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

ANÁLISE DE CARTEIRAS EM TEMPO DISCRETO

Fernando Hideki Kato

fkato@sti.com.br

Orientador: Prof. Dr. José de Oliveira Siqueira

São Paulo, 14 de abril de 2004

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

Diretora da Faculdade de Economia, Administração e Contabilidade
Profa. Dra. Maria Tereza Leme Fleury

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

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

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

ANÁLISE DE CARTEIRAS EM TEMPO DISCRETO

Fernando Hideki Kato

fkato@sti.com.br

Dissertação apresentada ao Departamento de Administração da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo como requisito parcial para a obtenção do título de Mestre em Administração.

Área de Concentração: Métodos Quantitativos e Informática

Orientador: Prof. Dr. José de Oliveira Siqueira

**São Paulo
2004**

© 2004 Fernando Hideki Kato

FICHA CATALOGRÁFICA

Elaborada pela Seção de Publicações e Divulgação do SBD/FEA/USP

Kato, Fernando Hideki

Análise de carteiras em tempo discreto / Fernando Hideki

Kato. -- São Paulo : FEA/USP, 2004.

144 f.

Dissertação (Mestrado) – Universidade de São Paulo, 2004

Bibliografia.

1. Finanças 2. Ações 3. Carteiras 4. Risco financeiro I. Faculdade de Economia, Administração e Contabilidade da USP II. Título.

CDD – 332

RESUMO

Nesta dissertação, o modelo de seleção de carteiras de Markowitz será estendido com uma análise em tempo discreto e hipóteses mais realísticas.

Um produto tensorial finito de densidades Erlang será usado para aproximar a densidade de probabilidade multivariada dos retornos discretos uniperiódicos de ativos dependentes. A Erlang é um caso particular da distribuição Gama. Uma mistura finita pode gerar densidades multimodais não-simétricas e o produto tensorial generaliza este conceito para dimensões maiores. Assumindo que a densidade multivariada foi independente e identicamente distribuída (i.i.d.) no passado, a aproximação pode ser calibrada com dados históricos usando o critério da máxima verossimilhança. Este é um problema de otimização em larga escala, mas com uma estrutura especial.

Assumindo que esta densidade multivariada será i.i.d. no futuro, então a densidade dos retornos discretos de uma carteira de ativos com pesos não-negativos será uma mistura finita de densidades Erlang. O risco será calculado com a medida *Downside Risk*, que é convexa para determinados parâmetros, não é baseada em quantis, não causa a subestimação do risco e torna os problemas de otimização uni e multiperiódico convexos.

O retorno discreto é uma variável aleatória multiplicativa ao longo do tempo. A distribuição multiperiódica dos retornos discretos de uma sequência de T carteiras será uma mistura finita de distribuições Meijer G. Após uma mudança na medida de probabilidade para a composta média, é possível calcular o risco e o retorno, que levará à fronteira eficiente multiperiódica, na qual cada ponto representa uma ou mais sequências ordenadas de T carteiras. As carteiras de cada sequência devem ser calculadas do futuro para o presente, mantendo o retorno esperado no nível desejado, o qual pode ser função do tempo. Uma estratégia de alocação dinâmica de ativos é refazer os cálculos a cada período, usando as novas informações disponíveis.

Se o horizonte de tempo tender a infinito, então a fronteira eficiente, na medida de probabilidade composta média, tenderá a um único ponto, dado pela carteira de Kelly, qualquer que seja a medida de risco.

Para selecionar um dentre vários modelos de otimização de carteira, é necessário comparar seus desempenhos relativos. A fronteira eficiente de cada modelo deve ser traçada em seu respectivo gráfico. Como os pesos dos ativos das carteiras sobre estas curvas são conhecidos, é possível traçar todas as curvas em um mesmo gráfico. Para um dado retorno esperado, as carteiras eficientes dos modelos podem ser calculadas, e os retornos realizados e suas diferenças ao longo de um *backtest* podem ser comparados.

Palavras-Chave: seleção de carteiras, otimização de carteiras, otimização em larga escala, produto tensorial, aproximação da densidade de probabilidade multivariada, mistura finita de distribuições Erlang, distribuição Gama Generalizada, *Downside Risk*, *Lower Partial Moment*, convexidade, medidas coerentes de risco, risco de falência, retornos aditivos e multiplicativos, mundos de retornos discretos e contínuos, multiperíodo, estratégia multiperiódica, critério de Kelly, seleção de modelo, convolução aditiva, convolução multiplicativa, transformada de Laplace, transformada de Mellin, função Meijer G, função Fox H.

ABSTRACT

Discrete Time Portfolio Analysis

In this thesis, Markowitz's portfolio selection model will be extended by means of a discrete time analysis and more realistic hypotheses.

A finite tensor product of Erlang densities will be used to approximate the multivariate probability density function of the single-period discrete returns of dependent assets. The Erlang is a particular case of the Gamma distribution. A finite mixture can generate multimodal asymmetric densities and the tensor product generalizes this concept to higher dimensions. Assuming that the multivariate density was independent and identically distributed (i.i.d.) in the past, the approximation can be calibrated with historical data using the maximum likelihood criterion. This is a large-scale optimization problem, but with a special structure.

Assuming that this multivariate density will be i.i.d. in the future, then the density of the discrete returns of a portfolio of assets with nonnegative weights will be a finite mixture of Erlang densities. The risk will be calculated with the Downside Risk measure, which is convex for certain parameters, is not based on quantiles, does not cause risk underestimation and makes the single and multiperiod optimization problems convex.

The discrete return is a multiplicative random variable along the time. The multiperiod distribution of the discrete returns of a sequence of T portfolios will be a finite mixture of Meijer G distributions. After a change of the distribution to the average compound, it is possible to calculate the risk and the return, which will lead to the multiperiod efficient frontier, where each point represents one or more ordered sequences of T portfolios. The portfolios of each sequence must be calculated from the future to the present, keeping the expected return at the desired level, which can be a function of time. A dynamic asset allocation strategy is to redo the calculations at each period, using new available information.

If the time horizon tends to infinite, then the efficient frontier, in the average compound probability measure, will tend to only one point, given by the Kelly's portfolio, whatever the risk measure is.

To select one among several portfolio optimization models, it is necessary to compare their relative performances. The efficient frontier of each model must be plotted in its respective graph. As the weights of the assets of the portfolios on these curves are known, it is possible to plot all curves in the same graph. For a given expected return, the efficient portfolios of the models can be calculated, and the realized returns and their differences along a backtest can be compared.

Keywords: portfolio selection, portfolio optimization, large-scale optimization, tensor product, approximation of the multivariate probability density function, finite mixture of Erlang distributions, Generalized Gamma distribution, Downside Risk, Lower Partial Moment, convexity, coherent risk measures, default risk, additive and multiplicative returns, worlds of discrete and continuous returns, multiperiod, multiperiodic strategy, Kelly's criterion, model selection, additive convolution, multiplicative convolution, Laplace transform, Mellin transform, Meijer G function, Fox H function.

Este trabalho é dedicado aos meus pais, Osamu e Alice, que sempre se esforçaram para me dar a melhor educação possível.

“One thing I have learned in a long life:
that all our science, measured against reality,
is primitive and childlike –
and yet it is the most precious thing we have.”
Albert Einstein (1879-1955)

BANCA EXAMINADORA

Prof. Dr. José de Oliveira Siqueira

Universidade de São Paulo
Faculdade de Economia, Administração e Contabilidade
Departamento de Administração

Prof. Dr. Adolpho Walter Pimazoni Canton

Universidade de São Paulo
Faculdade de Economia, Administração e Contabilidade
Departamento de Administração

Prof. Dr. Vladimir Belitsky

Universidade de São Paulo
Instituto de Matemática e Estatística
Departamento de Estatística

AGRADECIMENTOS

Ao meu orientador, Prof. Dr. José de Oliveira Siqueira, pela paciência e disposição para entender este trabalho.

Aos meus colegas Christian Johannes Zimmer e Marcos Ribeiro de Castro, pelas sugestões.

Aos inúmeros professores que tive, pelo conhecimento científico transmitido, o ativo mais precioso que nós temos.

CONTEÚDO

1.	INTRODUÇÃO	1
1.1.	MOTIVAÇÃO E IMPORTÂNCIA DO TEMA.....	1
1.2.	OBJETIVOS	1
1.3.	QUESTÃO DE PESQUISA.....	2
1.4.	METODOLOGIA E HIPÓTESES ADOTADAS.....	2
1.5.	ORGANIZAÇÃO DA DISSERTAÇÃO.....	3
1.6.	ABREVIACÕES.....	4
1.7.	GLOSSÁRIO	4
1.8.	NOTAÇÃO MATEMÁTICA.....	5
2.	REVISÃO DA LITERATURA	7
2.1.	INTRODUÇÃO	7
2.2.	RISCO.....	7
2.3.	ALOCAÇÃO DE CAPITAL UNIPERÍODICA.....	8
2.4.	MEDIDAS DE RISCO	9
3.	DISTRIBUIÇÕES, TRANSFORMADAS E FUNÇÕES ESPECIAIS	11
3.1.	INTRODUÇÃO	11
3.2.	FUNÇÕES RELACIONADAS.....	12
3.2.1.	Casos particulares.....	14
3.2.2.	Propriedades	14
3.3.	A DISTRIBUIÇÃO ERLANG E SUAS DISTRIBUIÇÕES RELACIONADAS	15
3.3.1.	Propriedades	15
3.4.	MISTURA E PRODUTO TENSORIAL.....	16
3.5.	CALIBRAÇÃO DE MISTURA E PRODUTO TENSORIAL.....	18
3.6.	NÚMERO DE AMOSTRAS E ERRO DE AMOSTRAGEM	26
3.7.	AS TRANSFORMADAS DE LAPLACE E DE MELLIN.....	26
3.7.1.	Casos particulares.....	28
3.7.2.	Propriedades da transformada de Mellin	29
3.8.	AS FUNÇÕES MEIJER G E FOX H.....	30
3.8.1.	Notações auxiliares para a Meijer G e para a Fox H	33
3.8.2.	Casos particulares.....	33
3.8.3.	Propriedades, integrais e transformadas	34
3.8.4.	Distribuições Meijer G e Fox H	36
3.9.	AValiação numérica da função Meijer G.....	36
3.10.	CONVOLUÇÃO ADITIVA	41

