligned} & E_{X^{(1)}} [X^{(1)} | X^{(1)} \leq BM^{(1)}] \\ &= E_{X^{(0)}} [(B^{(0)} - X^{(0)}) | X^{(1)} \leq BM^{(1)}, X^{(0)} \leq BM^{(0)}] \\ &= B^{(0)} - E_{X^{(0)}} [X^{(0)} | X^{(1)} \leq BM^{(1)}, X^{(0)} \leq BM^{(0)}]. \end{aligned}$$

Agora, levamos em conta que<sup>25</sup>

$$\begin{aligned} & \{\omega \in \Omega : X^{(0)}(\omega) \leq BM^{(0)}\} \cap \{\omega \in \Omega : X^{(1)}(\omega) \leq BM^{(1)}\} \\ &= \{\omega \in \Omega : X^{(0)}(\omega) \leq BM^{(0)}\} \\ & \quad \cap \left\{ \omega \in \Omega : (BM^{(0)} - X^{(0)}(\omega)) \cdot \mathbf{1}_{\{X^{(0)}(\omega) \leq BM^{(0)}\}} \leq BM^{(1)} \right\} \\ &= \{\omega \in \Omega : X^{(0)}(\omega) \leq BM^{(0)}, X^{(0)}(\omega) \geq BM^{(0)} - BM^{(1)}\}, \end{aligned}$$

o que significa que

$$\begin{aligned} & E_{X^{(1)}} [X^{(1)} | X^{(1)} \leq BM^{(1)}] \\ &= B^{(0)} - E[X^{(0)} | BM^{(0)} - BM^{(1)} \leq X^{(0)} \leq BM^{(0)}]. \end{aligned}$$

Assim, obtemos

$$\begin{aligned} & v_{n=3}^{\text{DBM}}(X^{(0)}) \\ &= E[X^{(0)}] \\ & \quad + E[X^{(0)} | X^{(0)} \leq BM^{(0)}] - B^{(0)} \\ & \quad + E[X^{(0)} | BM^{(0)} - BM^{(1)} \leq X^{(0)} \leq BM^{(0)}] - (B^{(0)} - B^{(1)}) \\ & \quad + E[X^{(0)} | BM^{(0)} - BM^{(1)} \leq X^{(0)} \leq BM^{(0)} - BM^{(1)} + BM^{(2)}] \\ & \quad - (B^{(0)} - B^{(1)} + B^{(2)}). \end{aligned}$$

Aplicando essa recursão para um  $n \geq 1$  qualquer, chegamos ao seguinte resultado:

**Lema 6.12** *O valor consistente com risco como valor do desvio do benchmark é dado por:*

$$\begin{aligned} v^{\text{DBM}}(X^{(0)}) &= E[X^{(0)}] \\ & \quad + \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} \left( E[X^{(0)} | lb(i) \leq X^{(0)} \leq ub(i)] \right. \\ & \quad \left. + \sum_{j=0}^i (-1)^{j+1} BM^{(j)} \right) \end{aligned}$$

<sup>25</sup>O lema 6.13 generaliza o seguinte passo.

com

$$lb(i) := \sum_{j=1}^{\lfloor \frac{i+1}{2} \rfloor} (BM^{(2(j-1))} - BM^{(2j-1)}),$$

$$ub(i) := BM^{(0)} + \sum_{j=1}^{\lfloor i/2 \rfloor} (BM^{(2j)} - BM^{(2j-1)}).$$

Observamos algumas diferenças marcantes em relação à formulação sem o desvio do *benchmark*:

1. O termo envolvendo o valor esperado somente soma a variável  $X^{(i)} \cdot \mathbf{1}_{\{X^{(i)} \leq BM^{(i)}\}}$ , i.e., não existe alternância.
2. Do valor esperado condicional de  $X^{(0)}$  é subtraído alternadamente o limite inferior e superior. Assim, em cada dois passos, é subtraída uma pequena parcela do valor inicial.

Esse resultado usa o seguinte lema que garante que o termo de valor esperado de  $X^{(i)}$  pode ser reformulado em termos de  $X^{(0)}$ . Este é um resultado que caracteriza como a parte relevante de  $X^{(0)}$  é cercada pela seqüência dos *benchmarks*.

**Lema 6.13** Para qualquer  $i \geq 1$ , vale que

$$\bigcap_{j=0}^i \{\omega \in \Omega : X^{(j)} \leq BM^{(j)}\}$$

$$= \{\omega \in \Omega : lb(i) \leq X^{(0)} \leq ub(i)\}.$$

**Prova.** Para qualquer  $j$ , temos

$$\begin{aligned} & \{X^{(j)} \leq BM^{(j)}\} \cap \{X^{(j+1)} \leq BM^{(j+1)}\} \cap \{X^{(j+2)} \leq BM^{(j+2)}\} \\ &= \{X^{(j)} \leq BM^{(j)}\} \\ & \cap \left\{ (BM^{(j)} - X^{(j)}) \cdot \mathbf{1}_{\{(BM^{(j)} - X^{(j)})\}} \leq BM^{(j+1)} \right\} \\ & \cap \left\{ (BM^{(j+1)} - X^{(j+1)}) \cdot \mathbf{1}_{\{X^{(j+1)} \leq BM^{(j+1)}\}} \leq BM^{(j+2)} \right\} \\ &= \{X^{(j)} \leq BM^{(j)}\} \\ & \cap \{-X^{(j)} \leq BM^{(j+1)} - BM^{(j)}\} \\ & \cap \{-X^{(j+1)} \leq BM^{(j+2)} - BM^{(j+1)}\} \\ &= \{X^{(j)} \leq BM^{(j)}\} \\ & \cap \{X^{(j)} \geq BM^{(j)} - BM^{(j+1)}\} \\ & \cap \{X^{(j)} \leq BM^{(j+2)} - BM^{(j+1)} - BM^{(j)}\}, \end{aligned}$$

que é equivalente ao conjunto

$$\{\omega \in \Omega : BM^{(j)} - BM^{(j+1)} \leq X^{(j)} \leq BM^{(j)} + (BM^{(j+2)} - BM^{(j+1)})\}.$$

Aplicando essa recursão para  $n \geq 1$ , começando em  $j = 0$ , obtemos o resultado desejado.

■

Se a medida de risco *expected regret* ( $ER$ ) fosse definida em analogia à medida *perda esperada* ( $ES$ ) de forma que

$$ER_{\text{BM}}(X) := E[BM - X | X \leq BM],$$

teríamos a seguinte expressão para a avaliação consistente com risco como desvio do *benchmark*:

$$V_n(X^{(0)}) = E[X^{(0)}] + ER_{\text{BM}}(X^{(0)}) - BM^{(0)} + S_{n-1}(X^{(0)}),$$

sendo que  $S_{n-1}(X^{(0)})$  representa a soma dos termos de maiores ordens na formulação recursiva,

$$S_{n-1}(X^{(0)}) := \sum_{i=1}^{n-1} \left\{ E[X^{(0)} | lb(i) \leq X^{(0)} \leq ub(i)] + \sum_{j=0}^i (-1)^{j+1} BM^{(j)} \right\}.$$

Finalmente, podemos também escrever o último termo de forma mais compacta:

$$\begin{aligned} \sum_{i=0}^{n-1} \sum_{j=0}^i (-1)^{i+1} BM^{(i)} &= -n \cdot BM^{(0)} \\ &+ (n-1) \cdot BM^{(1)} \\ &- (n-2) \cdot BM^{(2)} \\ &+ \dots \\ &= \sum_{i=0}^n (n-i) \cdot BM^{(i)}. \end{aligned}$$

### A escolha entre uma distribuição uniforme e uma beta

Por causa da simetria da distribuição uniforme em torno de qualquer um dos *benchmarks*, não obtemos valores diferentes para as duas formas de precificação coerente. Segue que

$$v^{\text{DBM}}(X_{\text{Uniforme}}^{(0)}) = v(X_{\text{Uniforme}}^{(0)}) = \frac{1}{3}.$$

A distribuição, por sua vez, somente é simétrica em torno do seu valor esperado, mas não em relação aos *benchmarks*  $BM^{(i)}$ ,  $i \geq 1$ . Calcula-se o valor consistente para a *Beta* (4, 4)

usando os seguintes passos intermediários

$$E[X^{(1)}] \approx 0,1367$$

$$E[X^{(2)}] \approx 0,0710$$

$$E[X^{(3)}] \approx 0,0344$$

$$E[X^{(4)}] \approx 0,0174$$

$$E[X^{(5)}] \approx 0,0086.$$

Finalmente, obtemos que

$$\begin{aligned} & v_{n=5}^{\text{DBM}}(X_{\text{Beta}}^{(0)}) \\ &= E[X^{(0)}] - E[X^{(1)}] + E[X^{(2)}] - E[X^{(3)}] + E[X^{(4)}] - E[X^{(5)}] \\ &\approx 0,41 \geq 0, \bar{3} = v^{\text{DBM}}(X_{\text{Uniforme}}^{(0)}) \end{aligned}$$

Essa forma de precificação converge bem mais rapidamente de que a alternativa sem consideração do desvio do *benchmark*. A figura (6.3) visualiza os passos do valor em relação aos graus de indiferença ao risco.

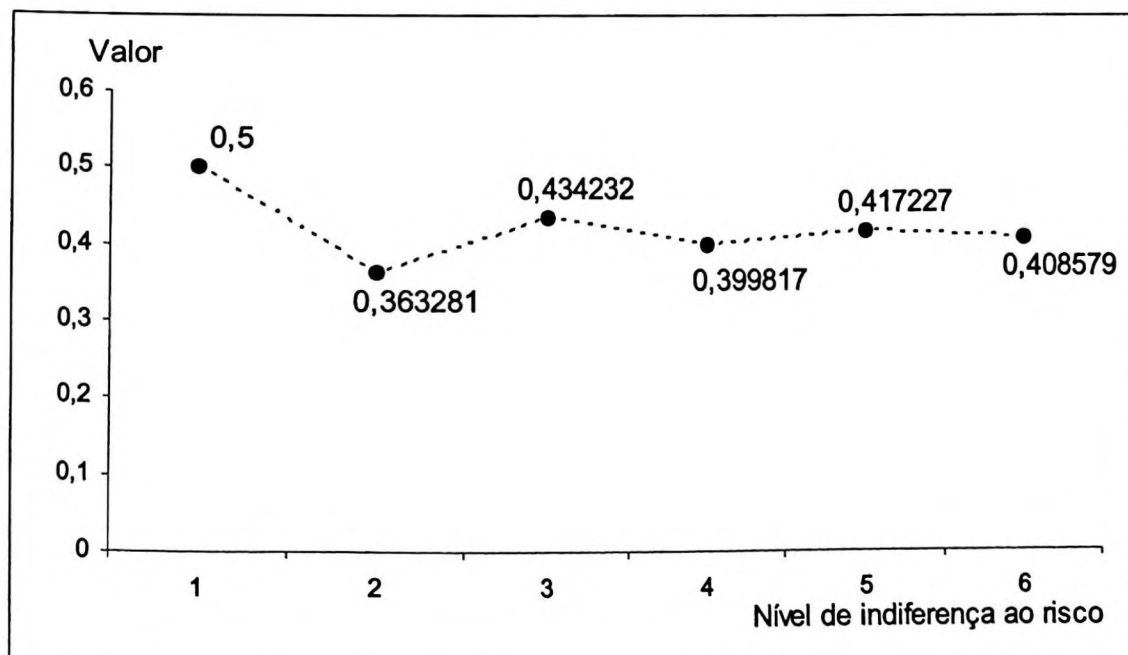


Figura 6.3: Os preços intermediários na determinação do preço consistente com desvio do *benchmark*.



### 6.3.4 Aplicação da consistência na precificação em mercados incompletos

#### Precificação de ativos básicos

Com a definição<sup>26</sup> de  $S_{t_0}^{(0)} := S_{t_0}$ , o valor em  $t_0$  do ativo  $S$  é calculado, pressupondo que  $BM^{(i)} = E_{\mathbf{P}_{S_{t_1}^{(i)}}} [S_{t_1}^{(i)}]$  da seguinte forma:<sup>27</sup>

$$\begin{aligned}
 v(S_{t_1}^{(0)}) &= E_{\mathbf{P}_{S_{t_1}^{(0)}}} [S_{t_1}^{(0)}] \\
 &\quad - E_{\mathbf{P}_{S_{t_1}^{(1)}}} [S_{t_1}^{(0)} \cdot 1_{\{S_{t_1}^{(0)} \leq E[S_{t_1}^{(0)}]\}}] \\
 &\quad + E_{\mathbf{P}_{S_{t_1}^{(2)}}} [S_{t_1}^{(1)} \cdot 1_{\{S_{t_1}^{(1)} \leq E[S_{t_1}^{(1)}]\}}] \\
 &= \frac{1}{4} \cdot 120 + \frac{1}{2} \cdot 100 + \frac{1}{4} \cdot 90 \\
 &\quad - \left( \frac{2}{3} \cdot 100 + \frac{1}{3} \cdot 90 \right) \\
 &\quad + (1 \cdot 90) \\
 &= 102,5 - 96,67 + 90 \\
 &= 95,83 \neq 100 = S_{t_0}^{(0)}.
 \end{aligned}$$

Somente no caso que a utilidade do investidor seja igual à do mercado, o preço do mercado coincide com o valor que o investidor atribui ao ativo  $S$  em  $t_0$ . No nosso exemplo é o caso que a fórmula de precificação do mercado não é a precificação consistente.

Para calcular o valor do derivativo em  $t_0$  sem ter uma carteira de proteção, podemos prosseguir da mesma forma e obteremos  $v(D_{t_0}) = 5$ , pois os valores dos estados não-desejáveis em relação ao *benchmark*  $BM^{(0)} = E_{\mathbf{P}_{D_{t_1}^{(0)}}} [D_{t_1}] = 5$  são nulos. Esse resultado é equivalente à decisão do investidor de atribuir o preço de Arrow-Debreu de  $\frac{1}{2}$  ao estado  $\omega_3$ .

#### Precificação de derivativos com o valor consistente sem o desvio do *benchmark*

Quando se desconsidera a possibilidade de criar uma carteira de proteção, o derivativo é tratado como um ativo básico sem dependência dos outros ativos básicos.<sup>28</sup> Mas, no

<sup>26</sup>Observa que a superposição (i) indica o passo da recursão e não o número do ativo como na seção 2.2.

<sup>27</sup>Vide equação 6.6.

<sup>28</sup>Segundo a notação de FRITELLI (2000b), o mercado seria totalmente incompleto.

nosso caso,  $D$  é função do ativo subjacente  $S$ , que é negociável no mercado. <sup>29</sup> Para escolher um preço  $x$  na faixa dos preços livres de arbitragem  $[0, \frac{20}{3}]$  temos que aplicar o nosso critério de decisão de forma que:<sup>30</sup>

- o valor esperado da carteira composta por ativos subjacentes mais o derivativo tem que ser positivo
- usamos  $r(X)$  como fórmula para avaliar o risco do ativo  $X$ .

Seguimos os passos típicos para determinar o preço do derivativo num mercado incompleto:<sup>31</sup>

1. Determinam-se as estratégias de mínimo risco para o comprador,  $\pi_{\text{Buy}}$ , e para o vendedor,  $\pi_{\text{Venda}}$ .
2. Calcula-se o preço do derivativo que permite que os valores esperados das carteiras do comprador e do vendedor sejam positivos.

Usamos primeiro o grau de consistência de 1, o que é equivalente ao uso da medida de risco *perda esperada* no caso que não consideramos o desvio do *benchmark*. Sabendo que o preço superior (preço de replicação do vendedor) é de  $\frac{20}{3} < 10$ , observamos que, para um  $\varepsilon > 0$  e com a definição 3.2,

$$V_t^{x, \pi, C, -D} := X_t^{x, \pi, C} - D_t,$$

temos

$$\begin{aligned} E[V_{\text{Venda}}] &:= E_{\mathbb{P}_{V_{t_1}}} \left[ V_{t_1}^{10-\varepsilon, \pi_{\text{Venda}}, 0, -D_{t_1}} \right] \\ &= \frac{1}{4} (10 - \varepsilon + 20\pi_{\text{Venda}} - 20) + \frac{1}{2} (10 - \varepsilon) \\ &\quad + \frac{1}{4} (10 - \varepsilon - 10\pi_{\text{Venda}}) \\ &= 5 - \varepsilon + \frac{5}{2} \pi_{\text{Venda}}, \end{aligned}$$

<sup>29</sup>Segundo FRITELLI (2000b), esse mercado seria apenas parcialmente incompleto.

<sup>30</sup>Pode-se pensar em não usar uma abordagem “esperança-risco”, mas em otimizar a função de “utilidade”  $v(X) := E[X] - r(X)$ . Porém, neste caso, precisamos determinar o parâmetro de aversão ao risco para poder escolher entre as diversas curvas de nível.

<sup>31</sup>Vide, por exemplo, SCHWEIZER (1992); SCHULMERICH; TRAUTMANN (2001) ou SOMMER (1998).

e

$$\begin{aligned}
E[V_{\text{Compra}}] &:= E_{\mathbb{P}_{V_{t_1}}} [V_{t_1}^{-10+\varepsilon, \pi_{\text{Compra}}, 0, D_{t_1}}] \\
&= \frac{1}{4}(- (10 - \varepsilon) + 20\pi_{\text{Compra}} + 20) + \frac{1}{2}(- (10 - \varepsilon)) \\
&\quad + \frac{1}{4}(- (10 - \varepsilon) - 10\pi_{\text{Compra}}) \\
&= -5 + \varepsilon + \frac{5}{2}\pi_{\text{Compra}}.
\end{aligned}$$

Como é na natureza do problema que existem estados nos quais o valor da carteira composta por ativos e derivativos é negativo, temos que ajustar um pouco a notação que usamos na seção 6.3.2, pois tínhamos pressuposto que  $V_t^{x, \pi, C, -D} \geq 0$ , q.c. Somos obrigados a introduzir o valor absoluto na fórmula do valor consistente sem desvio do *benchmark*, pois  $BM^{(0)} = 0$ :

$$\begin{aligned}
v_{n=1}(X^{(0)}) &= E[X^{(0)}] - \text{abs}(E[X^{(0)} | X^{(0)} \leq BM^{(0)}]) \\
&= E[X^{(0)}] + E[X^{(0)} | X^{(0)} \leq BM^{(0)}].
\end{aligned}$$

Com a notação  $x^- := (-1) \cdot x \cdot \mathbf{1}_{\{x \leq 0\}}$ , podemos escrever a perda esperada do vendedor como:

$$\begin{aligned}
&\mathbb{P}(V_{\text{Venda}} \leq 0) \cdot ES_{BM=0}[V_{\text{Venda}}] \\
&= - \mathbb{P}(V_{\text{Venda}} \leq 0) \cdot E[V_{\text{Venda}} | V_{\text{Venda}} \leq 0] \\
&= \frac{1}{4}(10 - \varepsilon + 20\pi_{\text{Venda}} - 20)^- + \frac{1}{2}(10 - \varepsilon)^- \\
&\quad + \frac{1}{4}(10 - \varepsilon - 10\pi_{\text{Venda}})^-.
\end{aligned}$$

No caso de

$$1 - \frac{\varepsilon}{10} \geq \pi_{\text{Venda}} \geq \frac{1}{2} + \frac{\varepsilon}{20},$$

temos que  $E[V_{\text{Venda}} | V_{\text{Venda}} \leq 0] = 0$ . Isso somente acontece se  $\varepsilon \leq \frac{10}{3}$ , ou seja, o preço do derivativo está acima do (ou igual ao) preço de super-replicação do vendedor. Por outro lado, se  $\varepsilon > \frac{10}{3}$ , o vendedor não consegue montar uma carteira de proteção que evite perdas em todos os estados: se ele escolher

$$\pi_{\text{Venda}} = 1 - \frac{\varepsilon}{10},$$

ele se protege contra perdas no estado  $\omega_3$ , mas terá uma perda esperada de

$$\begin{aligned}
-E[V_{\text{Venda}} | V_{\text{Venda}} \leq 0] &= \left(10 - \varepsilon + 20 \left(1 - \frac{\varepsilon}{10}\right) - 20\right)^- \\
&= (10 - 3\varepsilon)^-
\end{aligned}$$

que é positiva no caso de  $\varepsilon > \frac{10}{3}$  e devido ao resultado no estado  $\omega_1$ . Se ele decide aplicar

$$\pi_{\text{Venda}} = \frac{1}{2} + \frac{\varepsilon}{20} \quad (6.18)$$

no ativo  $S^{(1)}$ , ele evita perdas no estado  $\omega_1$ , mas no estado  $\omega_3$ , ele sofrerá uma perda de  $5 - \frac{3}{2}\varepsilon$  que também é positiva para  $\varepsilon > \frac{10}{3}$ .

Portanto, no caso  $\varepsilon > \frac{10}{3}$  a perda nos dois estados  $\omega_1$  e  $\omega_3$  seria igual e podemos escolher a estratégia  $\tilde{\pi}_{\text{Venda}} = \frac{1}{2} + \frac{\varepsilon}{20}$ . Com essa estratégia determinamos o valor esperado da carteira como

$$\begin{aligned} E[V_{\text{Venda}}] &= 0 + \frac{1}{2}(10 - \varepsilon) + \frac{1}{4}\left(10 - \varepsilon - 10\left(\frac{1}{2} + \frac{\varepsilon}{20}\right)\right) \\ &= \frac{1}{4}\left(20 - 2\varepsilon + 10 - \varepsilon - 5 - \frac{\varepsilon}{2}\right) \\ &= \frac{1}{4}\left(25 - \frac{7}{2}\varepsilon\right), \end{aligned}$$

que é não-negativo no caso de

$$\varepsilon \leq \frac{50}{7}.$$

Isso implica um preço mínimo da opção exigido pelo vendedor de

$$x_{\text{Venda}} = 10 - \varepsilon \geq \frac{20}{7}.$$

Do ponto de vista do comprador determina-se o preço máximo de compra de forma equivalente. Primeiro, achamos a estratégia  $\pi_{\text{Compra}}^*$  que é o argumento ótimo do seguinte problema:

$$\min_{\pi_{\text{Compra}}} -E\left[V_{t_1}^{-10+\varepsilon, \pi_{\text{Compra}}, 0, D_{t_1}} \mid V_{t_1}^{-10+\varepsilon, \pi_{\text{Compra}}, 0, D_{t_1}} \leq 0\right].$$

Como

$$\begin{aligned} &\mathbb{P}(V_{\text{Compra}} \leq 0) \cdot ES_{BM=0}[V_{\text{Compra}}] \\ &= -\mathbb{P}(V_{\text{Compra}} \leq 0) \cdot E[V_{\text{Compra}} \mid V_{\text{Compra}} \leq 0] \\ &= \frac{1}{4}(-10 + \varepsilon + 20\pi_{\text{Compra}} + 20)^- + \frac{1}{2}(-10 + \varepsilon)^- \\ &\quad + \frac{1}{4}(-10 + \varepsilon - 10\pi_{\text{Compra}})^-. \end{aligned}$$

No segundo estado,  $\omega_2$ , o comprador sofre uma perda sem possibilidades de proteção.<sup>32</sup>

Porém, nos dois outros estados vale que

$$\begin{aligned} -10 + \varepsilon + 20\pi_{\text{Compra}} + 20 \leq 0 &\Leftrightarrow \pi_{\text{Compra}} \leq -\frac{1}{2} - \frac{\varepsilon}{20}, \\ -10 + \varepsilon - 10\pi_{\text{Compra}} \leq 0 &\Leftrightarrow \pi_{\text{Compra}} \geq -1 + \frac{\varepsilon}{10}. \end{aligned}$$

<sup>32</sup>Exceto no caso que o valor da opção fosse 0, i.e.,  $\varepsilon = 10$ .

No caso de  $\varepsilon \geq \frac{10}{3}$ , vale que  $-1 + \frac{\varepsilon}{10} \geq -\frac{1}{2} - \frac{\varepsilon}{20}$  e, portanto, com a escolha de

$$\tilde{\pi}_{\text{Compra}} = -\frac{1}{2} - \frac{\varepsilon}{20}$$

evita-se qualquer perda nos estados  $\omega_1$  e  $\omega_3$ . Com essa estratégia calcula-se o valor esperado da carteira do comprador como

$$\begin{aligned} E[V_{\text{Compra}}] &= -5 + \varepsilon + \frac{5}{2}\pi_{\text{Compra}} \\ &= -5 + \varepsilon + \frac{5}{2}\left(-\frac{1}{2} - \frac{\varepsilon}{20}\right) \\ &= -\frac{25}{4} - \frac{35}{40}\varepsilon \end{aligned}$$

que é não-negativo no caso de

$$\varepsilon \geq \frac{50}{7}$$

o que implica que o preço máximo de compra aceito pelo comprador é de

$$x_{\text{Compra}} = 10 - \varepsilon \leq \frac{20}{7}.$$

O preço de negócio que se estabelecerá no mercado se tanto o comprador como o vendedor usarem a precificação consistente é de

$$x = D_{t_0} = \frac{20}{7} \approx 2,86.$$

Divido à possibilidade de montar uma carteira que protege (ao menos parcialmente) o investidor de flutuações dos preços de  $S$ , é possível diminuir o preço do derivativo em relação a um mercado totalmente incompleto.

Sabendo que se estabelece no mercado  $\varepsilon = \frac{50}{7}$ , obtemos, a partir da equação (6.18) que o vendedor do derivativo aplica a quantidade de

$$\pi_{\text{Venda}} = \frac{1}{2} + \frac{5}{14} \approx 0,86$$

no ativo com incerteza. Essa quantia é maior que a aplicação quando o investidor segue uma utilidade exponencial negativa (vide equação (4.5) na seção 4.2.1).

No caso da *ERD* da seção 6.2 obtivemos como estratégia ótima do vendedor do derivativo a aplicação de  $\pi_{\text{Venda}}^{\text{ERD}} = \frac{14}{19}$ . Isso resulta num valor esperado não-negativo da carteira no caso de  $\text{el}eq \frac{130}{19} \approx 6,84$ . Portanto, o preço de mercado seria de  $x^{\text{ERD}} \approx 3,16$ .