3.11.	COMBINAÇÃO LINEAR POSITIVA.....	45
3.12.	CONVOLUÇÃO MULTIPLICATIVA	48
3.13.	POTENCIAÇÃO.....	48
3.14.	VALOR ESPERADO.....	50
3.15.	<i>DOWNSIDE RISK</i>	52
3.16.	LEMAS AUXILIARES	56
3.17.	CONCLUSÕES	58
4.	RISCO	60
4.1.	INTRODUÇÃO	60
4.2.	PROBLEMAS DE MEDIDAS DE RISCO COM DISTRIBUIÇÕES GENÉRICAS	60
4.3.	MEDIDAS COERENTES DE RISCO	62
4.4.	<i>DOWNSIDE RISK</i>	63
4.4.1.	Convexidade.....	65
4.5.	FALÊNCIA.....	67
4.6.	CONCLUSÕES	69
5.	OTIMIZAÇÃO UNIPERÍODICA DE CARTEIRAS	70
5.1.	INTRODUÇÃO	70
5.2.	CONVOLUÇÃO ADITIVA	70
5.3.	COMBINAÇÃO LINEAR POSITIVA.....	71
5.4.	FRONTEIRA EFICIENTE	72
5.4.1.	Distribuição multivariada dos retornos dos ativos	73
5.4.2.	Distribuição dos retornos da carteira.....	74
5.4.3.	Retorno esperado.....	75
5.4.4.	<i>Downside Risk</i>	75
5.4.5.	Formulação do problema de otimização.....	76
5.5.	CONCLUSÕES	77
6.	OTIMIZAÇÃO MULTIPERÍODICA DE CARTEIRAS	78
6.1.	INTRODUÇÃO	78
6.2.	MUNDOS DE RETORNOS DISCRETOS E CONTÍNUOS	79
6.3.	VARIÁVEL ALEATÓRIA MULTIPLICATIVA	81
6.4.	DISTRIBUIÇÃO MULTIPERÍODICA DOS RETORNOS DA CARTEIRA	82
6.5.	CONVOLUÇÃO MULTIPLICATIVA	83
6.6.	ESPERANÇA E VARIÂNCIA DE VARIÁVEIS ALEATÓRIAS MULTIPLICATIVAS EM FUNÇÃO DO TEMPO	83
6.7.	DISTRIBUIÇÃO COMPOSTA MÉDIA	85
6.7.1.	Função utilidade potência.....	90
6.8.	DISTRIBUIÇÃO COMPOSTA MÉDIA COM LONGEVIDADE	91
6.8.1.	Concavidade da esperança.....	93
6.8.2.	Convexidade do <i>Downside Risk</i>	95
6.9.	REBALANCEAMENTO CALCULADO POR PROGRAMAÇÃO DINÂMICA	95
6.10.	FRONTEIRA EFICIENTE	96
6.10.1.	Distribuição multivariada dos retornos dos ativos	97

6.10.2. Distribuição dos retornos da carteira.....	97
6.10.3. Distribuição composta média	98
6.10.4. Valor esperado.....	99
6.10.5. <i>Downside Risk</i>	99
6.10.6. Formulação do problema de otimização.....	99
6.11. UMA ESTRATÉGIA DINÂMICA.....	101
6.12. CONCLUSÕES	101
 7. OTIMIZAÇÃO DE CARTEIRAS COM INFINITOS PERÍODOS	103
7.1. INTRODUÇÃO	103
7.2. LIMITE DA DISTRIBUIÇÃO COMPOSTA MÉDIA	103
7.2.1. Função utilidade logarítmica	106
7.3. O CRITÉRIO DE KELLY	106
7.4. FRONTEIRA EFICIENTE	110
7.5. DERIVADAS DO CRITÉRIO DE KELLY	111
7.6. CONCLUSÕES	112
 8. SELEÇÃO DE UM MODELO DE OTIMIZAÇÃO DE CARTEIRAS	114
8.1. INTRODUÇÃO	114
8.2. MOTIVAÇÃO.....	115
8.3. METODOLOGIA	115
8.4. CONCLUSÕES	117
 9. CONCLUSÕES	119
9.1. CONCLUSÕES	119
9.2. CONTRIBUIÇÕES ORIGINAIS	121
9.3. LIMITAÇÕES E SUGESTÕES PARA PESQUISAS FUTURAS	122
 APÊNDICE A – CADERNOS MATHEMATICA	124
EXEMPLO 3.5.1	124
EXEMPLO 3.5.2	125
EXEMPLO 3.9.1	127
EXEMPLO 3.10.1	128
EXEMPLO 3.11.1	129
EXEMPLO 6.7.1	130
CONJECTURA 6.8.1.....	132
 REFERÊNCIAS	133

1. INTRODUÇÃO

1.1. Motivação e importância do tema

Esta dissertação foi motivada pela existência de algumas críticas na teoria de alocação de capital, também chamada de seleção de carteiras ou de otimização de carteiras.

Embora a grande quantidade de dados históricos de cotações possibilite uma boa aproximação da densidade de probabilidade multivariada dos retornos, no modelo de seleção de carteiras de Markowitz (1952) são usados somente as médias, variâncias e covariâncias das taxas de retorno dos ativos. O uso da matriz de variância-covariância assume uma normal multivariada (ou, mais genericamente, uma distribuição elíptica) para as taxas de retorno. Mas a distribuição normal permite taxas de retorno menores que -100% , o que não é possível para cotações de ações.

A definição de risco como desvio-padrão penaliza ativos com caudas longas para a direita. Markowitz (1991a, p.194) reconheceu este problema e propôs a semivariância como uma medida de risco alternativa e preferível (Markowitz, 1991a, p.374).

Considerando-se que o período de tempo é um dia, então os investidores usualmente têm um horizonte de tempo maior que um período. Mas os modelos de alocação de capital multiperíódicos existentes são baseados em função utilidade, usam o desvio-padrão como medida de risco ou assumem que o horizonte de tempo é infinito. Essas são premissas que simplificam o cálculo. Livros didáticos em Finanças nem mesmo mencionam que o retorno discreto é uma variável multiplicativa ao longo do tempo, um fato de importância fundamental na análise multiperíódica.

Tendo disponíveis vários modelos de alocação de capital, um investidor deve selecionar um deles antes de aplicar seu dinheiro. É necessário definir um critério para a seleção de um dentre vários modelos.

Estes são problemas importantes, facilmente reconhecidos por todos da área de Finanças. O desejo de obter um modelo matematicamente mais exato e computacionalmente factível motivou a pesquisa do autor neste tema e resultou neste trabalho.

1.2. Objetivos

O objetivo primário desta dissertação é estender o modelo de seleção de carteiras de Markowitz com uma aproximação da densidade multivariada dos retornos discretos uniperíódicos dos ativos e usá-la para calcular o retorno esperado e o risco das carteiras. A medida de risco deve ter sentido financeiro e tirar

proveito da densidade da carteira.

Os objetivos secundários são: i) estender a análise uniperiódica para os casos multiperiódico e com infinitos períodos; e ii) propor um procedimento para comparar os desempenhos relativos entre modelos de seleção de carteiras.

1.3. Questão de pesquisa

Durante sua concepção, esta dissertação gerou as questões de pesquisa a seguir.

1. Como obter uma aproximação da densidade multivariada de variáveis aleatórias dependentes?
2. Como calcular a densidade dos retornos discretos de uma carteira a partir da densidade multivariada dos retornos discretos de seus ativos?
3. Como definir a medida de risco para tirar proveito da densidade dos retornos discretos da carteira?
4. Como calcular a densidade dos retornos discretos de uma seqüência de carteiras com base em suas densidades uniperiódicas?
5. Qual critério usar para comparar os desempenhos dos modelos propostos frente ao dos existentes?

A formulação da questão de pesquisa reúne estas questões.

- Como estender o modelo de seleção de carteiras de Markowitz com as hipóteses de que os retornos discretos dos ativos dependentes devem ser não-negativos, o risco deve ser calculado sobre os retornos indesejáveis e o tempo de investimento deve ser maior que um período; e como comparar os desempenhos de modelos de seleção de carteiras?

1.4. Metodologia e hipóteses adotadas

Esta dissertação é uma proposição de novos modelos de alocação de capital em tempo discreto e de um procedimento do tipo *backtest* para compará-los com os existentes. Vale a pena lembrar que novos modelos nunca podem ser considerados totalmente novos porque eles sempre são uma evolução de modelos anteriores. A criação de modelos é considerada uma pesquisa exploratória qualitativa que, em princípio, não requer obtenção de dados numéricos. Todos os capítulos teóricos desta dissertação são parte da metodologia.

Nesta dissertação, o problema de otimização de carteiras será estudado com a hipótese de que o tempo é discreto. As premissas ou hipóteses simplificadoras assumidas neste trabalho serão dadas a seguir.

- O investidor é racional (ele prefere mais retorno e menos risco).
- Não há custos de transação.
- Os pesos dos ativos nas carteiras são reais não-negativos.
- Há probabilidade zero de que o retorno de um ativo seja menor ou igual a zero (isto é, que a taxa de retorno seja menor ou igual a -100%).

- Há dependência estatística entre os retornos dos ativos em cada instante do tempo.
- O tempo é discreto (em algumas partes deste trabalho, o tempo será considerado contínuo).
- As distribuições são independentes e identicamente distribuídas (i.i.d.) ao longo do tempo.

As variáveis independentes são os preços dos ativos cotados em bolsa. Com base nos preços, os retornos são obtidos. O horizonte de tempo do investidor e seu nível desejado de retorno esperado ou de risco também são variáveis independentes. As variáveis dependentes são os pesos dos ativos nas infinitas carteiras que podem ser formadas. As carteiras mais importantes são aquelas sobre a fronteira eficiente e, em particular, aquela com o nível desejado de retorno esperado ou risco.

1.5. Organização da dissertação

Como será visto ao longo dos capítulos, a análise em tempo discreto permite uma clara separação dos passos envolvidos na otimização de carteiras.

No **Capítulo 2** será feita uma revisão da literatura sobre os modelos existentes de otimização de carteiras e alguns tópicos relacionados.

A parte matemática desta dissertação está concentrada no **Capítulo 3**. Isto evitará a interrupção do raciocínio financeiro nos capítulos subseqüentes. Serão apresentados a distribuição Erlang, as funções e distribuições relacionadas à Erlang e os conceitos de mistura e produto tensorial. O critério da máxima verossimilhança será usado na calibração dos parâmetros da densidade multivariada representada por um produto tensorial finito de Erlangs. A maldição da dimensionalidade surgirá neste problema de otimização. As funções especiais Meijer G e Fox H serão apresentadas, assim como uma fórmula para a avaliação numérica de um caso particular. As convoluções aditiva e multiplicativa serão apresentadas. Algumas propriedades envolvendo as densidades Erlang e Meijer G serão demonstradas. O *Downside Risk* será calculado para algumas distribuições. Um resultado interessante que será mostrado é o comportamento da variância e do desvio-padrão de uma variável aleatória multiplicativa em função do tempo contínuo.

Algumas medidas de riscos serão mostradas no **Capítulo 4**. As propriedades coerentes das medidas de risco serão discutidas. A medida escolhida para esta dissertação foi o *Downside Risk*. Embora não satisfaça todas as propriedades “coerentes”, ela não causa o problema de sub ou superestimação dos retornos “indesejáveis” ou “ruins” e torna convexos os problemas de otimização envolvendo a construção da fronteira eficiente. O risco de falência devido a carteiras alavancadas será discutido.

A otimização uniperiódica de carteiras será tratada no **Capítulo 5**. O retorno discreto é uma variável aditiva em cada instante de tempo. Se a densidade multivariada dos retornos dos ativos for um produto tensorial finito de Erlangs, então a densidade dos retornos de uma carteira com pesos positivos será uma mistura finita de Erlangs. O *Downside Risk* destas carteiras será calculado.

A otimização multiperiódica de carteiras será abordada no **Capítulo 6**. O retorno discreto é uma variável multiplicativa ao longo do tempo. Os mundos de retornos discretos e contínuos serão definidos. A densidade multiperiódica dos retornos discretos de uma seqüência de carteiras será calculada a partir das

densidades uniperiódicas. Será conveniente mudar a distribuição multiperiódica para a composta média. O retorno esperado e o *Downside Risk* serão calculados. A estratégia de alocação para o horizonte de tempo desejado e uma estratégia dinâmica serão discutidas.

No **Capítulo 7**, será assumido que o horizonte de tempo do investidor tende a infinito. Será mostrado que o critério de Kelly pode ser obtido com a ajuda do mundo de retornos contínuos ou com propriedades da transformada de Mellin. Quando o horizonte de tempo tender a infinito, a distribuição composta média de carteiras de pesos constantes tenderá a uma Delta de Dirac. A fronteira eficiente, se existir, será um único ponto e a medida de risco se tornará irrelevante para uma decisão financeira.

No **Capítulo 8**, será proposto um critério para auxiliar um investidor racional a selecionar um dentre vários modelos de otimização de carteiras. As fronteiras eficientes dos modelos devem ser traçadas em um mesmo gráfico, um nível de retorno deve ser fixado e um *backtest* deve ser realizado. As séries temporais dos retornos realizados e de suas diferenças serão usados para a comparação dos desempenhos relativos entre os modelos.

No **Capítulo 9**, serão apresentadas as conclusões, contribuições originais, limitações e sugestões para trabalhos futuros.

Neste trabalho, foi usado o programa Mathematica® versão 5.0, da empresa Wolfram Research.

1.6. Abreviações

i.i.d.	independente e identicamente distribuído
<i>apud</i>	em (citação indireta)
<i>et al.</i>	e outros
s.t.	<i>subject to</i> (sujeito às seguintes restrições)
p.	página
N.	número
V.	volume
CRRA	<i>Constant Relative Risk Aversion</i>
HARA	<i>Hyperbolic Absolute Risk Aversion</i>
CAPM	<i>Capital Asset Pricing Model</i>
VaR	<i>Value-at-Risk</i>
CVaR	<i>Conditional Value-at-Risk</i>

1.7. Glossário

Retorno	Preço no período t sobre o preço no período $t - 1$.
Taxa de retorno	Variação de preço sobre o preço original.
Retorno discreto	Retorno em um período discreto.

Retorno contínuo	Logaritmo (usualmente natural) do retorno discreto.
Mundo dos retornos discretos ou mundo multiplicativo	Mundo onde os retornos são capitalizados em períodos discretos.
Mundo dos retornos contínuos ou mundo aditivo	Mundo onde os retornos contínuos em diferentes instantes podem ser somados.
Variável aditiva	Variável que possui a operação de soma bem definida.
Variável multiplicativa	Variável que possui a operação de produto bem definida.
Convolução aditiva	Operação matemática usada para se obter a densidade da soma de duas variáveis independentes.
Convolução multiplicativa ou de escala	Operação matemática usada para se obter a densidade do produto de duas variáveis independentes.
Medida de risco	Função que mede os eventos indesejáveis ou arriscados de uma variável.
<i>Downside Risk</i>	Um exemplo de medida de risco.

1.8. Notação matemática

$\mathbb{C}$	Conjunto dos números complexos.
$\mathbb{R}$	Conjunto dos números reais.
$\mathbb{R}_+ = \{x \in \mathbb{R} x > 0\}$	Conjunto dos números reais positivos.
$\mathbb{R}_+^* = \{x \in \mathbb{R} x \geq 0\}$	Conjunto dos números reais não-negativos.
$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$	Conjunto dos números inteiros.
$\mathbb{N} = \{1, 2, \dots\}$	Conjunto dos números inteiros positivos.
$\mathbb{N}^* = \{0, 1, 2, \dots\}$	Conjunto dos números inteiros não-negativos.
$\mathbb{D} \subseteq \mathbb{R}$	Um subconjunto dos números reais.
P_t	Preço de um ativo ou carteira no tempo t .
$R_t = \frac{P_t}{P_{t-1}}$	Retorno de um ativo ou carteira no tempo t .
$r_t = R_t - 1$	Taxa de retorno de um ativo ou carteira no tempo t .
$R_f = 1 + r_f$	Retorno do ativo livre de risco (com variância nula).
r_f	Taxa de retorno do ativo livre de risco (com variância nula).
w_i	Peso do ativo i na carteira.
W_t	Riqueza do investidor no tempo t .
X	Variável aleatória.
$X \sim f(x)$	Variável aleatória X que segue a densidade $f(x)$.

X^t

$X^{\times t} = \underbrace{X \cdot X \cdot \dots \cdot X}_t$

$E[X] = E(f(x))$

$V[X] = V(f(x))$

$\mathcal{R}[X] = \mathcal{R}(f(x))$

$E[U(X)]$

$[X_1, X_2, \dots, X_D]^\top$

$\tilde{R} = \ln(R)$

$\tilde{f}(\tilde{R}) = e^{\tilde{R}} f(e^{\tilde{R}})$

$(f_1 \oplus f_2)(x)$

$(f_1 \otimes f_2)(x)$

$f^{\otimes n}(x)$

$f^{\oplus n}(x) = \underbrace{(f \oplus \dots \oplus f)}_n(x)$

$f^{\otimes n}(x) = \underbrace{(f \otimes \dots \otimes f)}_n(x)$

$f(x) \xrightarrow{\mathcal{L}} F(s) = \mathcal{L}\{f(x); s\}$

$f(x) \xrightarrow{\mathcal{M}} F(s) = \mathcal{M}\{f(x); s\}$

$f^{(n)}(x) = \frac{\partial^n f(x)}{\partial x^n}$

$\text{Re}(z)$

$i = \sqrt{-1}$

$\llbracket \cdot \rrbracket$

□

■

Variável aleatória X elevada à potência t .

Produto de t cópias independentes da variável aleatória X .

Esperança da variável aleatória X que segue a densidade $f(x)$.

Variância da variável aleatória X que segue a densidade $f(x)$.

Risco da variável aleatória X que segue a densidade $f(x)$.

Esperança da função utilidade da variável aleatória X .

Vetor transposto com D variáveis aleatórias.

Retorno contínuo no mundo dos retornos contínuos é o logaritmo natural do retorno discreto.

A densidade do retorno contínuo.

Convolução aditiva. Densidade da soma das variáveis aleatórias independentes X_1 e X_2 .

Convolução multiplicativa ou de escala. Densidade do produto das variáveis aleatórias independentes X_1 e X_2 .

Potenciação. Densidade da n -ésima potência da variável aleatória X .

Convolução aditiva de n densidades iguais.

Convolução multiplicativa ou de escala de n densidades iguais.

A transformada de Laplace de $f(x)$ é $F(s)$.

A transformada de Mellin de $f(x)$ é $F(s)$.

n -ésima derivada de $f(x)$.

Parte real de z .

Raiz quadrada de -1 .

Condições suficientes para a validade de uma expressão.

Fim da prova (quadrado de Halmos).

Fim do exemplo.

2. REVISÃO DA LITERATURA

2.1. Introdução

A quantidade de artigos e livros com tópicos relacionados à seleção ou otimização de carteiras é muito grande para uma ampla revisão da literatura. Neste capítulo, serão mencionados apenas alguns dos pontos mais importantes. Para um bom levantamento histórico, veja Bernstein (1996).

2.2. Risco

Risco é um conceito com várias definições. De acordo com um artigo de Borch (1969) sobre utilidade, a “Teoria do Risco” teve sua origem com o trabalho de Tetens (1789), que estudou ordenamentos que dependem da média e do desvio médio de funções de distribuição.

Fischer (1906, p.409, *apud* Grootveld e Hallerbach, 1999) é um autor que considerou “a chance dos ganhos caírem abaixo da linha de pagamento de juros”.

De acordo com Levy e Sarnat (1984, p.236), vários economistas identificaram o risco em investimentos com a dispersão dos retornos. Keynes, por exemplo, identificou o risco envolvido em um investimento com os possíveis desvios a partir do retorno médio. De acordo com Keynes (1937, *apud* Levy e Sarnat, 1984), um indivíduo que investe em um ativo cujos retornos têm uma distribuição bastante dispersa deve receber um prêmio para compensá-lo pelo risco tomado.

Tal como Keynes, Hicks (1946, *apud* Levy e Sarnat, 1984) também identificou a variância dos retornos com risco. Hicks enfatizou o fato de que quanto maior a dispersão dos retornos (para um dado nível de esperança), menor é a atratividade do investimento. Ele também enfatizou que quando os retornos são incertos, o terceiro momento centrado na esperança da distribuição, o índice de assimetria, pode também ser um fator significativo nas decisões do investidor.

Embora Marschak (1938, *apud* Levy e Sarnat, 1984) tenha dito que a decisão sob condições de risco deve refletir todos os momentos da distribuição, ele também notou que, em muitos casos, dois momentos, “a esperança matemática e o coeficiente de variação”, serão suficientes. Em outras palavras, Marschak também identificou (em alguns casos) o risco em investimentos com a variância, ou com o coeficiente de variação (desvio-padrão dividido pela média).

Domar e Musgrave (1944, *apud* Levy e Sarnat, 1984), por outro lado, identificaram o risco envolvido em um investimento sob condições de incerteza com a possibilidade de suportar uma perda. Por

isso, sugeriram que os investidores devem medir o risco somente com essa possibilidade: “De todas as possíveis questões que o investidor pode perguntar, a que parece mais importante para nós está relacionada com a probabilidade de a taxa de retorno real ser menor que zero, isto é, com a probabilidade de perda. Esta é a essência do risco.”. Estes dois autores desenvolveram um índice quantitativo para o risco, afetado tanto pela probabilidade da ocorrência de um resultado negativo quanto pelo tamanho da possível perda. Eles enfatizaram o segmento negativo da distribuição de probabilidade e, de acordo com o modelo, uma maior dispersão não necessariamente envolve um maior risco.

Baumol (1963, *apud* Levy e Sarnat, 1984) é outro autor que argumentou que a variância por si só não indica risco. De acordo com este autor, risco reflete principalmente a possibilidade de que a variável aleatória possa ter valores extremamente baixos. Se o retorno esperado de um investimento é alto relativo a seu desvio-padrão, Baumol sugeriu que a diferença entre o retorno esperado e k vezes o desvio-padrão ($E[R] - k\sigma$) fosse tomada como índice de risco, pois a probabilidade de uma variável aleatória ter um valor k desvios-padrão abaixo de sua média é limitada (pela Desigualdade de Chebyshev) a $1/k^2$.

Uma pesquisa realizada por Petty, Scott e Bird (1975, *apud* Park e Sharp-Bette, 1990, p.380) indicou que as pessoas nos negócios vêem risco associado primariamente com a probabilidade de não alcançar um retorno alvo. Quase 40% dos executivos corporativos entrevistados descrevem risco desta forma. Em outras palavras, a gerência está mais preocupada com variações negativas do que com a variação total possível dos investimentos. A segunda definição de risco mais popular encontrada na pesquisa estava relacionada com a variação dos retornos, que é equivalente à variância como medida de risco.

Bernstein (1996) fez uma longa pesquisa histórica sobre risco, mostrando sua evolução dos antigos gregos até os dias atuais. Pedersen e Satchell (1998) deram exemplos de várias medidas de risco.