### Precificação de derivativos com o valor consistente com o desvio do *benchmark*

Quando consideramos o desvio do *benchmark*, usando a fórmula do valor consistente  $v^{\text{DBM}}$  do lema 6.12, temos o mesmo resultado que no caso sem *benchmark* para o grau de consistência  $n = 1$ .<sup>33</sup>

$$\begin{aligned} v^{\text{DBM}}(X^{(0)}) &= E[X^{(0)}] + E[X^{(0)} | X^{(0)} \leq BM^{(0)}] - BM^{(0)} \\ &= E[X^{(0)}] + E[X^{(0)} | X^{(0)} \leq BM^{(0)}] \\ &= v(X^{(0)}). \end{aligned}$$

Porém, se escolhemos  $n = 2$ , a função-objetivo já é diferente:

$$\begin{aligned} r_{n=2}^{\text{DBM}}(X^{(0)}) &= E[X^{(0)} | X^{(0)} \leq BM^{(0)}] - BM^{(0)} \\ &\quad + E[X^{(0)} | BM^{(0)} - BM^{(1)} \leq X^{(0)} \leq BM^{(0)}] \\ &\quad - (BM^{(0)} - BM^{(1)}). \end{aligned}$$

Sabemos, além de  $BM^{(0)} = 0$ , que

$$BM^{(1)} = E[X^{(1)}] = 0 - E[X^{(0)} | X^{(0)} \leq BM^{(0)}],$$

o que implica que

$$\begin{aligned} r_{n=2}^{\text{DBM}}(X^{(0)}) &= E[X^{(0)} | E[X^{(0)} | X^{(0)} \leq 0] \leq X^{(0)} \leq 0] \\ &= E[X^{(0)} | -ES(X^{(0)}) \leq X^{(0)} \leq 0]. \end{aligned}$$

Podemos agora seguir os passos da última seção, começando com a otimização do *expected regret* do vendedor do derivativo:

$$r_{n=2}^{\text{DBM}}(V_{\text{Venda}}) := E_{\mathbb{P}_{V_{t_1}}} \left[ V_{t_1}^{10-\varepsilon, \pi_{\text{Venda}, 0, -D_{t_1}}} \left| 5 - \frac{3}{2}\varepsilon \leq V_{t_1}^{10-\varepsilon, \pi_{\text{Venda}, 0, -D_{t_1}}} \leq 0 \right. \right].$$

Descobrimos que, apesar de um conceito de risco e uma função-objetivo diferente, a decisão de aplicação é a mesma, pois a estratégia ótima que cria valores não-positivos e maiores que o negativo da *perda esperada* é a de mínima *perda esperada*, ou seja

$$\begin{aligned} \tilde{\pi}_{\text{Venda}}^{\text{DBM}} &= \tilde{\pi}_{\text{Venda}} \\ &= \frac{1}{2} + \frac{\varepsilon}{20}. \end{aligned}$$

<sup>33</sup>Observamos que os valores já estão com sinal adaptado ao caso de valores negativos. Usamos a variável genérica  $X$  em vez de  $V$ , pois queremos mostrar a diferença entre as duas abordagens, independente se a posição é do comprador e do vendedor.

No caso do comprador do derivativo vale o resultado análogo que a sua estratégia ótima é dada pela estratégia ótima no caso da *perda esperada*. Portanto, o preço de negociação no mercado também é o mesmo:

$$x^{\text{DBM}} = \frac{20}{7}.$$

## 6.4 O Equivalente Monetário Determinístico (EMD) de grau $n$

### 6.4.1 Introdução

Uma forma de levar em conta os momentos da distribuição  $\mathbb{Q}$  dos retornos  $X$  de um ativo financeiro seria expandi-la por um método como funções geratrizes de momentos, funções características, Edgeworth, Taylor etc. Como já mencionado na seção 3.2, isso se aplica no ambiente de precificação se  $\mathbb{Q}$  já é a distribuição ajustada (distorcida) pelo risco. Já na alternativa de distorção dos valores por funções de utilidade é  $U(X)$  que temos que expandir. Pressupondo que  $U \in \mathcal{C}^{(n)}$  ( $n$ -vezes diferenciável e contínua) em torno de  $\mu_0 = E[X]$ , então vale que

$$E[U(X)] = U(\mu_0) + \sum_{i=1}^n \frac{U^{(i)}(\mu_0)}{i!} \mu_i + o_n(\mu_0), \quad (6.19)$$

com  $\mu_i := E[(X - \mu_0)^i]$  sendo o  $i$ -ésimo momento centrado. Truncando após o  $n$ -ésimo passo, temos que  $o_n(\mu_0) = 0$ .

O equivalente monetário determinístico (certo) é o valor determinístico  $EMD(X) \in \mathbb{R}$  que deixa o agente indiferente frente ao valor de utilidade esperada da variável aleatória  $X$  (vide seção 3.2), e pressupondo que  $U$  é inversível:

$$EMD(X) := U^{-1}(E[U(X)]).$$

Uma expansão de  $U(EMD(X))$ , até primeira ordem, em torno de  $\mu_0$ , resulta, sob a hipótese de  $EMD(X) \simeq \mu_0$ , em

$$\begin{aligned} U(EMD(X)) &\simeq U(\mu_0) \\ &\simeq U(\mu_0) + U^{(1)}(\mu_0) \cdot [EMD(X) - \mu_0]. \end{aligned}$$

Como vale também

$$U(EMD(X)) = E[U(X)],$$

usamos a expansão de Taylor de (6.19) para escrever

$$U(\mu_0) + \sum_{i=1}^n \frac{U^{(i)}(\mu_0)}{i!} \mu_i + o_n(\mu_0) \simeq U(\mu_0) + U^{(1)}(\mu_0) \cdot [\text{EMD}(X) - \mu_0] \quad (6.20)$$

Resolvendo essa equação por  $\text{EMD}(X)$ , truncando após o  $n$ -ésimo passo e observando que  $\mu_1 = 0$ , podemos introduzir a seguinte definição:

**Definição 6.7 (EMD de grau  $n$ )** *O equivalente monetário determinístico de grau  $n$ ,  $\text{EMD}_n$ , para uma variável aleatória é dado por*

$$\text{EMD}_n(X) := \mu_0 + \frac{1}{U^{(1)}(\mu_0)} \cdot \sum_{i=2}^n \frac{U^{(i)}(\mu_0)}{i!} \cdot \mu_i.$$

O  $\text{EMD}$  é uma função  $\text{EMD} : L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$  que atribui a uma variável aleatória um valor de forma subjetiva, pois o valor depende de uma função de utilidade  $U$ . O  $\text{EMD}$  pode ser usado para analisar qual função (ou classe de funções) de utilidade pode ser considerada para determinar o valor de  $X$ .

## 6.4.2 Comportamento racional do investidor

Tipicamente o comportamento pressuposto de um agente econômico (investidor) é que ele deve aumentar a sua aversão à incerteza para uma variável com incerteza  $S$  com uma diminuição do seu patrimônio  $x$  (vide PRATT, 1964). Para caracterizar esse e outros comportamentos, vale lembrar as seguintes noções para  $U \in \mathcal{C}^\infty$ .

Índice de aversão à incerteza absoluto:<sup>34</sup>

$$a_2(x) := r(x) := -\frac{U^{(2)}(x)}{U^{(1)}(x)}.$$

Índice de prudência absoluto:<sup>35</sup>

$$a_3(x) := p(x) := -\frac{U^{(3)}(x)}{U^{(2)}(x)}.$$

Índice de temperança absoluto:<sup>36</sup>

$$a_4(x) := t(x) := -\frac{U^{(4)}(x)}{U^{(3)}(x)}.$$

<sup>34</sup>PRATT (1964)

<sup>35</sup>KIMBALL (1993) e EECKHOUDT; GOLLIER; SCHNEIDER (1995)

<sup>36</sup>KIMBALL (1993) e EECKHOUDT; GOLLIER; SCHNEIDER (1995)



PRATT; ZECKHAUSER (1987) continuaram essa linha de pesquisa, postulando que um investidor confrontado com uma incerteza  $S^{(1)}$  está mais averso a incorrer uma outra (mesmo que seja independente) incerteza  $S^{(2)}$  que um investidor que não está exposto à primeira incerteza  $S^{(1)}$ . O nome atribuído a esse conceito foi de “*proper risk aversion*”<sup>37</sup>. KIMBALL (1993) diferenciou o conjunto de incerteza  $S^{(1)}$  que deve causar um aumento da aversão para uma incerteza adicional  $S^{(2)}$  e chamou o conceito de “*standard risk aversion*”<sup>38</sup>. Para poder analisar o comportamento de modo geral, introduzimos o índice de aversão absoluto ao  $k$ -ésimo momento centrado.<sup>39</sup>

**Definição 6.8 (Índice de aversão absoluto ao  $k$ -ésimo momento centrado)** *O índice de aversão absoluto ao  $k$ -ésimo momento (centrado em  $\mu_0$ ) é definido por*

$$a_k(\mu_0) := - \left. \frac{U^{(k)}(x)}{U^{(k-1)}(x)} \right|_{x=\mu_0}.$$

Para evitar os comportamentos economicamente não-plausíveis podem ser impostas condições na função de utilidade de von Neumann-Morgenstern  $U$  do investidor. Para o caso da aversão ao segundo momento centrado (aversão ao risco simples de PRATT, 1964), é claro que tem que valer  $\frac{\partial a_2(x)}{\partial x} \leq 0$ , i.e., a aversão absoluta ao segundo momento centrado é decrescente com o patrimônio do investidor  $x$ . KIMBALL (1993), por sua vez, mostrou que para aumentar a aversão à incerteza de  $S^{(2)} \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$  na existência de  $S^{(1)} \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ , é suficiente e necessário que

$$a_4(x) \geq a_3(x) \geq a_2(x) \geq 0.$$

**Observação 6.14** *Vulnerabilidade à incerteza, por ser um conceito mais geral que aversão à incerteza padrão, precisa somente da condição mais fraca de*

$$\left\{ \begin{array}{l} a_3(x) \geq a_2(x), \\ a_4(x) \geq a_2(x). \end{array} \right\}$$

A condição de que a aversão ao terceiro momento centrado tem que ser maior que a aversão ao segundo momento centrado segue de um argumento de aproximação do prêmio da incerteza aplicado ao caso de um choque (não correlacionado com os ativos na carteira) no patrimônio do investidor.

<sup>37</sup> “*Aversão própria ao risco*” em Português.

<sup>38</sup> “*Aversão ao risco padrão*” em Português.

<sup>39</sup> O índice foi chamado de “*medida absoluta de preferência pelo risco*” por PAULSSON; SPROULE; WAGENER (2002) e já usado em CABALLÉ; POMANSKY (1996).

Como nós queremos uma regra para caracterizar um investidor com a sua função de utilidade  $U$  de forma geral, queremos saber quais as condições que devemos pressupor para ter um comportamento economicamente plausível. A partir dos resultados citados, especialmente da aversão à incerteza padrão, fazemos as seguintes conjecturas para definir um investidor racional até o  $n$ -ésimo nível:

**Conjectura 6.2** *O índice de aversão à incerteza de nível  $k$  tem as seguintes características,  $\forall k \geq 2, x \in \mathbb{R}^+$*

1. *Positivo:*

$$a_k(x) \geq 0,$$

2. *Decrescente:*

$$\frac{\partial a_k(x)}{\partial x} \leq 0,$$

3. *Aumento da importância dos índices de aversão:*

$$a_{k+1}(x) \geq a_k(x), \quad k > 0.$$

Uma condição equivalente à de  $\partial a_k$  ser decrescente é de pressupor que as derivadas das utilidades troquem de sinal, ou seja, que a utilidade defina uma mistura exponencial.<sup>40</sup> Com a seguinte definição, conseguimos substituir a conjectura 6.2:

**Definição 6.9** *Dizemos que um investidor é racional para um nível  $n \geq 2$  se*

1. *ele preferir momentos ímpares positivos e evitar momentos pares e ímpares negativos.*<sup>41</sup>  $\forall 0 < k \leq n$

$$U^{(k)} < 0, k \text{ par}, \tag{6.21}$$

$$U^{(k)} > 0, k \text{ ímpar},$$

2. *os coeficientes de aversão à incerteza são crescentes:  $\forall k > 0$ :*

$$a_{k+1}(x) \geq a_k(x).$$

<sup>40</sup>Vide também a proposição 3.2 em CABALLÉ; POMANSKY (1996).

<sup>41</sup>A característica de sinais alternados das derivadas da função de utilidade foi chamado de "aversão ao risco misto" por CABALLÉ; POMANSKY (1996).

Para verificar que essa hipótese justifica um investidor racional, observamos que para todo  $k \geq 2$  e  $x \in \mathbb{R}$  vale que

1.

$$0 \leq -\frac{U^{(k)}(x)}{U^{(k-1)}(x)} = a_k(x).$$

2.

$$\begin{aligned} \frac{\partial a_k(x)}{\partial x} &= \frac{\partial \left( -\frac{U^{(k)}(x)}{U^{(k-1)}(x)} \right)}{\partial x} \\ &= a_k(x) \cdot (a_k(x) - a_{k+1}(x)) \\ &= a_k(x)^2 - \frac{U^{(k+1)}(x)}{U^{(k-1)}(x)} \leq 0. \end{aligned}$$

Para ter certeza que vale a última desigualdade, lembramos que temos somente duas possibilidades: no caso de  $k$  ímpar, temos  $U^{(k+1)}(x) \leq 0$  e, portanto,  $U^{(k-1)}(x) \leq 0$ . No caso de  $k$  par, temos  $U^{(k+1)}(x) \geq 0$  e  $U^{(k-1)}(x) \geq 0$ . Como vale  $a_{k+1} \geq a_k$ , segue o resultado.