2.3. Alocação de capital uniperiódica

Markowitz, em seu artigo clássico (1952), estudou a seleção de carteiras em um plano retorno *versus* risco (mais especificamente, no plano variância dos retornos *versus* retorno esperado). A primeira grande contribuição de Markowitz foi ter percebido que o risco de uma carteira não é a soma dos riscos dos ativos componentes. A segunda foi ter especificado um bom critério para a escolha de carteiras. Usando conceitos estatísticos, Markowitz calculou o retorno e o risco de carteiras não-alavancadas e, então, traçou uma região ou compacto nesse gráfico. Os estimadores das médias, variâncias e correlações das taxas de retorno dos ativos componentes são usualmente obtidos a partir das taxas de retorno históricas (passadas). Note que a realização de inferências para o futuro, com base em dados históricos, requer a hipótese de que a densidade multivariada do futuro será a mesma do passado.

Para um dado nível de retorno esperado, um investidor racional (aquele que prefere mais retorno e menos risco) escolherá a carteira com o menor risco possível; e para um dado nível de risco, ele escolherá a carteira com maior retorno esperado. Estas carteiras formam a fronteira eficiente, que é a fronteira superior esquerda do compacto no gráfico retorno esperado *versus* risco. Ela também é chamada Pareto-ótima, não-

inferior, não-dominada ou admissível; veja Yu (1985, p.22).

Não é conhecida uma fórmula fechada para a fronteira eficiente para o caso no qual os pesos dos ativos devem ser não-negativos. Markowitz (1991a, p.154, 316) criou o método da linha crítica para resolver numericamente este problema. Ele também mencionou que Wolfe (1957, *apud* Markowitz, 1991a, p.337) criou uma variante do método Simplex para calcular o mínimo de uma função quadrática com matriz positiva semidefinida sujeito a restrições lineares.

Markowitz (2000, p.11, 39) chama o modelo no qual os pesos dos ativos podem ser negativos de Modelo de Black. Neste caso, a fronteira eficiente é uma hipérbole no plano taxa de retorno esperada *versus* desvio-padrão, e é uma parábola no plano taxa de retorno esperada *versus* variância. A equação da fronteira foi obtida por Merton (1972) e pode também ser encontrada em Ingersoll, Jr. (1987, p.82) e em Huang e Litzenberger (1988, p.63). Pelo teorema da separação em dois fundos, com duas carteiras desta fronteira, é possível obter todas as demais.

Com a inclusão de um ativo sem risco, a fronteira eficiente passa a ser a semi-reta que parte do ativo sem risco e tangencia a fronteira eficiente original no plano retorno esperado *versus* desvio-padrão. Este é o teorema da separação em um fundo.

2.4. Medidas de risco

Existem várias operacionalizações matemáticas, definições ou medidas para o conceito de risco. Algumas dessas definições serão dadas a seguir, considerando-se que X é uma variável aleatória do tipo taxa de retorno, $p(x)$ é sua função densidade de probabilidade, $P(x) = \text{Prob}(X \leq x)$ é sua função densidade de probabilidade acumulada, μ é a sua esperança, α é um quantil especificado (ex.: 5%), b é uma taxa de retorno dada e $dP(x)$ é a notação da integral de Lebesgue-Stieltjes.

Variância (Markowitz, 1952).

$$V[X] = \sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 dP(x) \quad (2.1)$$

Desvio-padrão.

$$\sigma = \sqrt{\sigma^2} = \left(\int_{-\infty}^{\infty} (x - \mu)^2 dP(x) \right)^{\frac{1}{2}} \quad (2.2)$$

Semivariância (Markowitz, 1991a, p.193).

$$S_b[X] = \frac{1}{n} \sum_{i=1}^n (\max \{b - x_i, 0\})^2 \quad (2.3)$$

$$S_b[X] = \int_{-\infty}^b (b - x)^2 dP(x) \quad (2.4)$$

Probabilidade de perda.

$$L_0[X] = \int_{-\infty}^{0-} dP(x) \quad (2.5)$$

Valor-em-risco (Artzner *et al.*, 1999; Delbaen, 2000; Acerbi e Tasche, 2002a).

$$VaR_\alpha[X] = -\inf \{x \mid P(x) > \alpha\} \quad (2.6)$$

Valor-em-risco (Acerbi e Tasche, 2002b).

$$VaR_\alpha[X] = -\sup \{x \mid P(x) \leq \alpha\} \quad (2.7)$$

A definição de valor-em-risco (VaR) é não-trivial nos casos em que há distribuições Delta de Dirac ou em que a densidade é zero em um intervalo logo após a taxa de retorno do quantil α . Neste último caso, há uma abscissa inferior e uma superior para o mesmo quantil. Nas definições (2.6) e (2.7), a abscissa superior é usada. O sinal negativo é devido à variável aleatória ser do tipo taxa de retorno.

Outra medida de risco frequentemente mencionada em artigos recentes é o *Conditional Value-at-Risk*, ou *Expected Shortfall*, cuja definição precisa é complexa e pode ser encontrada em Acerbi e Tasche (2002b).

3. DISTRIBUIÇÕES, TRANSFORMADAS E FUNÇÕES ESPECIAIS

3.1. Introdução

Os conceitos matemáticos necessários para esta dissertação estão concentrados neste capítulo. Isto evitará a interrupção do raciocínio financeiro nos capítulos seguintes.

Neste trabalho, serão usadas variáveis aleatórias do tipo retorno, que terão que ser adicionadas e multiplicadas. Dadas as densidades de duas variáveis aleatórias independentes X e Y , será necessário determinar as densidades de $X+Y$ e de XY . Isto será feito com as convoluções aditiva e multiplicativa, por meio das transformadas de Laplace e de Mellin. Também será necessário determinar a densidade de $w_1X + w_2Y$, onde $w_1 \geq 0$, $w_2 \geq 0$ e $w_1 + w_2 = 1$, quando X e Y são dependentes, e encontrar uma forma de representar esta dependência.

Uma variável aleatória é dita aditiva se ela possui a operação de soma bem definida, isto é, pode-se somar duas variáveis do mesmo tipo. Uma variável aleatória é dita multiplicativa se ela possui a operação de multiplicação bem definida, isto é, pode-se multiplicar duas variáveis, resultando em outra com sentido físico ou econômico.

Uma mistura de distribuições é uma soma ponderada de distribuições e possibilita um melhor ajuste de dados observados. A soma dos pesos deve ser unitária. Alguns pesos podem ser negativos, mas em problemas de ajuste, eles podem ser forçados a ser não-negativos porque garantir a não-negatividade da densidade é um problema difícil. É mais simples usar pesos não-negativos e aumentar o número de parcelas na mistura. O produto tensorial é uma extensão deste conceito para várias dimensões. Ele permite a representação da densidade multivariada de variáveis aleatórias dependentes por meio da soma ponderada de produtos de densidades univariadas. A separabilidade das funções ajuda no cálculo das integrais que surgem quando é necessário obter a densidade da combinação linear de variáveis aleatórias dependentes.

Após definir a família de densidade da mistura ou do produto tensorial, é necessário ajustar ou calibrar os pesos e os parâmetros das densidades componentes de acordo com um critério de ajuste. Nesta dissertação, será usado o critério da máxima verossimilhança. O oposto da verossimilhança é uma função convexa sobre os pesos da mistura e do produto tensorial. Mas esta função é, em geral, não-convexa sobre os parâmetros das densidades componentes.

A família da Gama Generalizada inclui várias distribuições famosas, e um caso particular é a

distribuição Gama. Se o parâmetro de forma da distribuição Gama for um inteiro positivo, então ela se torna a distribuição Erlang, e se esse parâmetro for igual a 1, então se obtém a distribuição Exponencial. A mistura de Erlangs possui propriedades importantes para esta dissertação.

A soma de duas ou mais variáveis aleatórias independentes que seguem uma Erlang ou uma mistura finita de Erlangs resulta em uma variável aleatória que segue uma mistura finita de Erlangs. A mistura finita de Erlangs é uma distribuição fechada para a convolução aditiva. Se a densidade multivariada de variáveis aleatórias dependentes for um produto tensorial finito de Erlangs, então a densidade de uma combinação linear positiva destas variáveis será uma mistura finita de Erlangs.

O produto de duas ou mais variáveis aleatórias independentes que seguem Erlangs gera uma variável aleatória que segue uma distribuição baseada em um caso particular da função Meijer G. Como a Erlang é um caso particular desta classe de distribuições Meijer G, esta última é fechada em relação à convolução multiplicativa (ou de escala).

A função Meijer G é um caso particular da função Fox H, que é a função especial mais genérica conhecida atualmente. A Fox H pode representar muitas funções elementares, outras funções especiais e integrais definidas. A função Meijer G é usada internamente por programas tais como Mathematica® e Maple® para o cálculo de integração algébrica.

3.2. Funções relacionadas

Antes de apresentar a distribuição Erlang, as transformadas integrais e as funções especiais, é necessário definir a função Gama e algumas funções relacionadas.

Definição 3.2.1. Função Gama. A função Gama $\Gamma : \mathbb{C} - \{0, -1, -2, \dots\} \rightarrow \mathbb{C}$ é definida por

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt, \text{ para } \operatorname{Re}(z) > 0, \quad (3.1)$$

e pela propriedade

$$\Gamma(z) = \frac{\Gamma(z+1)}{z}, \text{ para } \operatorname{Re}(z) \leq 0, z \notin \{0, -1, -2, \dots\}. \quad (3.2)$$

O gráfico dessa função, para argumentos reais, pode ser visto na **Figura 3.2.1**. Para argumentos inteiros positivos, a função Gama se torna a função fatorial do predecessor do argumento.

Figura 3.2.1. Função Gama. Gráfico da função Gama para argumentos reais. Os pólos ocorrem em $0, -1, -2, \dots$.

Definição 3.2.2. Função Gama Incompleta Superior. A função Gama Incompleta Superior $\Gamma : \mathbb{C} - \{0, -1, -2, \dots\} \times \mathbb{R}_+^* \rightarrow \mathbb{C}$ é definida por

$$\Gamma(z, a) = \int_a^{\infty} t^{z-1} e^{-t} dt, \text{ para } \operatorname{Re}(z) > 0, \quad (3.3)$$

e pela propriedade

$$\Gamma(z, a) = \frac{\Gamma(z+1, a) - e^{-a} a^z}{z}, \text{ para } \operatorname{Re}(z) \leq 0, z \notin \{0, -1, -2, \dots\}. \quad (3.4)$$

Definição 3.2.3. Função Gama Incompleta Inferior. A função Gama Incompleta Inferior $\gamma : \mathbb{C} - \{0, -1, -2, \dots\} \times \mathbb{R}_+^* \rightarrow \mathbb{C}$ é definida como

$$\gamma(z, a) = \Gamma(z) - \Gamma(z, a). \quad (3.5)$$

Definição 3.2.4. Função Psi. A derivada do logaritmo natural da função Gama é chamada de função Psi $\psi : \mathbb{C} - \{0, -1, -2, \dots\} \rightarrow \mathbb{C}$ e é definida como

$$\psi(z) = \frac{d}{dz} \ln(\Gamma(z)) = \frac{\Gamma'(z)}{\Gamma(z)}. \quad (3.6)$$

Definição 3.2.5. Função PolyGama. As derivadas de ordem superior da função Psi são representadas pela função PolyGama $\psi^{(n)} : \mathbb{C} - \{0, -1, -2, \dots\} \rightarrow \mathbb{C}$, que é definida como

$$\operatorname{PolyGamma}(n, z) = \psi^{(n)}(z) = \frac{d^n}{dz^n} \psi(z), \quad (3.7)$$

onde $n = 0, 1, \dots$.