### 6.4.3 Utilidade esperada e preferência por momentos

O uso das preferências na primeira parte da definição é justificado a partir das análises tradicionais da utilidade esperada por expansão de Taylor, que implica o uso das derivadas da utilidade  $U^{(k)}$  como “preços” dos momentos, vide (6.20). Porém, é importante tomar cuidado no uso equivalente da utilidade esperada e da preferência pelos momentos. Como BROCKETT; KAHANE (1992) mostraram, para quaisquer utilidade  $U$  que satisfaz (6.21), existem variáveis aleatórias  $X_1, X_2$  com relação qualquer entre os momentos, e.g.,

$$\begin{aligned} \mu_0(X_1) &> \mu_0(X_2), \\ \mu_2(X_1) &< \mu_2(X_2), \quad \text{etc.} \end{aligned}$$

tal que

$$E[U(X_1)] < E[U(X_2)].^{42}$$

Isso significa que mesmo que valha a implicação heurística “utilidade esperada  $\Rightarrow$  preferência por momentos”, o inverso não é sempre verdadeiro. Por isso, seria também precipitado chamar os coeficientes de aversão à incerteza  $a_k$  de coeficientes de aversão

Variável	Esperança	Variância	Assimetria	$E[U_1]$	$E[U_2]$	$E[U_3]$
$X_1$	4,9	6,35	0,9	0,955	1,676	2,135
$X_2$	5	6,25	1	0,874	1,647	2,076

Tabela 6.1: Inconsistência entre preferência por momentos e utilidade esperada. As funções de utilidade são:  $U_1(x) = 1 - \exp(-x)$ ,  $U_2(x) = \ln(1+x)$ ,  $U_3(x) = x^{\frac{1}{2}}$ .

aos momentos. O exemplo 2 de BROCKETT; KAHANE (1992), replicado na tabela 6.1, mostra o raciocínio.

Os autores mencionam que esse fato pode ser contraintuitivo, citando literatura que pressupõe uma equivalência entre preferência por momentos e utilidade esperada. Para esclarecer o paradoxo, lembramos que a preferência por momentos não é simplesmente por momentos, mas por momentos ponderados com as derivadas da função utilidade, pois usa-se uma aproximação por série de Taylor. Portanto, se usamos um critério de decisão baseado em momentos, é importante ponderá-los. Temos com o *EMD* uma ferramenta que usa elementos dos dois lados: os momentos centrados são “precificados” (ou ponderados) pelas aversões a todos os momentos anteriores. E como o comportamento racional é refletido nos coeficientes de aversão, mantém-se essa característica na avaliação por *EMD*. É conveniente reescrever o  $EMD_n$  de forma que destaque esses índices e mostre a conexão com os momentos da distribuição:

$$\begin{aligned}
EMD_n(X) &= \mu_0 + \frac{1}{U^{(1)}(\mu_0)} \cdot \sum_{i=2}^n \frac{U^{(i)}(\mu_0)}{i!} \mu_i \\
&= \mu_0 + \frac{1}{1} \frac{U^{(2)}(\mu_0)}{U^{(1)}(\mu_0)} \mu_2 + \frac{1}{2} \frac{U^{(3)}(\mu_0)}{U^{(1)}(\mu_0)} \mu_3 + \dots \\
&= \mu_0 + \frac{1}{1} \frac{U^{(2)}(\mu_0)}{U^{(1)}(\mu_0)} \mu_2 + \frac{1}{2} \frac{U^{(3)}(\mu_0)}{U^{(2)}(\mu_0)} \frac{U^{(2)}(\mu_0)}{U^{(1)}(\mu_0)} \mu_3 \\
&\quad + \frac{1}{3} \frac{U^{(4)}(\mu_0)}{U^{(3)}(\mu_0)} \frac{U^{(3)}(\mu_0)}{U^{(2)}(\mu_0)} \frac{U^{(2)}(\mu_0)}{U^{(1)}(\mu_0)} \mu_4 + \dots \\
&= \mu_0 + \sum_{i=2}^n \frac{(-1)^{i+1}}{i!} \mu_i \prod_{k=2}^i a_k(\mu_0).
\end{aligned}$$

Podemos também fazer uma discriminação entre os momentos pares e ímpares

$$\begin{aligned}
EMD_n(X) &= \mu_0 + \sum_{i=2}^n \frac{(-1)^{i+1}}{i!} \mu_i \mathbf{1}_{\{i \text{ ímpar}\}} \prod_{k=2}^i a_k(\mu_0) \\
&+ \sum_{i=2}^n \frac{(-1)^{i+1}}{i!} \mu_i \mathbf{1}_{\{i \text{ par}\}} \prod_{k=2}^i a_k(\mu_0) \\
&= \mu_0 + \sum_{i=1}^{\lceil \frac{n-1}{2} \rceil} \frac{1}{(2i+1)!} \mu_{2i+1} \prod_{k=2}^{2i+1} a_k(\mu_0) \\
&- \sum_{i=1}^{\lfloor \frac{n-1}{2} \rfloor} \frac{1}{(2i)!} \mu_{2i} \prod_{k=2}^{2i} a_k(\mu_0),
\end{aligned}$$

sendo que  $\lceil \cdot \rceil$  é o maior e  $\lfloor \cdot \rfloor$  o menor inteiro, respectivamente.

Observamos que  $\frac{1}{i!} \mu_i \prod_{k=2}^i a_k(\mu_0)$  pode ser interpretado como o preço do  $i$ -ésimo momento centrado e, assim, o  $EMD_n(X)$  compõe-se do acúmulo dos valores positivos dos momentos ímpares positivos,

$$EMD_n^+(X) = \sum_{i=1}^{\lceil \frac{n-1}{2} \rceil} \frac{1}{(2i+1)!} \cdot \mathbf{1}_{\{\mu_{2i+1} \geq 0\}} \cdot \mu_{2i+1} \prod_{k=1}^{2i+1} a_k(\mu_0),$$

menos os valores negativos causados pelos momentos pares e pelos momentos ímpares negativos,

$$\begin{aligned}
EMD_n^-(X) &= \sum_{i=1}^{\lfloor \frac{n-1}{2} \rfloor} \frac{1}{(2i)!} \mu_{2i} \prod_{k=2}^{2i} a_k(\mu_0) \\
&- \sum_{i=1}^{\lfloor \frac{n-1}{2} \rfloor} \frac{1}{(2i+1)!} \cdot \mathbf{1}_{\{\mu_{2i+1} < 0\}} \cdot \mu_{2i+1} \prod_{k=1}^{2i+1} a_k(\mu_0).
\end{aligned}$$

Ou seja,

$$EMD_n(X) = EMD_n^+(X) - EMD_n^-(X).$$

Segue que quanto maior um momento par centrado (ou ímpar centrado negativo), menor o  $EMD_n(X)$  e, portanto, menos atrativo é  $X$  para o investidor com a função de utilidade  $U$ . Isso significa que o  $EMD_n$  é equivalente a um sistema de preferência por momentos (ponderados).

Dependendo da ordem de expansão, a preferência de  $X_2$  usando valor esperado-variância não precisa ser revertida quando se usa o critério de utilidade esperada. Esse fato é pertinente também quando se usa a expansão por Taylor. Na tabela 6.2, mostramos os resultados da expansão por Taylor até ordem 3 e do  $EMD_3$ . Observamos que a decisão não seria invertida:  $X_2$  continua preferível.<sup>43</sup>

<sup>43</sup>O erro de aproximação da função de utilidade por uma série de Taylor foi analisado, por exemplo,

	$EMD_3$			Taylor até 3. Ordem		
Variável	$U_1$	$U_2$	$U_3$	$U_1$	$U_2$	$U_3$
$X_1$	1,875	4,370	4,581	0,970	1,685	2,068
$X_2$	2,042	4,488	4,693	0,973	1,706	2,097

Tabela 6.2: Manutenção da decisão com expansão até ordem 3.

Como o cálculo completo da utilidade esperada mostra uma inversão da decisão, temos que aumentar a ordem de expansão para chegar ao resultado previsto. A tabela 6.3 relata o quarto momento central  $\mu_4 = E[(X - \mu_0)^4]$  das duas distribuições e os valores da expansão da utilidade esperada por Taylor até ordem 4 junto com o  $EMD_4$ .

	$EMD_4$			Taylor até 4. Ordem		
Variável	$U_1$	$U_2$	$U_3$	$U_1$	$U_2$	$U_3$
X1	0,192	4,321	4,554	0,957	1,677	2,062
X2	-4,467	4,308	4,595	0,929	1,676	2,076

Tabela 6.3: Inversão da decisão com expansão até ordem 4 para as funções de utilidade  $U_1$  e  $U_2$ .

Agora podemos observar uma inversão da decisão nas utilidades  $U_1$  e  $U_2$  causada pelo valor alto do quarto momento centrado, ao qual o investidor é averso. Um problema surge quando se usa o  $U_3$ . Neste caso, vale ainda que  $EMD_4(X_2) \geq EMD_4(X_1)$  e a expansão por Taylor mostra o mesmo problema. O grau de expansão é portanto não suficiente neste caso. Fica em aberto para investigar em futuras pesquisas a seguinte questão:

**Afirmção 6.15** *Quais as condições às variáveis  $X_1, X_2 \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ ,  $p \in (0, \infty)$  e à função utilidade  $U$  para que não se inverte a ordem entre as variáveis por  $EMD_n$  quando se usa a utilidade:*

$$EMD_n(X_1) \geq EMD_n(X_2) \Leftrightarrow E[U(X_1)] \geq E[(U(X_2))].$$

Se soubéssemos estas condições, poderíamos substituir a hipótese da utilidade esperada pelo conceito de  $EMD_n$ , tendo no mesmo tempo uma forma simples para ponderar os momentos. Isso significa que temos um critério de ordenação por momentos.

O objetivo proposto no uso do  $EMD_n$  é o de ordenar as funções de utilidade para uma distribuição  $X$  dada. Mas, é também possível calcular o  $EMD_n$  para uma única função utilidade  $U$  dada. Neste caso escolha-se uma parametrização que esteja em concordância com um comportamento racional do investidor. Neste caso deve se usar o critério de  $EMD_n \geq 0$  para que uma função de utilidade seja aceitável.

**Exemplo 6.5** ( $EMD_n$  para utilidade quadrática) *Pressupõe-se  $n = 2$  e uma função de utilidade quadrática  $U(x) = ax - \frac{b}{2}x^2$ ,  $a, b \in \mathbb{R}$ . Como a função quadrática não leva em conta momentos maiores de 2, temos somente um índices de aversão à incerteza.<sup>44</sup>*

$$a_2(x) = -\frac{U^{(2)}(x)}{U^{(1)}(x)} = \frac{b}{2 \cdot (a - bx)}.$$

O  $a_2$  é maior que 0 se e somente se no caso  $b > 0$  vale que  $a > bx$  e no caso  $b < 0$  vale que  $a < bx$ .<sup>45</sup> Dependendo da parametrização, a primeira condição da racionalidade é satisfeita. Mas, observamos que a segunda condição não é satisfeita:

$$\frac{\partial a_2(x)}{\partial x} = \frac{b}{2} \cdot \left( \frac{1}{a - bx} \right)' = \frac{b^2}{2} (a - bx)^{-2} > 0,$$

<sup>44</sup>Obviamente, poderíamos analisar também o índice  $a_1(x) = -\frac{U^{(1)}(x)}{U(x)}$ . Mas, este índice é basicamente a elasticidade da utilidade.

<sup>45</sup>Apesar que é pressupõe-se  $b > 0$  para poder interpretar  $U$  como função de utilidade, pode ser, de forma geral, analisado  $b < 0$  também.

o que significa que os índices de aversão não são decrescentes. Este resultado é causado pela falta do índice de aversão ao terceiro momento  $a_3$ , pois segue que

$$\frac{\partial a_k(x)}{\partial x} = (a_k(x))^2 \geq 0.$$

O cálculo do  $EMD_2$  resulta em

$$\begin{aligned} EMD_2(X) &= \mu_0 - \mu_2 \cdot a_2(\mu_0) \\ &= \mu_0 - \mu_2 \frac{1}{a-b} \frac{b}{2}. \end{aligned}$$

Para que um investidor com utilidade quadrática aceite um ativo com retorno  $X$  tem que valer

$$a > \frac{\mu_2}{2 \cdot \mu_0} + b \cdot \mu_0.$$

Esta condição é mais forte que a condição necessária para racionalidade:  $a > b\mu_0$ . No caso que o investidor atribui um preço de  $b = 1$  à variação, no uso do  $EMD_2$  a metade do inverso do índice de Sharpe tem que ser adicionado ao fator multiplicativo do valor esperado para classificar o retorno como favorável.

**Exemplo 6.6** ( $EMD_n$  para utilidade logarítmica) Com  $U(x) = \ln(x)$ ,  $x > 0$ , temos para todo  $k \geq 1$ :

$$U^{(k)}(x) = (-1)^{k+1} \frac{(k-1)!}{x^k},$$

e

$$\begin{aligned} a_k(x) &= \frac{k-1}{x} > 0, \\ \frac{\partial a_k(x)}{\partial x} &= -\frac{k-1}{x^2} < 0 \end{aligned}$$

Isso implica no fato que um investidor com utilidade logarítmica é um investidor racional, independentemente de qualquer parametrização. Portanto, faz sentido pressupor que essa seja uma função de escolha que um investidor utilizará. No caso  $n = 2$  e  $X \sim N(\mu_0, \sigma)$ , temos para o  $EMD_2$ :

$$\begin{aligned} EMD_2(X) &= \mu_0 - \mu_2 \cdot a_2(\mu_0) \\ &= \mu_0 - \mu_2 \cdot \frac{1}{\mu_0}. \end{aligned}$$

Isso implica em  $EMD_2(X) > 0 \Leftrightarrow \mu_0^2 > \mu_2$ , ou seja, o investidor demandaria de um retorno  $X$  que a sua variação tem que ser menor que o quadrado do seu valor esperado.