Definição 3.2.6. Função Zeta Generalizada. A função Zeta Generalizada $\zeta : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ é definida como

$$\zeta(s, a) = \sum_{\substack{i=0 \\ i \neq -a}}^{\infty} (a+i)^{-s}. \quad (3.8)$$

3.2.1. Casos particulares

Alguns casos particulares das funções definidas anteriormente serão dados a seguir.

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \quad (3.9)$$

$$\Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2} \quad (3.10)$$

$$\text{Constante de Euler} = \gamma = -\psi(1) = 0.57721566... \quad (3.11)$$

3.2.2. Propriedades

Algumas propriedades das funções definidas serão dadas a seguir.

$$\Gamma(n) = (n-1)!, \quad n \in \{1, 2, \dots\}, \quad 0! = 1 \quad (3.12)$$

$$\Gamma(x+1) = x\Gamma(x), \quad x \neq 0, -1, -2, \dots \quad (3.13)$$

$$2^{2x-1}\Gamma(x)\Gamma\left(x+\frac{1}{2}\right) = \sqrt{\pi}\Gamma(2x) \quad (3.14)$$

$$\Gamma(mx) = (2\pi)^{\frac{1-m}{2}} m^{mx-\frac{1}{2}} \prod_{j=0}^{m-1} \Gamma\left(x + \frac{j}{m}\right), \quad m = 1, 2, \dots \quad (3.15)$$

$$\Gamma\left(n + \frac{p}{q}\right) = \frac{1}{q^n} \Gamma\left(\frac{p}{q}\right) \prod_{j=0}^{n-1} (p + jq), \quad n \in \{0, 1, 2, \dots\}, \quad \frac{p}{q} \neq 0, -1, -2, \dots \quad (3.16)$$

$$\psi(x) = \psi(1) + (x-1) \sum_{i=0}^{\infty} [(i+1)(x+i)]^{-1}, \quad x \neq 0, -1, -2, \dots \quad (3.17)$$

$$\psi(n) = \psi(1) + 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n-1}, \quad n \in \{1, 2, \dots\} \quad (3.18)$$

$$\psi^{(n)}(z) = (-1)^{n+1} n! \zeta(n+1, z), \quad n \in \{1, 2, \dots\}, \quad \text{Re}(z) > 0 \quad (3.19)$$

Comentário 3.2.1. A propriedade (3.14) é a fórmula de duplicação de Legendre e (3.15) é a fórmula de multiplicação de Gauss. A propriedade (3.16) pode ser encontrada em Prudnikov *et al.* (1990, p.759).

3.3. A distribuição Erlang e suas distribuições relacionadas

Definição 3.3.1. Função densidade de probabilidade Gama. A função densidade de probabilidade Gama, com um parâmetro de escala a e um parâmetro de forma b , é a função $\gamma_b^a : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ definida como

$$\gamma_b^a(x) = \frac{a^b x^{b-1}}{\Gamma(b)e^{ax}}, \quad (3.20)$$

onde $a > 0$, $b > 0$ e $x \geq 0$.

A notação $\gamma_b^a(x)$ foi escolhida por economizar espaço horizontal no texto e por ser análoga à notação da função base de Bernstein-Bézier.

Definição 3.3.2. Função densidade de probabilidade Erlang. Se o parâmetro de forma da densidade Gama é um inteiro positivo, então ela é chamada de Erlang e é a função $\gamma_b^a : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ definida como

$$\gamma_b^a(x) = \frac{a^b x^{b-1}}{\Gamma(b)e^{ax}}, \quad (3.21)$$

onde $a > 0$, $b \in \{1, 2, \dots\}$ e $x \geq 0$.

Definição 3.3.3. Função densidade de probabilidade Exponencial. Se $b = 1$, então se obtém a densidade Exponencial, que é a função $f_a : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ definida como

$$f_a(x) = \gamma_1^a(x) = \frac{a}{e^{ax}}, \quad (3.22)$$

onde $a > 0$ e $x \geq 0$.

Definição 3.3.4. Função densidade de probabilidade Gama Generalizada. A densidade Gama Generalizada é a função $\gamma_{b,c}^a : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ definida como

$$\gamma_{b,c}^a(x) = \frac{|c| a^b x^{bc-1}}{\Gamma(b)e^{ax^c}}, \quad (3.23)$$

onde $a > 0$, $b > 0$, $c \neq 0$ e $x \geq 0$.

A distribuição Gama Generalizada foi descoberta várias vezes por pesquisadores de diferentes áreas ao longo do século XX; veja Hegyi (1999). Ela tem como casos particulares as distribuições Gama, Erlang, Exponencial, Chi-Quadrado, Chi, Maxwell-Boltzmann, Rayleigh, Normal Generalizada, Meia Normal, Weibull, o caso particular mais conhecido da Lévy e a Gama Generalizada Inversa. Tem como casos limites a Lognormal e a Pareto.

Durante o desenvolvimento desta dissertação, essas e algumas outras distribuições interessantes foram estudadas, mas não serão definidas aqui porque não serão usadas.

3.3.1. Propriedades

A seguir, serão apresentadas algumas propriedades importantes das densidades Gama e Erlang. É necessário lembrar que $a > 0$, e que $b > 0$ para a Gama e $b \in \{1, 2, \dots\}$ para a Erlang.

$$\gamma_b^a(\lambda x) = \frac{1}{\lambda} \gamma_b^{a\lambda}(x), \lambda > 0 \quad (3.24)$$

$$\frac{\partial \gamma_b^a(x)}{\partial x} = a(\gamma_{b-1}^a(x) - \gamma_b^a(x)) \quad (3.25)$$

$$\frac{\partial \gamma_b^a(x)}{\partial a} = \frac{b}{a}(\gamma_b^a(x) - \gamma_{b+1}^a(x)) \quad (3.26)$$

$$\int_{0+}^y \gamma_b^a(\lambda x) dx = \frac{1}{\lambda} - \frac{1}{a\lambda} \sum_{j=1}^b \gamma_j^a(\lambda y), b \in \{1, 2, \dots\}, \lambda > 0 \quad (3.27)$$

$$E(\gamma_b^a(x)) = \frac{b}{a} \quad (3.28)$$

$$V(\gamma_b^a(x)) = \frac{b}{a^2} \quad (3.29)$$

$$\int_0^\infty \ln(x) \gamma_b^a(x) dx = \psi(b) - \ln(a) \quad (3.30)$$

$$\int x^n e^{ax} dx = \frac{e^{ax}}{a} \left(x^n - \frac{nx^{n-1}}{a} + \frac{n(n-1)x^{n-2}}{a^2} - \dots + \frac{(-1)^n n!}{a^n} \right) + c, n \in \{0, 1, 2, \dots\}, c \in \mathbb{R} \quad (3.31)$$

$$\begin{aligned} \Gamma(b, x) &= \int_x^\infty t^{b-1} e^{-t} dt = e^{-x} (x^{b-1} + (b-1)x^{b-2} + (b-1)(b-2)x^{b-3} \dots + (b-1)!) \\ &= \Gamma(b) \sum_{i=1}^b \gamma_i^1(x), b \in \{1, 2, \dots\}, x \geq 0 \end{aligned} \quad (3.32)$$

3.4. Mistura e produto tensorial

Uma mistura é uma distribuição univariada construída a partir da soma ponderada de distribuições. Os pesos devem ter soma unitária e podem ser negativos, mas como qualquer distribuição, a densidade da mistura deve ser não-negativa sobre seu domínio. Se o número de parcelas é finito, então a mistura é chamada de mistura finita. Usualmente, mas não sempre, as distribuições componentes são da mesma família.

Definição 3.4.1. Mistura finita. A mistura finita é a distribuição de probabilidades cuja densidade é uma função $f : \mathbb{D} \rightarrow \mathbb{R}_+^*$ dada por

$$f(x) = \sum_{i=1}^n p_i f_i(x; \theta_i, \lambda_i, \dots), \quad (3.33)$$

onde $n \in \mathbb{N}$ é o número de parcelas na mistura tal que $1 \leq n < \infty$; $p_i \in \mathbb{R}$, para $1 \leq i \leq n$, são pesos tais que $\sum_{i=1}^n p_i = 1$; e $f_i : \mathbb{D} \rightarrow \mathbb{R}_+^*$ são densidades de probabilidade com parâmetros $\theta_i, \lambda_i, \dots$, para $1 \leq i \leq n$.

O conceito de mistura pode ser estendido para múltiplas dimensões por meio da soma ponderada de distribuições multivariadas. Um caso particular de mistura multivariada é o produto tensorial, que é uma distribuição multivariada construída a partir da soma ponderada do produto de D distribuições univariadas, onde D é o número de variáveis aleatórias. Se $D = 1$, então o produto tensorial se reduz a uma mistura.

Definição 3.4.2. Produto tensorial finito. *O produto tensorial finito é a distribuição de probabilidades cuja densidade é uma função $m : (\mathbb{D}_1 \times \dots \times \mathbb{D}_D) \rightarrow \mathbb{R}_+^*$ dada por*

$$m(x_1, \dots, x_D) = \sum_{i_1=1}^{n_1} \dots \sum_{i_D=1}^{n_D} p_{i_1, \dots, i_D} f_{1, i_1, \dots, i_D}(x_1; \theta_{1, i_1, \dots, i_D}, \lambda_{1, i_1, \dots, i_D}, \dots) \dots f_{D, i_1, \dots, i_D}(x_D; \theta_{D, i_1, \dots, i_D}, \lambda_{D, i_1, \dots, i_D}, \dots), \quad (3.34)$$

onde $D \in \mathbb{N}$ é o número de variáveis aleatórias; $1 \leq n_d < \infty$, para $1 \leq d \leq D$; $p_{i_1, \dots, i_D} \in \mathbb{R}$, para $1 \leq i_d \leq n_d$ e $1 \leq d \leq D$, são pesos tais que

$$\sum_{i_1=1}^{n_1} \dots \sum_{i_D=1}^{n_D} p_{i_1, \dots, i_D} = 1;$$

e $f_{d, i_1, \dots, i_D} : \mathbb{D}_d \rightarrow \mathbb{R}_+^*$ são densidades de probabilidade univariadas com parâmetros $\theta_{d, i_1, \dots, i_D}, \lambda_{d, i_1, \dots, i_D}, \dots$, para $1 \leq i_d \leq n_d$ e $1 \leq d \leq D$.

Nesta dissertação, serão usados a mistura finita e produto tensorial finito de Erlangs. Como muitos pesos serão iguais a zero, uma notação mais compacta é preferível, que será dada a seguir.

Definição 3.4.3. Definição alternativa para o produto tensorial finito de Erlangs. *O produto tensorial finito de Erlangs é a distribuição de probabilidades cuja densidade $m : (\mathbb{R}_+^*)^D \rightarrow \mathbb{R}_+^*$ é dada por*

$$m(x_1, \dots, x_D) = \sum_{i=1}^n p_i \gamma_{b_{i,1}}^{a_{i,1}}(x_1) \dots \gamma_{b_{i,D}}^{a_{i,D}}(x_D), \quad (3.35)$$

onde $D \in \mathbb{N}$ é o número de variáveis aleatórias; $n \in \mathbb{N}$ é o número de parcelas tal que $1 \leq n < \infty$; $p_i \in \mathbb{R}$, para $1 \leq i \leq n$, são pesos tais que $\sum_{i=1}^n p_i = 1$; e $\gamma_{b_{i,d}}^{a_{i,d}}(x_d)$ são densidades Erlang tais que $x_d \geq 0$, $a_{i,d} > 0$ e $b_{i,d} \in \{1, 2, \dots\}$, para todos os i e d , $1 \leq i \leq n$, $1 \leq d \leq D$.