Quando confrontados com incertezas nas distribuições dos retornos, a abordagem bayesiana tipicamente aplicada compara a utilidade esperada criada com as distribuições alternativas com a utilidade esperada criada com a distribuição prévia. Para isso funcionar tem que ser pressuposta uma função de utilidade do investidor. Para uma função de utilidade dada, o  $EMD_n$  atribui valores para as diversas distribuições e pode ser escolhida a distribuição que atinge o maior valor. Neste caso não é necessário a comparação com o valor de uma distribuição prévia. Apesar que não será tratado no presente trabalho desta forma, é possível aplicar uma forma mais sofisticada ainda: para toda distribuição alternativa determina-se qual é a função utilidade que maximiza o  $EMD_n$ . Isso significa, que procura-se a utilidade que melhor considera os momentos da distribuição sob análise. Depois, os diversos  $EMD_n$  podem ser comparados e será escolhida a combinação utilidade-distribuição com o maior valor do  $EMD_n$ .

**Exemplo 6.7 (Decisão entre utilidade quadrática e logarítmica)** *Se o investidor estiver confrontado com um retorno aleatório  $X$ , mas não sabe qual modelo (ou utilidade) deve escolher, ele pode usar o  $EMD$  para determinar a utilidade mais adequada. No caso de  $n = 2$  com  $U_1(x) = ax - \frac{b}{2}x^2$  e  $U_2(x) = \ln(x)$ ,  $x > 0$ , o critério de escolha se reduz a determinar se*

$$\frac{1}{a - b \cdot \mu_0} \cdot \frac{b}{2} > \frac{1}{\mu_0} \quad \Leftrightarrow$$

$$a > \frac{3}{2} \cdot b \cdot \mu_0.$$

#### 6.4.4 O $EMD$ , a irracionalidade da fronteira eficiente e a decisão de alocação

O objetivo dessa seção é mostrar algumas conexões do  $EMD$  com problemas em finanças que possam ser analisadas em futuras pesquisas.

Os resultados de BROCKETT; KAHANE (1992) significam que mesmo para dois ativos  $X_1 \sim N(\mu_1, \sigma_1^2)$  e  $X_2 \sim N(\mu_2, \sigma_2^2)$  como uma preferência simples (sem ponderação) dos momentos pode ser irracional do ponto de vista da utilidade esperada. Mas, se a variância for corretamente ponderada elimina-se a irracionalidade. Este fato tem um impacto na criação de uma fronteira eficiente no espaço valor esperado - variância. Basicamente, esta fronteira ordena em níveis de desvio-padrão e retorno esperado crescente os ativos e as suas combinações lineares. As combinações lineares de ativos que compõem a

fronteira é determinada a partir da relação entre os dois primeiros momentos sem considerar o preço destes momentos. Mesmo que a fronteira seja eficiente no sentido que para um desvio-padrão dado não existe combinação de ativos com maior retorno de que aquele dado pela carteira na curva, é bem possível que a curva não seja racional no sentido da utilidade esperada. Deixamos isso mais claro no seguinte exemplo:

**Exemplo 6.8 (Fronteira eficiente com irracionalidade)** *Sejam  $X_i, i = 1, 2, 3$  ativos com retornos esperados, desvio-padrões e utilidades esperadas crescentes para uma função de utilidade  $U$  racional, ou seja*

$$\begin{aligned}\mu_1 &< \mu_2 < \mu_3, \\ \sigma_1 &< \sigma_2 < \sigma_3, \\ E[U(X_1)] &< E[U(X_2)] < E[U(X_3)].\end{aligned}$$

*Com o aumento do coeficiente de aversão ao desvio-padrão ao longo da fronteira, a escolha no mundo retorno esperado-variância seria de começar com o ativo  $X_3$ , trocá-lo pelo  $X_2$  e ficar, finalmente, com  $X_1$ . Porém, o critério de decisão por utilidade esperada não muda: sempre será escolhido  $X_3$ .*

O exemplo mostra o problema de mudar a escolha dependendo do nível de aversão ao desvio-padrão. Mas, existe um problema maior: a escolha de ativos para criar as carteiras na fronteira nem precisa ser racional. No exemplo acima é bem possível, segundo o resultado da seção anterior, existir um ativo  $X_4$  com  $\mu_4 = \mu_2$  e  $\sigma_4 > \sigma_3$ . Na ausência de correlação, este ativo nunca faria parte de carteira alguma na fronteira eficiente. Porém, nada impede que

$$E[U(X_4)] > E[U(X_3)].$$

Pode-se pensar em criar uma fronteira eficiente de retorno esperado contra desvio-padrão que seja racional. Uma possibilidade para fazer isto é de ajustar o desvio-padrão de cada ativo pelo “preço” dele.<sup>46</sup> Como o preço do momento escolhe-se aquele derivada da expansão por Taylor, i.e., as derivadas do ordem equivalente ao momento ou o preço dado

<sup>46</sup>Lembramos que o coeficiente de aversão ao desvio-padrão usado para criar a fronteira eficiente pode ser visto como o preço pelo desvio-padrão quando se usa uma utilidade quadrática. Mas, esta função de utilidade não é considerada racional, portanto, a escolha de ativos não será. O que queremos fazer agora é poder criar uma decisão baseada na relação retorno esperado contra desvio-padrão que seja racional.

pelo *EMD*. Com o retorno esperado de  $\mu_0$ , o desvio-padrão tem que ser ajustado para  $\frac{1}{2}\sigma \cdot a_2(\mu_0)$ . Obviamente, a fronteira eficiente assim criada vai ser diferente da fronteira eficiente sem o ajuste.

Relacionado com o problema da fronteira eficiente é a pergunta de como a decisão de alocação entre um ativo com incerteza e um ativo sem incerteza depende das mudanças nos momentos centrais  $\mu_i$ . Para entender esta dependência deve ser analisada

$$\frac{\partial \lambda}{\partial \mu_i} = \frac{\partial \{\arg \max_{\lambda} E[U(W_0 \cdot (\lambda \cdot (r_{\text{Exc}} - r_{\text{rf}}) + r_{\text{rf}}))]\}}{\partial \mu_i},$$

sendo  $W_0$  o patrimônio inicial e  $r_{\text{exc}}$  o retorno excessivo do ativo com incerteza ao retorno “livre de incerteza”  $r_{\text{rf}}$ . Em PAULSSON; SPROULE; WAGENER (2002) foi feita uma análise parecida considerando mudanças<sup>47</sup> na distribuição do retorno em excesso. Os autores conseguiram mostrar que uma mudança na alocação  $\lambda$  somente acontece se a aversão à incerteza relativa ultrapassa um valor crítico.

---

<sup>47</sup>Foram analisados aumentos na dominância de primeira ordem, no *spread* preservando a média e no *downside risk*.

# Capítulo 7

## Conclusão

O objetivo dessa tese é propor métodos de decisão sob incerteza, levando em conta erros de modelo e de estimação dos parâmetros de um modelo escolhido. Como uma decisão ótima é um problema de otimização, analisamos detalhadamente o método de dualidade para processos aleatórios gerais, que se baseia na teoria de martingales. Um exemplo básico do mercado financeiro em forma de uma árvore trinária é usado para visualizar os comportamentos das abordagens.

Além da análise do método de otimização por dualidade, um assunto pouco tratado no Brasil, são propostas quatro abordagens originais para lidar com os problemas de erro estimação e de modelo:

1. A medida ERD é usada na otimização de uma carteira por utilidade (problema primal). Ela é aplicável para erros de modelo e de estimação.
2. As funções de penalidade são usadas para garantir uma estabilidade na composição da carteira no caso de erros de estimação. As abordagens de estabilidade e o uso de funções de penalidade são novidades.
3. O método de precificação consistente é proposto para apreçar qualquer variável aleatória. Essa forma de valoração é independente de uma função de utilidade, mas se baseia numa hipótese de comportamento do investidor.
4. O método do  $EMD_n$  identifica se uma função utilidade é *razoável* sob o ponto de vista do comportamento de um investidor caracterizado pela sua aversão à incerteza. Essa abordagem generaliza também o índice de aversão ao *risco* de PRATT (1964).

Discutimos e desenvolvemos as nossas próprias propostas sobre diversos assuntos relacionados ao problema central, tal como

- As noções de incerteza, risco e ambigüidade na seção 3.1.
- As visão abstrata que qualquer método de decisão se reduz a uma medida de distância e as suas relações (capítulo 4).
- O problema de inconsistência da medida de risco *perda esperada* com o critério de dominância estocástica na seção 6.3.2.
- A inconsistência da simples preferência por momentos com a utilidade esperada na seção 6.4.3.

Como foram propostas algumas abordagens originais, é da natureza de uma idéia nova não ser possível esgotar os pontos interessantes do assunto. Mas, ao longo do desenvolvimento, percebe-se que podem ser analisadas algumas questões que também servem para relacionar as diversas abordagens:

- Qual o limite analítico do preço consistente para algumas das distribuições mais comuns?
- Qual a conexão da *precificação consistente* com momentos e diversos métodos de expansão, tanto com o *EMD*.
- Qual o preço de um derivativo com o método de *EMD*?
- Qual o comportamento do *ERD* e do preço consistente num modelo dinâmico?
- Quais os resultados do uso do *EMD* para compor carteiras e determinar uma fronteira eficiente?

Deixamos esses problemas para futuras pesquisas.

# Apêndice A

## Resultados de dualidade para espaços convexos

O princípio básico da dualidade é que cada estrutura convexa tem uma descrição dual que determina unicamente cada objeto no espaço inicial (primal). Recuperando a definição de convexidade de um espaço  $\mathcal{X} \subseteq \mathbb{K}^n : \forall e, f \in \mathcal{X} \text{ e } \lambda \in [0, 1] \text{ vale que } \lambda \cdot e + (1 - \lambda) \cdot f \in \mathcal{X}$ , observamos que  $\bar{\mathcal{X}}$  é também convexo, caso contenha todos os limites das seqüências convergentes de  $\mathcal{X}$ . Uma descrição mais restrita, que aumenta as possibilidades da determinação do espaço dual, dá-se quando o espaço é localmente convexo.

**Definição A.1 (Localmente convexo)** Dizemos que  $E \subseteq \mathbb{K}^n$  é localmente convexo se vale para cada  $e \in \mathcal{X}$  que podemos achar um conjunto de elementos  $U \subseteq \mathbb{K}^n$  com  $e \in U$  de tal forma que  $U \cap \mathcal{X}$  é convexo.

De forma geral, podemos descrever o espaço dual de  $\mathcal{X}$ , i.e.,  $\mathcal{X}^*$ , como o espaço que contém todos os funcionais contínuos sobre elementos em  $\mathcal{X}$ .

O cone dual  $\mathcal{C}^*$  de um espaço  $\mathcal{C} \subseteq \mathbb{K}^n$  é um espaço dual específico e de fácil visualização:

$$\mathcal{C}^* := \{x^* \in \mathbb{K}^n \mid (x^* \cdot x) \leq 0, \forall x \in \mathcal{C}\},$$

sendo que  $(\cdot \cdot)$  é o produto corretamente definido para os elementos  $x^*, x$ . No caso de  $n < \infty$  e  $\mathbb{K}$  linear e espaço de Hilbert, temos o produto escalar definido, i.e.,  $(x^* \cdot x) = \langle x^* \cdot x \rangle = \sum_{i=1}^n x_i^* x_i$ . Uma visualização é dada pela figura A.1.