Se os parâmetros de escala das Erlangs são constantes para cada variável aleatória, isto é, $a_{i,d} = a_d$, para todos os i e d , $1 \leq i \leq n$, $1 \leq d \leq D$, então é possível definir o produto tensorial finito de Erlangs com parâmetros de escala constantes. Isto será feito a seguir.

Definição 3.4.4. Produto tensorial finito de Erlangs com parâmetros de escala constantes. *O produto tensorial finito de Erlangs com parâmetros de escala constantes é a distribuição de probabilidades cuja densidade $m : (\mathbb{R}_+^*)^D \rightarrow \mathbb{R}_+^*$ é dada por*

$$m(x_1, \dots, x_D) = \sum_{i=1}^n p_i \gamma_{b_{i,1}}^{a_1}(x_1) \dots \gamma_{b_{i,D}}^{a_D}(x_D), \quad (3.36)$$

onde $D \in \mathbb{N}$ é o número de variáveis aleatórias; $n \in \mathbb{N}$ é o número de parcelas tal que $1 \leq n < \infty$; $p_i \in \mathbb{R}$, para $1 \leq i \leq n$, são pesos tais que $\sum_{i=1}^n p_i = 1$; e $\gamma_{b_{i,d}}^{a_d}(x_d)$ são densidades Erlang tais que $x_d \geq 0$, $a_d > 0$ e $b_{i,d} \in \{1, 2, \dots\}$, para todos os i e d , $1 \leq i \leq n$, $1 \leq d \leq D$.

O produto tensorial finito de Erlangs com parâmetros de escala constantes será usado neste trabalho para aproximar a densidade multivariada uniperiódica dos retornos dependentes dos ativos.

Como a mistura finita de Erlangs pode ter um número arbitrário de parcelas, então ela pode aproximar, com precisão arbitrária, qualquer densidade contínua definida sobre um intervalo fechado dos reais não-negativos. Este é um teorema análogo ao de Weierstrass, cuja prova pode ser encontrada em Powell (1981, p.61).

3.5. Calibração de mistura e produto tensorial

Dados uma função distribuição de probabilidades e um conjunto de observações ou amostras, uma tarefa comum em Estatística é o ajuste ou calibração dos parâmetros da distribuição para que esta represente as amostras. Isto pode ser feito para os casos uni e multivariado. É necessário definir um critério de ajuste. Existem vários critérios disponíveis tais como a minimização do erro quadrático médio e o método dos momentos. Nesta dissertação, será usado o critério da máxima verossimilhança. Quanto maior a verossimilhança, melhor o ajuste da densidade às amostras. Uma hipótese necessária para a utilização deste critério é que deve haver independência entre as amostras, mas é possível que haja dependência entre as variáveis no caso multivariado.

Um resultado relativo a misturas e produtos tensoriais finitos é que se os parâmetros das distribuições são constantes, isto é, somente os pesos são variáveis no problema de otimização, então a função verossimilhança é côncava sobre eles.

Teorema 3.5.1. Concavidade da função verossimilhança de uma mistura de distribuições com parâmetros fixos. Sejam $x_1, x_2, \dots, x_S$, onde $S \geq 1$, observações unidimensionais independentes e uma mistura com densidade $f : \mathbb{D} \rightarrow \mathbb{R}_+^*$ dada por

$$f(x) = \sum_{i=1}^n p_i f_i(x; \theta_i, \lambda_i, \dots),$$

onde $1 \leq n \leq \infty$; $p_i \in \mathbb{R}$, para $1 \leq i \leq n$, são pesos tais que $\sum_{i=1}^n p_i = 1$; e $f_i : \mathbb{D} \rightarrow \mathbb{R}_+^*$ são densidades com parâmetros $\theta_i, \lambda_i, \dots$, para $1 \leq i \leq n$. Se esses parâmetros são constantes, então a função verossimilhança

$$L(p_1, \dots, p_n; x_1, \dots, x_S, f_1, \dots, f_n, \theta_1, \dots, \theta_n, \lambda_1, \dots, \lambda_n, \dots) = \prod_{s=1}^S \sum_{i=1}^n p_i f_i(x_s; \theta_i, \lambda_i, \dots)$$

é côncava sobre os pesos p_i .

Prova. Sejam os vetores no espaço de pesos

$$v_1 = [p_{1,1}, p_{1,2}, \dots, p_{1,n}]^T, \quad v_2 = [p_{2,1}, p_{2,2}, \dots, p_{2,n}]^T \text{ e}$$

$$v_3 = \frac{v_1 + v_2}{2} = \left[\frac{p_{1,1} + p_{2,1}}{2}, \frac{p_{1,2} + p_{2,2}}{2}, \dots, \frac{p_{1,n} + p_{2,n}}{2} \right]^T = [p_{3,1}, p_{3,2}, \dots, p_{3,n}]^T.$$

Como o domínio dos pesos é convexo, v_3 pertence ao domínio. Os valores da função verossimilhança nos

vetores v_1 , v_2 e v_3 são, respectivamente,

$$\begin{aligned} L_1 &= L(v_1; x_s, f_i, \theta_i, \lambda_i, \dots) = \prod_{s=1}^S \sum_{i=1}^n p_{1,i} f_i(x_s; \theta_i, \lambda_i, \dots), \\ L_2 &= L(v_2; x_s, f_i, \theta_i, \lambda_i, \dots) = \prod_{s=1}^S \sum_{i=1}^n p_{2,i} f_i(x_s; \theta_i, \lambda_i, \dots), \\ L_3 &= L(v_3; x_s, f_i, \theta_i, \lambda_i, \dots) = \prod_{s=1}^S \sum_{i=1}^n p_{3,i} f_i(x_s; \theta_i, \lambda_i, \dots). \end{aligned}$$

Os termos $f_i(x_s; \mu_i, \sigma_i, \dots)$ são constantes. Como os vetores v_1 e v_2 são arbitrários, é suficiente provar que

$$L_3 \geq \frac{(L_1 + L_2)}{2} \quad (3.37)$$

ou que (porque o logaritmo natural é uma função crescente)

$$\ln(L_3) \geq \frac{\ln(L_1) + \ln(L_2)}{2} = \ln(\sqrt{L_1 L_2}). \quad (3.38)$$

Qualquer vetor do tipo $kv_1 + (1-k)v_2$, $0 \leq k \leq 1$, pode ser atingido por bissecções sucessivas do intervalo entre v_1 e v_2 . Removendo o logaritmo de (3.38), é necessário provar que

$$L_3 \geq \sqrt{L_1 L_2}$$

ou

$$\prod_{s=1}^S \sum_{i=1}^n \frac{p_{1,i} + p_{2,i}}{2} f_i(x_s; \theta_i, \lambda_i, \dots) \geq \sqrt{\left(\prod_{s=1}^S \sum_{i=1}^n p_{1,i} f_i(x_s; \theta_i, \lambda_i, \dots) \right) \left(\prod_{s=1}^S \sum_{i=1}^n p_{2,i} f_i(x_s; \theta_i, \lambda_i, \dots) \right)}$$

ou

$$\begin{aligned} \prod_{s=1}^S \frac{\sum_{i=1}^n p_{1,i} f_i(x_s; \theta_i, \lambda_i, \dots) + \sum_{i=1}^n p_{2,i} f_i(x_s; \theta_i, \lambda_i, \dots)}{2} &\geq \\ \prod_{s=1}^S \sqrt{\left(\sum_{i=1}^n p_{1,i} f_i(x_s; \theta_i, \lambda_i, \dots) \right) \left(\sum_{i=1}^n p_{2,i} f_i(x_s; \theta_i, \lambda_i, \dots) \right)} &. \end{aligned} \quad (3.39)$$

Para cada índice s , há dois números não-negativos em cada lado de (3.39). Como a média aritmética de números não-negativos é sempre maior ou igual à média geométrica, a seguinte desigualdade é verdadeira.

$$\frac{\sum_{i=1}^n p_{1,i} f_i(x_s; \theta_i, \lambda_i, \dots) + \sum_{i=1}^n p_{2,i} f_i(x_s; \theta_i, \lambda_i, \dots)}{2} \geq \sqrt{\left(\sum_{i=1}^n p_{1,i} f_i(x_s; \theta_i, \lambda_i, \dots) \right) \left(\sum_{i=1}^n p_{2,i} f_i(x_s; \theta_i, \lambda_i, \dots) \right)}.$$

Multiplicando-se os S termos, o sinal da desigualdade é mantido e a expressão (3.39) é obtida.

□

Corolário 3.5.1. Concavidade da função verossimilhança do produto tensorial de distribuições com parâmetros fixos. Sejam $x_1 = [x_{1,1}, \dots, x_{1,D}]^\top$, $x_2 = [x_{2,1}, \dots, x_{2,D}]^\top$, ..., $x_S = [x_{S,1}, \dots, x_{S,D}]^\top$, onde $S \geq 1$, observações independentes com $D \geq 1$ variáveis aleatórias e um produto tensorial com uma densidade $m : (\mathbb{D}_1 \times \dots \times \mathbb{D}_D) \rightarrow \mathbb{R}_+^*$ dada por

$$m(x_1, \dots, x_D) = \sum_{i_1=1}^m \dots \sum_{i_D=1}^{n_D} p_{i_1, \dots, i_D} f_{1, i_1, \dots, i_D}(x_1; \theta_{1, i_1, \dots, i_D}, \lambda_{1, i_1, \dots, i_D}, \dots) \dots f_{D, i_1, \dots, i_D}(x_D; \theta_{D, i_1, \dots, i_D}, \lambda_{D, i_1, \dots, i_D}, \dots),$$

onde $1 \leq n_d \leq \infty$ para $1 \leq d \leq D$; $p_{i_1, \dots, i_D} \in \mathbb{R}$, para $1 \leq i_d \leq n_d$ e $1 \leq d \leq D$, são pesos tais que

$$\sum_{i_1=1}^m \dots \sum_{i_D=1}^{n_D} p_{i_1, \dots, i_D} = 1;$$

e $f_{d, i_1, \dots, i_D} : \mathbb{D}_d \rightarrow \mathbb{R}_+^*$ são densidades de probabilidade univariadas com parâmetros $\theta_{d, i_1, \dots, i_D}, \lambda_{d, i_1, \dots, i_D}, \dots$, para $1 \leq i_d \leq n_d$ e $1 \leq d \leq D$. Se esses parâmetros são constantes, então a função verossimilhança

$$L(p_{i_1, \dots, i_D}; x_s, f_{1, i_1, \dots, i_D}, \theta_{1, i_1, \dots, i_D}, \lambda_{1, i_1, \dots, i_D}, \dots) = \prod_{s=1}^S \sum_{i_1=1}^{n_1} \dots \sum_{i_D=1}^{n_D} p_{i_1, \dots, i_D} f_{1, i_1, \dots, i_D}(x_{s,1}; \theta_{1, i_1, \dots, i_D}, \lambda_{1, i_1, \dots, i_D}, \dots) \dots f_{D, i_1, \dots, i_D}(x_{s,D}; \theta_{D, i_1, \dots, i_D}, \lambda_{D, i_1, \dots, i_D}, \dots)$$

é côncava sobre os pesos $p_{i_1, \dots, i_D}$.