Observe também que, a priori, somente vale  $(\mathcal{C}^*)^* \supseteq \mathcal{C}$ , mas, por exemplo no caso de  $n < \infty$ , temos também que  $(\mathcal{C}^*)^* \subseteq \mathcal{C}$ . O cone dual pode ser geometricamente interpretado

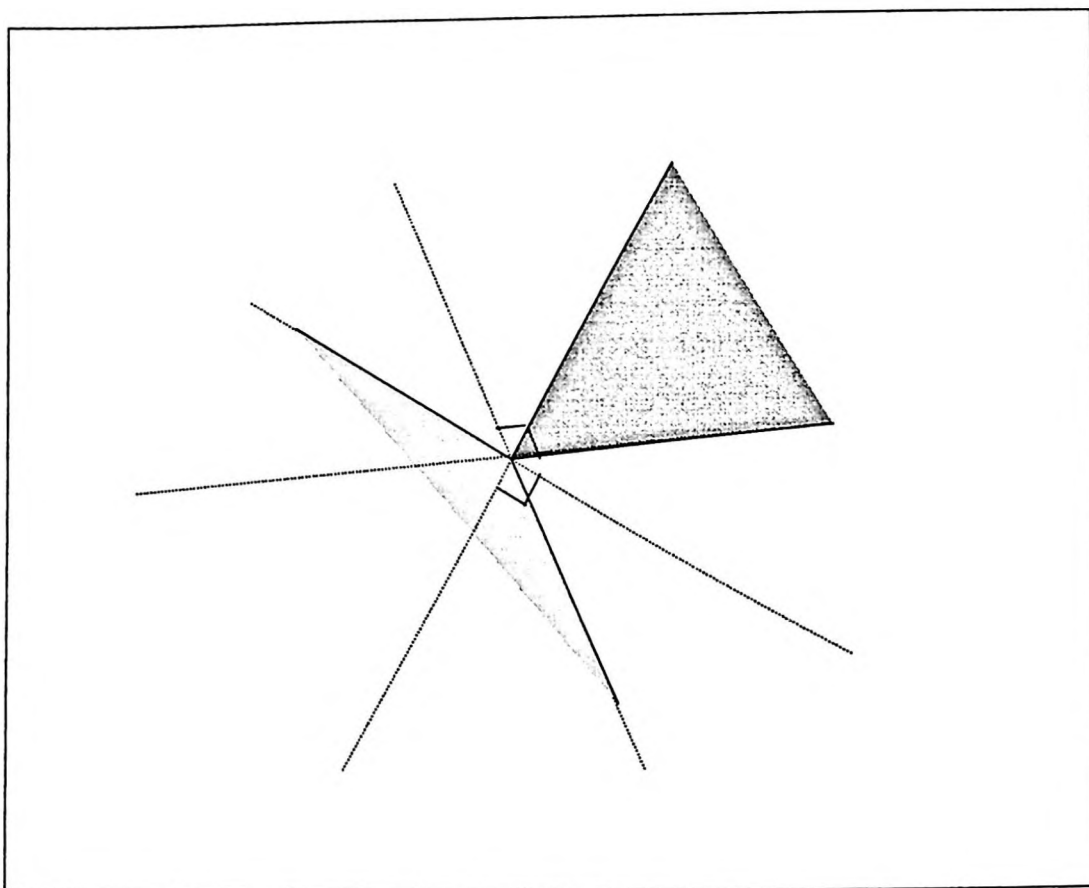


Figura A.1: O cone  $C \subseteq \mathbb{R}_+^2$  e seu cone dual  $C^*$ .

como todos os elementos do espaço  $\mathbb{K}^n$ , que mantêm um ângulo obtuso, enquanto o espaço do cone polar  $C^+$  é dado pelos elementos  $x \in \mathbb{K}^n$  que criam um ângulo agudo, i.e.,

$$C^+ := \{x^* \in \mathbb{K}^n \mid (x^* \cdot x) \geq 0, \forall x \in C\}.$$

**Definição A.2 (Semi-continuidade de baixo)** Uma função  $f : X \rightarrow Y$  é chamada de semi-continua de baixa no ponto  $x \in X$  se segue do fato de  $x \in f^{-1}(A)$  que  $B(x, \delta) \subseteq f^{-1}(A)$  para cada  $A \subseteq Y$  aberta e  $\delta > 0$ .

Isso significa que, no ponto  $x$ , os valores da função  $f(x)$  não terão um salto dentro de um pequeno ambiente de  $x$ .

A formulação geral do teorema da dualidade é dado, por exemplo, em EKELAND; TEMAM (1976).

**Teorema A.1** Seja  $(E, F)$  um par de espaços vetoriais localmente convexos em dualidade separativa e  $(\mathcal{X}, \mathcal{Y})$  um par de subconjuntos convexos. Seja também  $\mathcal{L} : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ ,  $(x, y) \mapsto \mathcal{L}(x, y)$  uma função convexa em  $y$ , côncava em  $x$  e (semi-) contínua em  $(x, y)$ .

Vale então, se  $\mathcal{X}$  ou  $\mathcal{Y}$  é compacto e o outro conjunto é completo, que existe um ponto de inversão  $(\bar{x}, \bar{y}) \in (\mathcal{X}, \mathcal{Y})$  tal que

$$\begin{aligned} \mathcal{L}(\bar{x}, \bar{y}) &= \sup_{x \in \mathcal{X}} \inf_{y \in \mathcal{Y}} \mathcal{L}(x, y) \\ &= \inf_{y \in \mathcal{Y}} \sup_{x \in \mathcal{X}} \mathcal{L}(x, y). \end{aligned}$$

Querendo aplicar o resultado geral para o caso de  $V_T$  ser somente mensurável, temos que achar um conjunto  $E$  conveniente para encaixar o conjunto<sup>1</sup>

$$\mathcal{X}(x) = \{V_T \in L_+^0 \mid V_T \leq x + (\pi \cdot S)_T, \pi \in \mathcal{A}(x)\}.$$

Como  $L^1(\Omega, \mathcal{F}, \mathbb{P})$  é o candidato natural para encaixar as medidas neutralizadoras de risco, teríamos o problema de pressupor que  $\mathcal{X}(x) \subseteq E = L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ , ou seja, pressupomos que os valores finais são uniformemente limitados. Querendo evitar essa hipótese, podemos pensar em usar  $E = L^0$  diretamente como espaço para encaixar  $\mathcal{X}$ , mas isso cria outros problemas:

- $L^0$  e  $L^1$  não têm nenhuma forma de bipolaridade.
- $L^0$  nem é espaço localmente convexo.

A saída, proposta por exemplo em KRAMKOV; SCHACHERMAYER (1999), é observar que  $\mathcal{X} \subseteq L_+^0$  e  $\mathcal{M}(S) \subseteq L_+^1 \subseteq L_+^0$ .

Poderíamos agora aplicar diretamente o resultado geral da análise convexa, mas no caso geral temos que ter certeza que o ponto ótimo  $Y_T^*$  realmente é uma medida probabilística e não somente uma medida signada:  $E \left[ \frac{dQ^*}{dP} \right] \leq 1$  (vide KRAMKOV; SCHACHERMAYER, 1999). A solução dada para evitar massa incompleta é aumentar o conjunto de medidas neutralizadoras de risco de  $\mathcal{M}(S)$  para um conjunto  $\mathcal{Y}(y)$ .

O último passo na aplicação do teorema A.1 é verificar se  $\mathcal{X}(x)$  ou  $\mathcal{Y}(y)$  são compactos. Como isso não é verdade para nenhum dos dois conjuntos, DELBAEN; SCHACHERMAYER (1994) propuseram o uso de uma convergência ponto a ponto para substituir o critério de compacidade.

O ponto importante para a dualidade funcionar é a bipolaridade entre  $\mathcal{X}(x)$  e  $\mathcal{Y}(y)$ , i.e.,

$$\begin{aligned} \mathcal{X}(x)^0 &= \mathcal{Y}(y), \\ \mathcal{Y}(y)^0 &= \mathcal{X}(x)^{00} = \mathcal{X}(x). \end{aligned}$$

<sup>1</sup>Pressupomos sem perda de generalidade aqui que  $V_T \geq 0$ , somente para facilitar a notação.



Cada elemento de  $\mathcal{X}(x)$  tem um ângulo fechado com cada elemento de  $\mathcal{Y}(y)$ , o que é formalizado nas seguintes relações, equivalentes às relações de bipolaridade:

$$X \in \mathcal{X}(x) \Leftrightarrow E_{\mathbf{P}}[X \cdot Y] \leq 1, \forall Y \in \mathcal{Y}(y),$$

$$Y \in \mathcal{Y}(y) \Leftrightarrow E_{\mathbf{P}}[X \cdot Y] \leq 1, \forall X \in \mathcal{X}(x).$$

## Apêndice B

### A ERD para algumas distribuições

<i>Distribuição</i>	<i>Função de densidade</i>
normal	$f(x \mu;\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right],$ $x \in \mathbb{R}, \sigma > 0$
logística	$f(x \alpha;\beta) =$ $\frac{1}{\beta} \exp\left[1 + \exp\left(-\frac{x-\alpha}{\beta}\right)\right]^{-2} \exp\left(-\frac{x-\alpha}{\beta}\right),$ $x \in \mathbb{R}$
Laplace ( <i>double exponential</i> )	$f(x \alpha;\beta) = \frac{1}{2\beta} \exp\left(-\frac{ x-\alpha }{\beta}\right),$ $x \in \mathbb{R}$
Gumbel	$f(x \alpha;\beta) = \frac{1}{\beta} \exp\left[-\frac{x-\alpha}{\beta} - \exp\left(-\frac{x-\alpha}{\beta}\right)\right],$ $x \in \mathbb{R}$
uniforme contínua	$f(x \alpha;\beta) = \frac{1}{(\beta-\alpha)},$ $\alpha \leq x \leq \beta$
uniforme discreta	$\Pr(x n) = \frac{1}{n},$ $x = 1, 2, \dots, n$

Tabela B.1: As distribuições sob análise.

<i>Distribuição</i>	<i>Entropia</i>
normal	$\frac{1}{2} \ln(2\pi e\sigma^2)$
logística	$\ln(\beta) + 2$
Laplace	$1 + \ln(2) + \ln(\beta)$
Gumbel	$1 + \ln(\beta) + \gamma, \gamma \approx 0,5772$
uniforme contínua	$\ln(\beta - \alpha)$
uniforme discreta	$\ln(n)$

Tabela B.2: Entropias das distribuições.

<i>Distribuição</i>	<i>ERD</i>	Valor aproximado
normal	$\frac{1}{\sqrt{2e\pi}}$	0,241971
logística	$\frac{\pi}{\sqrt{3 \cdot e^2}}$	0,245471
Laplace	$\frac{1}{\sqrt{2 \cdot e}}$	0,26013
Gumbel	$\frac{\pi}{\sqrt{6 \cdot \exp(1+\gamma)}}$	0,264914
uniforme contínua	$\frac{1}{\sqrt{12}}$	0,288675
uniforme discreta	$\sqrt{\frac{1}{12} - \frac{1}{12n^2}}$	0,288675, $n = 500$

Tabela B.3: *ERD* para as distribuições.

## Apêndice C

### Convexidade de ERD nas estratégias

Sejam  $\omega_1, \omega_2 \in \mathcal{A}_K(x)$  e  $\alpha \in (0, 1)$ . Pressupõe, primeiramente, que  $D = 0$  e que  $V_T \sim N(\mu, \sigma)$ . O volume de uma carteira em  $T$  criada por  $\omega$  é dado por

$$\begin{aligned}
 \text{vol}(V_T^{x, \omega, 0, 0}) &= \exp \left\{ - \int \frac{1}{\sqrt{2\pi\omega\sigma}} \cdot e^{-\left(\frac{x-\omega\mu}{\omega\sigma}\right)^2} \right. \\
 &\quad \cdot \left. \left( \log \left( \frac{1}{\sqrt{2\pi\omega\sigma}} \right) + \left( \frac{x-\omega\mu}{\omega\sigma} \right)^2 \right) dx \right\} \\
 &= \exp \left\{ - \frac{\log \left( \frac{1}{\sqrt{2\pi\omega\sigma}} \right)}{\sqrt{2\pi\omega\sigma}} \int e^{-\left(\frac{x-\omega\mu}{\omega\sigma}\right)^2} dx \right. \\
 &\quad \cdot \left. \frac{1}{\sqrt{2\pi}(\omega\sigma)^2} \int e^{-\left(\frac{x-\omega\mu}{\omega\sigma}\right)^2} (x-\omega\mu)^2 dx \right\} \\
 &= \exp \left\{ \frac{\log(\sqrt{2\pi\omega\sigma})}{\sqrt{2\pi\omega\sigma}} \int e^{-\left(\frac{x-\omega\mu}{\omega\sigma}\right)^2} dx \right\} \\
 &\quad \cdot \exp \left\{ \frac{1}{\sqrt{2\pi}(\omega\sigma)^2} \int e^{-\left(\frac{x-\omega\mu}{\omega\sigma}\right)^2} (x-\omega\mu)^2 dx \right\} \\
 &= \sqrt{2\pi\omega\sigma} \exp \left\{ \frac{1}{\sqrt{2\pi}(\omega\sigma)^2} \int e^{-\left(\frac{x-\omega\mu}{\omega\sigma}\right)^2} (x-\omega\mu)^2 dx \right\}.
 \end{aligned}$$

O segundo termo é calculado pela integração por partes:

$$\begin{aligned}
 &\frac{1}{\sqrt{2\pi}(\omega\sigma)^2} \int \underbrace{e^{-\left(\frac{x-\omega\mu}{\omega\sigma}\right)^2}}_{=f'} \underbrace{(x-\omega\mu)^2}_{=g} dx \\
 &= \frac{1}{\sqrt{2\pi}(\omega\sigma)^2} \left( f(x) \cdot g(x) \Big|_{-\infty}^{\infty} - \int f(x) \cdot g'(x) dx \right)
 \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\sqrt{2\pi}(\omega\sigma)^2} \left( \frac{\omega\sigma}{-2(x-\omega\mu)} e^{-\left(\frac{x-\omega\mu}{\omega\sigma}\right)^2} \cdot (x-\omega\mu)^2 \Big|_{-\infty}^{\infty} \right. \\
&\quad \left. - \int \frac{\omega\sigma}{2 \cdot (x-\omega\mu)} e^{-\left(\frac{x-\omega\mu}{\omega\sigma}\right)^2} \cdot (-2)(x-\omega\mu) dx \right) \\
&= \frac{1}{\sqrt{2\pi}\omega\sigma} \int e^{-\left(\frac{x-\omega\mu}{\omega\sigma}\right)^2} dx \\
&= 1
\end{aligned}$$

segundo o resultado do anexo B:

$$\text{vol}(V_T^{x,\omega,0,0}) = \sqrt{2\pi}\omega\sigma$$

e o fato de que ERD é convexo na estratégia, pois o volume é linear e o desvio-padrão também.

## Referências Bibliográficas

- ACERBI, C. Spectral Measures of Risk: a coherent representation of subjective risk aversion. **Journal of Banking and Finance**, v.26, p.1505–1518, 2002.
- ACERBI, C.; TASCHE, D. On the Coherence of Expected Shortfall. **Journal of Banking and Finance**, v.26, p.1487–1503, 2002.
- ALLAIS, M. Le Comportement de L'homme Rationnel Devant Le Risque: critique des postulats e axiomes de l'école américaine. **Econometrica**, v.21, p.503–546, 1953.
- AMENDINGER, J. **Initial Enlargement of Filtration and Additional Information in Financial Markets**. 1999. Tese de Doutorado — Technische Universität Berlin.
- ANDERSON, E. W.; HANSEN, L. P.; SARGENT, T. J. **Robustness, Detection and the Price of Risk**. University of North Carolina, 2000. Working paper.
- ANSCOMBE, F.; AUMANN, R. A Definition of Subjective Probability. **Annals of Mathematical Statistics**, v.34, p.199–205, 1963.
- ARTZNER, P.; DELBAEN, F.; EBER, J.-M.; HEATH, D. Thinking Coherently. **Risk Magazine**, v.10, n.11, p.68–71, Nov. 1997.
- ARTZNER, P.; DELBAEN, F.; EBER, J.-M.; HEATH, D. Coherent measures of risk. **Mathematical Finance**, v.9, n.3, p.203–228, July 1999.
- AVRAMOV, D. Stock Return Predictability and Model Uncertainty. **Journal of Financial Economics**, v.64, n.3, p.423–458, June 2002.
- BALZER, L. A. Measuring Investment Risk: a review. **Journal of Investing**, v.3, n.3, p.47–58, Fall 1994.
- BARRY, C. B. Portfolio Analysis under Uncertain Means, Variances, and Covariances. **The Journal of Finance**, v.29, n.2, p.515–522, May 1974.
- BAWA, V. S. Optimal Rules for Ordering Uncertain Prospects. **Journal of Financial Economics**, v.2, n.1, p.95–121, 1975.

- BAWA, V. S.; BROWN, S. J.; KLEIN, R. W. **Estimation Risk and Optimal Portfolio Choice**. Amsterdam: North Holland, 1979.
- BECHERER, D. **Rational Hedging and Valuation with Utility-Based Preferences**. 2001. Tese de Doutorado — Technische Universität Berlin.
- BEN-TAL, A. The Entropic Penalty Approach to Stochastic Programming. **Mathematics of Operation Research**, v.10, n.2, p.263–279, May 1985.
- BEN-TAL, A.; TEBoulLE, M. Penalty Functions and Duality in Stochastic Programming Via Phi-Divergence Functionals. **Mathematics of Operations Research**, v.12, n.2, p.224–240, May 1987.
- BERGER, J. Minimax Estimation of a Multivariate Normal Mean under Polynomial Loss. **Journal of Multivariate Analysis**, v.8, p.173–180, June 1978.
- BERGER, M. **Geometry II**. Berlin: Springer-Verlag, 1987.
- BERNARDO, A.; LEDOIT, O. **Approximate Arbitrage**. University of California, 1999. Working paper.
- BERNARDO, A.; LEDOIT, O. Gain, Loss and Asset Pricing. **Journal of Political Economy**, v.108, n.1, p.144–172, Feb. 2000.
- BEST, M. J.; GRAUER, R. On the Sensitivity of Mean-Variance Efficient Portfolios to Changes in Asset Means: some analytical and computational results. **Review of Financial Studies**, v.4, p.315–342, 1991.
- BILLINGSLEY, P. **Probability and Measure**. 3.ed. New York: John Wiley and Sons, 1995.
- BLACK, F.; LITTERMAN, R. **Asset Allocation: combining investor views with market equilibrium**. Goldman Sachs and Co., Fixed Income Research, 1990. Working paper.
- BLACK, F.; LITTERMAN, R. Global Portfolio Optimization. **Financial Analysts Journal**, p.28–43, September-Oct. 1992.
- BLAVATSKYY, P. R. **A Stochastic Expected Utility Theory**. Universität Zürich: Institute for Empirical Research in Economics, 2005. Working paper. (231).
- BOYLE, P. P.; SIU, T. K.; YANG, H. Risk and Probability Measures. **Risk Magazin**, v.7, p.53–57, July 2002.
- BROCKETT, P. L.; KAHANE, Y. Risk, Return, Skewness and Preference. **Management Science**, v.38, n.6, p.851–866, June 1992.
- BROWN, S. The Effect of Estimation Risk on Capital Market Equilibrium. **Journal of Financial and Quantitative Analysis**, v.14, n.2, p.215–220, June 1979.
- CABALLÉ, J.; POMANSKY, A. Mixed Risk Aversion. **Journal of Economic Theory**, v.71, p.485–513, 1996.
- CAMERER, C. Individual Decision Making. In: KAGEL, J. H.; ROTH, A. E. (Ed.). **The Handbook of Experimental Economics**. Princeton: Princeton University Press, 1995. p.587–703.



- CAVADINI, F.; SBUELZ, A.; TROJANI, F. **Incorporating MR, ER and Robustness in Mean Variance Portfolio Choice**. University of Lugano, 2001. Seminarpaper.
- CERNY, A.; HODGES, S. **The Theory of Good-Deal Pricing in Financial Markets**. Imperial College, 2000. Working paper.
- CHEN, S.-N.; BROWN, S. J. Estimation Risk and Simple Rules for Optimal Portfolio Selection. **The Journal of Finance**, v.38, n.4, p.1087–1093, September 1983.
- CHOPRA, V. K.; ZIEMBA, W. T. The Effect of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice. **The Journal of Portfolio Management**, p.6–11, Winter 1993.
- COCHRANE, J. H.; SAAÍ-REQUEJO, J. Beyond Arbitrage: good-deal asset price bounds in incomplete markets. **Journal of Political Economy**, v.108, n.1, p.79–119, 2000.
- COSTA, O. L. V.; PAIVA, A. C. Robust Portfolio Selection Using Linear-Matrix Inequalities. **Journal of Economic Dynamics and Control**, v.26, p.889–909, 2002.
- COVER, T. M.; THOMAS, J. A. **Elements of Information Theory**. John Wiley, New York, 1991.
- COVER, T. M.; THOMAS, J. A. **Elements of Information Theory: solutions manual**. John Wiley, New York, 1992.
- COX, J. C.; HUANG, C.-F. Optimal Consumption and Portfolio Policies When Asset Prices Follow a Diffusion Process. **Journal of Economic theory**, v.49, p.33–83, 1989.
- COX, J. C.; HUANG, C.-F. A Variational Problem Arising in Financial Economics. **Journal of Mathematical Economics**, v.20, p.465–487, 1991.
- CSISZAR, I. Information-Type Measures of Difference of Probability Distributions and Indirect Observations. **Studia Sci. Math. Hungar.**, v.2, p.299–318, 1967.
- CSISZAR, I. Generalized Projections for Non-Negative Functions. **Acta Mathematica Hungaria**, v.68, n.1-2, p.161–185, 1995.
- CVITANIC, J.; KARATZAS, I. On Dynamic Measures of Risk. **Finance and Stochastics**, v.3, n.4, p.451–482, Oct. 1999.
- DANIELSSON, J.; JORGENSEN, B. N.; SARMA, M.; VRIES, C. G. de. **Comparing Risk Measures**. Eindhoven: Eindhoven University of Technology, 2005. Working paper.
- DAVIS, M. H. A. Option Pricing in Incomplete Markets. In: DEMPSTER, M. A. H.; PLISKA, S. R. (Ed.). **Mathematics of Derivative Securities**. Cambridge University Press, 1998. p.217–225.
- DELBAEN, F. **Coherent risk measures on general probability spaces**. Eidgenössische Technische Hochschule (ETH) Zürich, 2000. Preprint.
- DELBAEN, F.; GRANDITS, P.; RHEINLÄNDER, T.; SAMPERI, D.; SCHWEIZER, M.; STRICKER, C. Exponential Hedging and Entropic Penalties. **Mathematical Finance**, v.12, n.2, p.99–123, Apr. 2002.

- DELBAEN, F.; RHEINLÄNDER, T.; SCHWEIZER, M. **Exponential Hedging and Entropic Penalties**. Eidgenössische Technische Hochschule (ETH) Zürich, 2000. Working paper.
- DELBAEN, F.; SCHACHERMAYER, W. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, v.300, p.463–520, 1994.
- DELBAEN, F.; SCHACHERMAYER, W. **A Simple Counter-Example of to Several Problems in the Theory of Asset Pricing Which Arises in Many Incomplete Markets**. Eidgenössische Technische Hochschule Zürich, 1998. Working paper.
- DENNEBERG, D. **Non-Additive Measure and Integral**. Amsterdam: Kluwer Academic Publisher, 1994.
- DICKENSON, J. P. The Reliability of Estimation Procedures in Portfolio Analysis. *Journal of Financial and Quantitative Analysis*, v.9, p.447–462, September 1979.
- DUFFIE, D. **Dynamic asset pricing theory**. 2.ed. New Jersey: Princeton University Press, 1996.
- EECKHOUDT, L.; GOLLIER, C.; SCHNEIDER, T. Risk-Aversionm Prudence and Termperance: a unified approach. *Economic Letters*, v.48, n.3-4, p.331–336, June 1995.
- EICHHORN, D.; GUPTA, F.; STUBBS, E. Using Constraints to Improve the Robustness of Asset Allocation. *Journal of Portfolio Management*, v.24, p.41–48, Spring 1998.
- EISENHAUER, J. G. A Measure of Relative Dispersion. *Teaching Statistics*, v.13, n.2, p.37–38, 1993.
- EKELAND, I.; TEMAM, R. **Convex Analysis and Variational Problems**. Amsterdam: North Holland, 1976.
- ELKAROUI, N.; QUENEZ, M.-C. Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM Journal of Control and Computation*, v.33, n.1, p.26–66, Jan. 1995.
- EMBRECHTS, P.; KLÜPPELBERG, C.; MIKOSCH, T. **Extremal Events in Finance and Insurance**. Berlin: Springer, 1997.
- FENCHEL, W. On Conjugate Convex Functions. *Canadian Journal of Mathematics*, v.1, p.73–77, 1949.
- FISHBURN, P. C. Mean-Risk Analysis with Risk Associated with Below-Target Returns. *American Economic Review*, v.67, n.2, p.116–126, 1977.
- FLEMING, W. H.; RISHEL, R. W. **Deterministic and Stochastic Optimal Control**. Berlin: Springer, 1975.
- FÖLLMER, H.; LEUKERT, P. Quantile hedging. *Finance and Stochastics*, v.3, p.251–273, 1999.
- FÖLLMER, H.; LEUKERT, P. Efficient hedging: cost versus shortfall risk. *Finance and Stochastics*, v.4, n.2, p.117–146, 2000.

- FÖLLMER, H.; SCHIED, A. Convex Measures of Risk and Trading Constraints. *Finance and Stochastics*, v.6, n.4, p.429–447, Oct. 2002.
- FRANKFURTER, G. M.; PHILLIPS, H. E.; SEAGLE, J. P. Portfolio Selection: the effects of uncertain means, variances, and covariances. *Journal of Financial and Quantitative Analysis*, v.6, p.1251–1262, September 1971.
- FRITELLI, M. The minimal entropy martingale measure and the valuation problem in incomplete markets. *Mathematical Finance*, v.10, n.1, p.39–52, January 2000.
- FRITELLI, M. Introduction to a Theory of Value Coherent with the No-Arbitrage Principle. *Finance and Stochastics*, v.4, p.275–297, 2000.
- FRITELLI, M. Representing Sublinear Risk Measures and Pricing Rules. Università di Milano Bicocca, 2002. Working paper.
- FRITELLI, M.; GIANIN, E. R. Putting Order in Risk Measures. *Journal of Banking and Finance*, v.26, p.1473–1486, 2002.
- FROST, P. A.; SAVARINO, J. E. For Better Performance: constrain portfolio weights. *The Journal of Portfolio Management*, v.Fall, p.29–34, 1988.
- FROST, P.; SAVARINO, J. An Empirical Bayes Approach to Efficient Portfolio Selection. *Journal of Financial and Quantitative Analysis*, v.21, n.3, p.293–305, 1986.
- GALE, D.; KUHN, H. W.; TUCKER, A. W. Linear Programming and the Theory of Games. In: KOOLMANS, T. C. (Ed.). *Activity Analysis of Production and Allocation*. New York: John Wiley, 1951. p.298–316.
- GIBBS, A. L.; SU, F. E. On Choosing and Bounding Probability Metrics. Harvey Mudd College, 2002. Working paper.
- GIORGI, E. D. Reward-Risk Portfolio Selection and Stochastic Dominance. Institute for Empirical Research in Economics, 2002. Working paper No. 121.
- GOLDFARB, D.; IYENGAR, G. Robust Portfolio Selection Problems. Columbia University, 2002. Working paper: CORC Report 2002-2003.
- GORDAN, P. Über Die Auflösung Linearer Gleichungen mit Reellen Coefficienten. *Mathematische Annalen*, v.6, p.23–28, 1873.
- GRAUER, R.; SHEN, F. Do Constraints Improve Portfolio Performance? *Journal of Banking and Finance*, v.24, p.1253–1274, 2000.
- HANSEN, L. P.; JAGANNATHAN, R. Implications of Security Market Data for Models of Dynamic Economies. *Journal of Political Economy*, v.99, n.2, p.225–262, Apr. 1991.
- HARRISON, J. M.; KREPS, D. M. Martingales and Arbitrage in Multiperiod Securities Markets. *Journal of Economic Theory*, v.20, p.381–408, 1979.
- HE, H.; PEARSON, N. D. Consumption and Portfolio Policies with Incomplete Markets and Short-Sale Constraints: the finite dimensional case. *Mathematical Finance*, v.1, n.3, p.1–10, 1991.