Prova. Para cada índice s , o produto $f_{1, i_1, \dots, i_D}(x_{s,1}; \theta_{1, i_1, \dots, i_D}, \lambda_{1, i_1, \dots, i_D}, \dots) \dots f_{D, i_1, \dots, i_D}(x_{s,D}; \theta_{D, i_1, \dots, i_D}, \lambda_{D, i_1, \dots, i_D}, \dots)$ é um número constante e a prova é análoga à do teorema anterior.

□

O corolário e o teorema anteriores podem ser provados com a desigualdade de Jensen; veja Boyd e Vandenberghe (2004, p.77, 79).

A escolha da distribuição e dos parâmetros a serem otimizados é crucial para a factibilidade computacional da maximização da verossimilhança. A escolha de uma densidade multivariada em forma de produto tensorial leva a uma explosão do número de variáveis a serem otimizadas. Este efeito é conhecido como “maldição da dimensionalidade”. Mas o fato de que o oposto da verossimilhança é uma função convexa sobre os pesos não-negativos da mistura alivia o problema do esforço computacional. O oposto da verossimilhança é, em geral, não-convexo sobre os parâmetros das distribuições e, portanto, é desejável o uso da menor quantidade de parâmetros possível.

Os pesos da mistura ou do produto tensorial podem ser negativos. Entretanto, nos problemas de otimização desta dissertação nos quais aparecem como variáveis de otimização, os pesos serão forçados a ser não-negativos porque garantir a não-negatividade da densidade é um problema difícil. É mais simples usar pesos não-negativos e aumentar o número de parcelas na mistura ou no produto tensorial.

A seguir, serão discutidas algumas variações de como o problema de otimização pode ser formulado para o caso unidimensional, isto é, para a mistura de densidades.

Problema de otimização 3.5.1. Calibração de uma mistura finita positiva de distribuições com parâmetros fixos. Existem S observações unidimensionais independentes. As distribuições componentes e

seus parâmetros são dados. Assumir ou forçar os pesos dos componentes da mistura a ser não-negativos garante a não-negatividade da densidade. A soma dos pesos deve ser 1 para que a área sob a função densidade seja igual a 1. O espaço dos pesos não-negativos é convexo. Mudar o sinal da função verossimilhança a torna convexa sobre os pesos. Portanto, este é um problema de otimização convexa.

$$\begin{aligned} \min_{p_i} -L(p_i; x_s, f_i, \mu_i, \sigma_i, \dots) &= -\prod_{s=1}^S \sum_{i=1}^n p_i f_i(x_s; \mu_i, \sigma_i, \dots) \\ \text{s.t.} \\ p_i &\geq 0, \quad 1 \leq i \leq n \\ \sum_{i=1}^n p_i &= 1 \end{aligned}$$

A vantagem de se formular problemas convexos é devido à solução ser única (desconsiderando-se as regiões com patamares) e à existência de algoritmos eficientes para resolvê-los.

Problema de otimização 3.5.2. Calibração de uma mistura finita positiva de distribuições Erlang. Os parâmetros de escala não são necessariamente os mesmos e podem variar de 0 até um dado valor máximo $a_{\max}$. Os parâmetros de forma são inteiros entre 1 e um dado $b_{\max}$.

$$\begin{aligned} \min_{a_i, p_i} -L(a_i, p_i) &= -\prod_{s=1}^S \sum_{i=1}^{b_{\max}} p_i \gamma_i^{a_i}(x_s) \\ \text{s.t.} \\ 0 < a_i &\leq a_{\max}, \quad 1 \leq i \leq b_{\max} \\ p_i &\geq 0, \quad 1 \leq i \leq b_{\max} \\ \sum_{i=1}^{b_{\max}} p_i &= 1 \end{aligned}$$

No problema de otimização anterior, os parâmetros de escala e de forma têm um valor máximo. Mas fixando o valor máximo de b , o parâmetro de escala a será automaticamente otimizado. Os parâmetros de escala podem ser forçados a ser iguais em todas as parcelas da mistura.

Problema de otimização 3.5.3. Calibração de uma mistura finita positiva de distribuições Erlang com parâmetros de escala constantes. Os parâmetros de escala de todos os componentes da mistura têm o mesmo valor. Os parâmetros de forma são inteiros entre 1 e um dado $b_{\max}$.

$$\begin{aligned} \min_{a, p_i} -L(a, p_i) &= -\prod_{s=1}^S \sum_{i=1}^{b_{\max}} p_i \gamma_i^a(x_s) \\ \text{s.t.} \\ a &> 0 \\ p_i &\geq 0, \quad 1 \leq i \leq b_{\max} \\ \sum_{i=1}^{b_{\max}} p_i &= 1 \end{aligned}$$

No problema de otimização anterior, a função objetivo é não-convexa somente sobre o parâmetro de escala a . O número de parcelas na mistura é $b_{\max}$. Testes numéricos indicam que somente alguns poucos pesos p_i serão diferentes de zero. O número de parcelas na mistura pode ser diminuído da forma a seguir.

Problema de otimização 3.5.4. Calibração de uma mistura finita positiva de distribuições Erlang com parâmetros de escala constantes. O parâmetro de escala é o mesmo para todos os componentes da mistura. Os parâmetros de forma são variáveis inteiras entre 1 e um dado $b_{\max}$. O número de parcelas n é também uma variável de otimização. Este é um problema de otimização mista não-convexa.

$$\begin{aligned} \min_{n,a,b_i,p_i} -L(n,a,b_i,p_i) &= -\prod_{s=1}^S \sum_{i=1}^n p_i \gamma_{b_i}^a(x_s) \\ \text{s.t.} \\ n &\in \{1,2,\dots\} \\ a &> 0 \\ b_i &\in \{1,\dots,b_{\max}\} \quad , \quad 1 \leq i \leq n \\ p_i &\geq 0 \quad , \quad 1 \leq i \leq n \\ \sum_{i=1}^n p_i &= 1 \end{aligned}$$

Para o caso multivariado, as mesmas variações dos problemas de otimização podem ser feitas. A seguir, será apresentada a formulação escolhida.

Problema de otimização 3.5.5. Calibração de um produto tensorial finito positivo de distribuições Erlang com parâmetros de escala constantes. Existem $D \geq 1$ variáveis aleatórias. O parâmetro de escala a_d é o mesmo para cada variável aleatória d , $1 \leq d \leq D$. Em cada variável aleatória, os parâmetros de forma são variáveis inteiras entre 1 e $b_{\max,d}$. O número de parcelas n é também uma variável. Este é um problema de otimização mista não-convexa.

$$\begin{aligned} \min_{n,a_d,b_{i,d},p_i} -L(n,a_d,b_{i,d},p_i) &= -\prod_{s=1}^S \sum_{i=1}^n p_i \gamma_{b_{i,1}}^{a_1}(x_{s,1}) \gamma_{b_{i,2}}^{a_2}(x_{s,2}) \dots \gamma_{b_{i,D}}^{a_D}(x_{s,D}) \\ \text{s.t.} \\ n &\in \{1,2,\dots\} \\ a_d &> 0 \quad , \quad 1 \leq d \leq D \\ b_{i,d} &\in \{1,\dots,b_{\max,d}\} \quad , \quad 1 \leq i \leq n, 1 \leq d \leq D \\ p_i &\geq 0 \quad , \quad 1 \leq i \leq n \\ \sum_{i=1}^n p_i &= 1 \end{aligned}$$

O valor de $b_{\max}$ é importante para o problema de otimização. O valor esperado e a variância de uma distribuição Erlang são

$$E(\gamma_b^a(x)) = \frac{b}{a} \text{ e } V(\gamma_b^a(x)) = \frac{b}{a^2}.$$

Se a média da variável aleatória é aproximadamente 1 e o desvio-padrão é aproximadamente 1%, então os valores a e b serão aproximadamente 10000. O valor de $b_{\max}$ pode ser, por exemplo, 15000.

Se o **Problema de otimização 3.5.3** for usado, então haverá 15000 pesos a serem otimizados. Se houver D variáveis aleatórias, então haverá 15000^D pesos a serem otimizados. Este efeito é conhecido por “maldição da dimensionalidade”. Mas este problema tem uma estrutura especial porque os pesos, em sua maioria, serão iguais a zero.

Se o **Problema de otimização 3.5.4** for usado, então a função objetivo será não-convexa sobre o parâmetro de escala e convexa sobre os n pesos, onde n é da ordem de grandeza de S ; e haverá n parâmetros de forma, que são variáveis de otimização inteiras.

No caso multivariado, com D variáveis aleatórias, a função objetivo do **Problema de otimização 3.5.5** será não-convexa sobre os D parâmetros de escala e convexa sobre os n pesos, onde n é da ordem de grandeza de S ; e haverá nD parâmetros de forma, que são variáveis de otimização inteiras.

O problema de minimização de uma função não-convexa é considerado difícil porque pode haver várias soluções ótimas locais. O problema de minimização de uma função convexa sobre um domínio convexo é considerado tratável; veja Boyd e Vandenberghe (2004, p.16).

Para resolver o **Problema de otimização 3.5.5**, foi necessário criar um algoritmo de otimização sob medida e implementá-lo computacionalmente. As variáveis são otimizadas iterativamente, com técnicas adequadas para cada tipo. Detalhes sobre o algoritmo estão fora do escopo desta dissertação.

Exemplo 3.5.1. Mistura finita de Erlangs. Os parâmetros de uma mistura univariada de Erlangs com parâmetros de escala constantes foram escolhidos arbitrariamente e alguns pontos foram sorteados. Esses pontos foram usados para calibrar os parâmetros de uma mistura finita de Erlangs com o critério da máxima verossimilhança. Então a mistura original e a calibrada foram comparadas.

A mistura de Erlangs escolhida foi

$$f_{orig}(x) = 0.3\gamma_{440}^{440}(x) + 0.3\gamma_{450}^{440}(x) + 0.4\gamma_{525}^{440}(x).$$

A seguir, $S = 40$ pontos foram sorteados desta mistura. Estes pontos estão listados no **Apêndice A**. Duas otimizações como a formulada no **Problema de otimização 3.5.4** foram feitas, uma com $b_{\max} = 525$ e outra com $b_{\max} = 750$. As misturas obtidas foram

$$f_{525}(x) = 0.1300\gamma_{448}^{443.37}(x) + 0.5107\gamma_{449}^{443.37}(x) + 0.3593\gamma_{525}^{443.37}(x)$$

e

$$f_{750}(x) = 0.1917\gamma_{607}^{621.40}(x) + 0.3712\gamma_{636}^{621.40}(x) + 0.0559\gamma_{637}^{621.40}(x) + 0.1494\gamma_{706}^{621.40}(x) + 0.2319\gamma_{750}^{621.40}(x).$$

Os logaritmos naturais das verossimilhanças obtidas foram

$$\text{LogLikelihood}_{525} = 42.306195$$

e

$$\text{LogLikelihood}_{750} = 42.588620.$$

Estas três densidades e os pontos sorteados podem ser vistos juntos na figura a seguir.