- HOLTON, G. Subjective Value at Risk. *Financial Engineering News*, v.1, n.1, p.8–9, 1997.
- HORST, J. R. T.; ROON, F. A. D.; WERKER, B. J. M. *Incorporating Estimation Risk in Portfolio Choice*. Tilburg University, 2001. Working paper.
- HUANG, C.-F.; LITZENBERGER, R. H. *Foundations for Financial Economics*. Amsterdam: North-Holland, 1988.
- IYENGAR, G. *Robust Dynamic Programming*. Columbia University, 2002. Working paper: CORC Report. (TR-2002-07).
- JOBSON, J. D.; KORKIE, B. Estimation for Markowitz Efficient Portfolios. *Journal of the American Statistical Association*, v.75, n.371, p.544–554, September 1980.
- JOHNSON, R. A.; WICHERN, D. W. *Applied Multivariate Statistical Analysis*. 3.ed. New Jersey: Prentice-Hall, 1992.
- JORION, P. Bayes-Stein Estimation for Portfolio Analysis. *The Journal of Financial and Quantitative Analysis*, v.21, n.3, p.279–292, September 1986.
- KACPERCZYK, M. *Asset Allocation under Distribution Uncertainty*. University of Michigan, 2003. Working paper.
- KAHNEMANN, D.; TVERSKY, A. Prospect Theory: an analysis of decision under risk. *Econometrica*, v.47, p.263–291, 1979.
- KALLBERG, J. G.; ZIEMBA, W. T. Mis-Specification in Portfolio Selection Problems. In: BAMBERG, G.; SPREMANN, K. (Ed.). *Risk and Capital*. Berlin: Springer, 1984. (Lecture Notes in Economics and Mathematical Systems).
- KALYMON, B. A. Estimation Risk and the Portfolio Selection Model. *The Journal of Financial and Quantitative Analysis*, v.6, p.559–582, Jan. 1971.
- KARATZAS, I. *Lectures on the mathematics of finance*. Rhode Island: CRM Monograph Series, 1996.
- KARATZAS, I.; LEHOCZKY, J. P.; SHREVE, S. E. Optimal Portfolio and Consumption Decisions for a "small Investor" on a Finite Horizon. *SIAM Journal of Control and Optimisation*, v.25, p.1557–1586, 1987.
- KARATZAS, I.; LEHOCZKY, J. P.; SHREVE, S.; XU, G.-L. Martingale and Duality Methods for Utility Maximization in an Incomplete Market. *Siam Journal of Control and Optimization*, v.29, n.3, p.702–730, May 1991.
- KARATZAS, I.; SHREVE, S. *Brownian Motion and Stochastic Calculus*. 2.ed. Berlin: Springer, 1991.
- KARNI, E.; SCHMEIDLER, D. Utility Theory with Uncertainty. In: HILDENBRAND, W.; SONNENSCHNEIN, H. (Ed.). *Handbook of Mathematical Economics*. North-Holland: Elsevier, 1991. p.1763–1831.
- KEMPF, A.; KREUZBERG, K.; MEMMEL, C. *How to Incorporate Estimation Risk Into Markowitz Optimization*. University of Köln, 2001. Working paper.

- KIMBALL, M. S. Standard Risk Aversion. *Econometrica*, v.61, n.3, p.589–611, May 1993.
- KNIGHT, F. H. *Risk, Uncertainty, and Profit*. England: Houghton Mifflin, 1921.
- KRAMKOV, D. O.; SCHACHERMAYER, W. The Asymptotic Elasticity of Utility Functions and Optimal Investment in Incomplete Markets. *Annals of Applied Probability*, v.9, p.904–950, 1999.
- KREPS, D. M. Arbitrage and equilibrium in economies with infinitely many commodities. *Journal of Mathematical Economics*, v.8, p.15–35, 1981.
- KUHN, H. W. Solvability and Consistency for Linear Equations and Inequalities. *American Mathematics Monthly*, v.63, p.217–232, 1956.
- LAKNER, P. Utility Maximization with Partial Information. *Stochastic Processes and Applications*, v.56, p.247–274, 1995.
- LAUPRETE, G. J. *Portfolio Risk Minimization under Departures from Normality*. 2001. Tese de Doutorado — Massachusetts Institute of Technology.
- LAUPRETE, G. J.; SAMAROV, A. M.; WELSCH, R. E. *Robust Portfolio Optimization*. Cambridge, USA: Massachusetts Institute of Technology, 2003. Working paper.
- LIESE, F.; VAJDA, I. *Convex Statistical Distances*. Leipzig: Teubner-Verlag, 1987.
- LOISTL, O. The Erroneous Approximation of Expected Utility by Means of a Taylor's Series Expansion: analytic and computational results. *The American Economic Review*, v.66, n.5, p.904–910, Dec. 1976.
- LUENBERGER, D. G. *Investment Science*. New York: Oxford University Press, 1998.
- MACCHERONI, F.; MARINACCI, M.; RUSTICHINI, A. *A Variational Formula for the Relative Gini Concentration Index*. International Center for Economic Research, 2004. Working paper.
- MACCHERONI, F.; MARINACCI, M.; RUSTICHINI, A.; TABOGA, M. *Portfolio Selection with Monotone Mean-Variance Preferences*. International Center for Economic Research, 2004. Working paper.
- MAENHOUT, P. J. *Robust Portfolio Rules, Hedging and Asset Pricing*. Fontainebleau: INSEAD, 2001. Working paper.
- MAGILL, M. J. P.; QUINZII, M. *Theory of Incomplete Markets*. Cambridge, Massachusetts: MIT Press, 1996. v.1.
- MARKOWITZ, H. *Portfolio Selection: efficient diversification of investments*. New York: John Wiley and Sons, 1959.
- MARKOWITZ, H. M. Portfolio Selection. *Journal of Finance*, v.7, n.1, p.77–91, Mar. 1952.
- MARKOWITZ, H.; USMEN, N. Diffuse Priors Vs. Resampled Frontiers: an experiment. *Journal of Investment Management*, 4th quarter 2003.

- MERTON, R. C. An Intertemporal Capital Asset Pricing Model. *Econometrica*, v.41, p.867–887, September 1973.
- MERTON, R. C. On Estimating the Expected Return on the Market. *Journal of Financial Economics*, v.8, p.323–361, Dec. 1980.
- MERTON, R. C. *Continuous-Time Finance*. Berlin: Blackwell, 1990.
- MICHAUD, R. O. The Markowitz Optimization Enigma: is 'optimized' optimal? *Financial Analysts Journal*, p.31–42, Jan./Feb. 1989.
- MICHAUD, R. O. *Efficient Asset Management: a practical guide to stock portfolio optimization and asset allocation*. Boston: HBS Press, 1998.
- MUSIELA, M.; ZARIPHOUPOULOU, T. A Valuation Algorithm for Indifference Prices in Incomplete Markets. *Finance and Stochastics*, v.8, n.3, p.399–414, August 2004.
- NEUMANN, J. V. Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, v.100, p.295–320, 1928.
- NEUMANN, J. V.; MORGENSTERN, O. *Theory of Games and Economic Behavior*. 3.ed. New Jersey: Princeton University Press, 1953.
- ØKSENDAL, B. *Stochastic Differential Equations*. 5.ed. Berlin: Springer, 2000.
- OLIVEIRA, G. A. de; FERRAZ, C. N. Quanta Incerteza Há Na Fronteira Eficiente? O Bootstrap Na Avaliação de Performance. São Paulo: Bank Boston, 2001. Working paper apresentado no 1. encontro Brasileiro de Finanças em São Paulo.
- PALMQUIST, J.; URYASEV, S.; KROKHNNAL, P. *Portfolio Optimization with Conditional Value at Risk Objective and Constraints*. 1999.
- PÁSTOR, L.; STAMBAUGH, R. F. Costs of Equity and Model Mispricing. *Journal of Finance*, v.54, n.1, p.67–121, Feb. 1999.
- PÁSTOR, L.; STAMBAUGH, R. F. Comparing Asset Pricing Models: an investment perspective. *Journal of Financial Economics*, v.56, n.3, p.335–381, June 2000.
- PAULSSON, T.; SPROULE, R.; WAGENER, A. *The Demand for a Risky Asset: signing, jointly and severally, the effects of three distributional shifts*. Québec: Bishop's University, 2002. Working paper.
- PFLUG, G. C. How to Measure Risk? In: LEOPOLD-WILDBUGER, U.; FEICHTINGER, G.; KISTNER, K.-P. (Ed.). *Modelling and Decisions in Economics*. Heidelberg: Physica-Verlag, 1999. p.39–59.
- PLISKA, S. R. A Stochastic Calculus Model for Continuous Trading: optimal portfolio. *Mathematics Operation Research*, v.11, p.371–382, 1986.
- PRATT, J. W. Risk Aversion in the Small and in the Large. *Econometrica*, v.32, n.1-2, p.122–136, Jan.-Apr. 1964.
- PRATT, J. W.; ZECKHAUSER, R. Proper Risk Aversion. *Econometrica*, v.55, p.143–154, 1987.

- REZA, F. M. **An Introduction to Information Theory**. New York: Dover, 1961.
- ROCKAFELLAR, R. T. **Convex Analysis**. Princeton, New Jersey: Princeton University Press, 1970.
- ROGERS, L. C. G. **Duality in Constrained Optimal Investment and Consumption Problems: a synthesis**. 2001. Lecture held at the Workshop on Financial Mathematics and Econometrics, Montréal.
- ROGERS, L. C. G.; WILLIAMS, D. **Diffusions, Markov Processes, and Martingales**. John Wiley and Sons, New York, 1987. v.2: Ito Calculus.
- ROSS, S. A. A Simple Approach to the Valuation of Risky Streams. **Journal of Business**, v.51, n.3, p.453–475, July 1978.
- ROTHSCHILD, M.; STIGLITZ, J. E. Increasing Risk: i. a definition. **Journal of Economic Theory**, v.225-243, n.2, p.102–125, 1970.
- ROUGE, R.; ELKAROUI, N. Pricing via utility maximization and entropy. **Mathematical Finance**, v.10, n.2, p.259–276, Apr. 2000.
- SAVAGE, L. J. **The Foundations of Statistics**. 1.ed. New York: Dover Publications, 1954.
- SCHACHERMAYER, W. **Optimal Investment in Incomplete Markets When Wealth Can Become Negative**. Technische Universität Wien, 2000. Working paper.
- SCHACHERMAYER, W. **Utility Maximisation in Incomplete Markets**. Technische Universität Wien, 2004. Working paper.
- SCHACHERMAYER, W. **Portfolio Optimization in Incomplete Financial Markets**. Technische Universität Wien, 2004. Working paper.
- SCHERER, B. Portfolio Resampling: review and critique. **Financial Analysts Journal**, p.98–109, Nov./Dec. 2002.
- SCHULMERICH, M.; TRAUTMANN, S. **Local Expected Shortfall-Hedging. Part I: the discrete model**. Mainz: Johannes Gutenberg-Universität, 2001. Working paper.
- SCHWEIZER, M. Mean-variance hedging for general claims. **The Annals of Applied Probability**, v.2, n.1, p.171–179, 1992.
- SOMMER, D. **Pseudo-Arbitrage**. Universität Bonn, 1998. Nota de discussão. (B-442).
- SONDERMANN, D. Reinsurance in Arbitrage-Free Markets. **Insurance: Mathematics and Economics**, v.10, n.3, p.191–202, 12 1991.
- STEIN, C. Confidence Sets for the Mean of a Multivariate Normal Distribution. **Journal of the Royal Statistical Society**, v.24, p.265–296, 1962.
- STIEMKE, E. Über Positive Lösungen Homogener Linearer Gleichungen. **Mathematische Annalen**, v.76, p.340–342, 1915.

- STOER, J.; WITZGALL, C. **Convexity and Optimization in Finite Dimensions.** New York,Berlin: Springer, 1970.
- STUTZER, M. A Bayesian Approach to Diagnosis of Asset Pricing Models. **Journal of Econometrics**, v.68, n.2, p.367–397, August 1993.
- URYASEV, S.; ROCKAFELLAR, T. **Optimization of Conditional Value-at-Risk.** University of Florida, 1999. Working paper.
- WANG, S. S. Premium Calculation by Transforming the Layer Premium Density. **Astin Bulletin**, v.26, n.1, p.71–92, 1996.
- WIRCH, J. L.; HARDY, M. R. **Ordering of Risk Measures for Capital Adequacy.** Simon Fraser University: Simon Fraser University, 2000. Working paper.
- ZELLNER, A. **An Introduction to Bayesian Inference in Econometrics.** Malabar: Robert and Krieger, 1987.
- ZIMMER, C. J.; NIEDERHAUSER, B. M. Determining an Efficient Frontier in a Stochastic Moments Setting. **Revista Brasileira de Finanças**, v.2, n.1, p.91–116, June 2004.