Figura 3.5.1. Calibração de uma mistura finita de distribuições Erlang com parâmetros de escala constantes. Os 40 pontos indicados no eixo x foram sorteados a partir de f_{orig} e usados para calibrar f_{525} e f_{750} .

Exemplo 3.5.2. Produto tensorial finito de Erlangs. Um produto tensorial finito bivariado de Erlangs foi escolhido arbitrariamente e alguns pontos foram sorteados. Estes pontos foram usados para calibrar os parâmetros de um produto tensorial finito de Erlangs com o critério da máxima verossimilhança. Então o produto tensorial escolhido e o calibrado foram comparados.

O produto tensorial escolhido foi

$$m_{orig}(x_1, x_2) = 0.4\gamma_{440}^{440}(x_1)\gamma_{300}^{300}(x_2) + 0.3\gamma_{500}^{440}(x_1)\gamma_{375}^{300}(x_2) + 0.3\gamma_{550}^{440}(x_1)\gamma_{300}^{300}(x_2).$$

A seguir, $S = 60$ pontos foram sorteados deste produto tensorial. Estes pontos estão listados no **Apêndice A**. Duas otimizações como a formulada no **Problema de otimização 3.5.5** foram feitas, uma com $b_{\max,1} = b_{\max,2} = 550$ e outra com $b_{\max,1} = b_{\max,2} = 750$. As misturas obtidas foram

$$\begin{aligned} m_{550}(x_1, x_2) = & 0.3431\gamma_{441}^{438.26}(x_1)\gamma_{441}^{431.16}(x_2) + 0.0407\gamma_{449}^{438.26}(x_1)\gamma_{418}^{431.16}(x_2) + 0.0341\gamma_{450}^{438.26}(x_1)\gamma_{417}^{431.16}(x_2) + \\ & 0.1184\gamma_{506}^{438.26}(x_1)\gamma_{550}^{431.16}(x_2) + 0.0847\gamma_{507}^{438.26}(x_1)\gamma_{550}^{431.16}(x_2) + 0.0395\gamma_{509}^{438.26}(x_1)\gamma_{521}^{431.16}(x_2) + \\ & 0.0392\gamma_{541}^{438.26}(x_1)\gamma_{511}^{431.16}(x_2) + 0.0020\gamma_{542}^{438.26}(x_1)\gamma_{511}^{431.16}(x_2) + 0.0875\gamma_{550}^{438.26}(x_1)\gamma_{399}^{431.16}(x_2) + \\ & 0.2107\gamma_{550}^{438.26}(x_1)\gamma_{433}^{431.16}(x_2) \end{aligned}$$

e

$$\begin{aligned} m_{750}(x_1, x_2) = & 0.2366\gamma_{590}^{595.77}(x_1)\gamma_{598}^{582.44}(x_2) + 0.0182\gamma_{600}^{595.77}(x_1)\gamma_{610}^{582.44}(x_2) + 0.1102\gamma_{613}^{595.77}(x_1)\gamma_{559}^{582.44}(x_2) + \\ & 0.0523\gamma_{629}^{595.77}(x_1)\gamma_{613}^{582.44}(x_2) + 0.0584\gamma_{685}^{595.77}(x_1)\gamma_{700}^{582.44}(x_2) + 0.1789\gamma_{688}^{595.77}(x_1)\gamma_{750}^{582.44}(x_2) + \\ & 0.0097\gamma_{695}^{595.77}(x_1)\gamma_{587}^{582.44}(x_2) + 0.0067\gamma_{738}^{595.77}(x_1)\gamma_{691}^{582.44}(x_2) + 0.0393\gamma_{739}^{595.77}(x_1)\gamma_{691}^{582.44}(x_2) + \\ & 0.0763\gamma_{748}^{595.77}(x_1)\gamma_{537}^{582.44}(x_2) + 0.0196\gamma_{749}^{595.77}(x_1)\gamma_{538}^{582.44}(x_2) + 0.0075\gamma_{750}^{595.77}(x_1)\gamma_{540}^{582.44}(x_2) + \\ & 0.0876\gamma_{750}^{595.77}(x_1)\gamma_{589}^{582.44}(x_2) + 0.0986\gamma_{750}^{595.77}(x_1)\gamma_{590}^{582.44}(x_2). \end{aligned}$$

Os logaritmos naturais das verossimilhanças obtidas foram

$$\text{LogLikelihood}_{550} = 125.292272$$

e

$$\text{LogLikelihood}_{750} = 128.840960 .$$

Estas três densidades e os pontos sorteados podem ser vistos na figura a seguir.

Figura 3.5.2. Calibração de um produto tensorial finito bivariado de Erlangs com parâmetros de escala constantes. A função densidade de probabilidade multivariada $m_{orig}(x_1, x_2)$ foi escolhida arbitrariamente, 60 pontos foram sorteados e usados para calibrar as densidades $m_{550}(x_1, x_2)$ e $m_{750}(x_1, x_2)$ com o critério da máxima verossimilhança.

■

Os pontos aleatórios do exemplo anterior foram obtidos em dois passos. Inicialmente, uma das parcelas do produto tensorial foi sorteada. No exemplo, há 3 parcelas, com probabilidades 0.4, 0.3 e 0.3. Em seguida, como o produto de Erlangs de cada parcela é uma densidade multivariada formada por um produto de variáveis independentes, foi sorteado um número entre 0 e 1 de uma distribuição uniforme para cada

variável aleatória e, então, calculada a inversa de sua função acumulada. Este procedimento é justificado com uma árvore de probabilidades, e vale para um produto tensorial finito de distribuições univariadas arbitrárias.

Como pode ser visto nos dois exemplos anteriores, quando S é finito, cada variável da solução sempre terá pelo menos um componente Erlang com seu respectivo $b_{\max}$ porque, caso contrário, o resultado seria a solução de um problema com menos graus de liberdade.

Comentário 3.5.1. As soluções obtidas são locais.

3.6. Número de amostras e erro de amostragem

No **Problema de otimização 3.5.5**, o número de variáveis a serem otimizadas é muito maior que o número de amostras. À primeira vista, isto parece ser um problema, mas, como foi mostrado nos exemplos, o número de variáveis diferentes de zero “se adapta” e se torna da mesma ordem de grandeza do número de observações.

Um problema com critérios de ajuste é o erro de amostragem. Se o número de amostras é finito (ou menor que a população), então os parâmetros estimados da distribuição podem ser tendenciosos. Por exemplo, no caso da distribuição normal univariada, se existirem poucas amostras, então, na média, o critério da máxima verossimilhança dará baixa importância às caudas porque há poucas amostras lá. Pelo critério, na média, as caudas serão mais finas do que elas realmente são, e a variância será menor do que ela realmente é. Este efeito pode ser mostrado com um experimento Monte Carlo simples. Por causa da simetria, a média não é tendenciosa. Mas o estimador da variância pode ser melhorado. Existem diferentes formas de se fazer isso, mas a maneira usual é multiplicá-lo pelo fator $n/(n-1)$, onde n é o número de amostras. Este problema diminui com o aumento do número de amostras.

Nesta dissertação, não será tentada a obtenção de estimadores melhores (menos tendenciosos ou mais eficientes) para os parâmetros da mistura e do produto tensorial de Erlangs. Este ainda é um problema de pesquisa.

3.7. As transformadas de Laplace e de Mellin

As transformadas integrais de Laplace e de Mellin, juntamente com os teoremas da convolução nos casos aditivo e multiplicativo são úteis para se determinar, respectivamente, a densidade resultante da soma e do produto de variáveis aleatórias independentes.

Comentário 3.7.1. Alguns autores chamam a convolução aditiva de *convolution product*.

Definição 3.7.1. Transformada de Laplace. A transformada de Laplace de uma função $f(x)$ é a função $F(s)$ definida como

$$f(x) \xrightarrow{\mathcal{L}} F(s) = \mathcal{L}\{f(x); s\} = \int_{0+}^{\infty} e^{-sx} f(x) dx. \quad (3.40)$$

Teorema 3.7.1. Convolução aditiva. Se

$$f(x) \xrightarrow{\mathcal{L}} F(s), \quad g(x) \xrightarrow{\mathcal{L}} G(s),$$

$f(x)$ e $g(x)$ têm como domínio o conjunto dos reais não-negativos e $F(s)$ e $G(s)$ convergem absolutamente em $s = a$, então

$$(f \oplus g)(x) = \int_{0+}^{x-} f(t)g(x-t)dt \xrightarrow{\mathcal{L}} F(s)G(s), \quad s \geq a \quad (3.41)$$

Prova. Veja Widder (1989, p.453). □

Se o domínio das funções $f(x)$ e $g(x)$ for o conjunto dos reais, então o limite inferior da integral (3.41) deve ser substituído por $-\infty$.

Corolário 3.7.1. Soma de variáveis aleatórias independentes. Se X_1 e X_2 são variáveis aleatórias independentes definidas sobre o domínio dos reais não-negativos, e tais que

$$X_1 \sim f_1(x) \xrightarrow{\mathcal{L}} F_1(s) \text{ e } X_2 \sim f_2(x) \xrightarrow{\mathcal{L}} F_2(s),$$

então

$$X_1 + X_2 \sim (f_1 \oplus f_2)(x) = \int_0^x f_1(t)f_2(x-t)dt \xrightarrow{\mathcal{L}} F_1(s)F_2(s).$$

Prova. Aplicação direta do teorema anterior. □

Definição 3.7.2. Transformada de Mellin. A transformada de Mellin de uma função $f(x)$ é a função $F(s)$ definida como

$$f(x) \xrightarrow{\mathcal{M}} F(s) = \mathcal{M}\{f(x); s\} = \int_0^\infty x^{s-1}f(x)dx. \quad (3.42)$$

Teorema 3.7.2. Convolução multiplicativa ou de escala. Se

$$f(x) \xrightarrow{\mathcal{M}} F(s), \quad g(x) \xrightarrow{\mathcal{M}} G(s),$$

e $f(x)$ e $g(x)$ têm como domínio o conjunto dos reais não-negativos, então

$$(f \otimes g)(x) = \int_0^\infty f(t)g\left(\frac{x}{t}\right)\frac{1}{t}dt \xrightarrow{\mathcal{M}} F(s)G(s).$$

Prova. Veja Ditkin e Prudnikov (1965, p.70) ou Sneddon (1972, p.270). □

Corolário 3.7.2. Produto de variáveis aleatórias independentes. Se X_1 e X_2 são variáveis aleatórias independentes definidas sobre o domínio dos reais não-negativos, e tais que

$$X_1 \sim f_1(x) \xrightarrow{\mathcal{M}} F_1(s) \text{ e } X_2 \sim f_2(x) \xrightarrow{\mathcal{M}} F_2(s),$$

então

$$X_1 X_2 \sim (f_1 \otimes f_2)(x) = \int_0^\infty f_1(t) f_2\left(\frac{x}{t}\right) \frac{1}{t} dt \xrightarrow{\mathcal{M}} F_1(s) F_2(s).$$

Prova. Aplicação direta do teorema anterior. □

A importância dessas transformadas se deve ao teorema da convolução nos casos aditivo e multiplicativo. Para diminuir o esforço computacional neste trabalho, é desejável que a soma e a combinação linear de variáveis aleatórias que seguem uma distribuição $\mathcal{P}_1$ resulte sempre em uma que segue uma distribuição $\mathcal{P}_2$; e que o produto de variáveis aleatórias que seguem a distribuição $\mathcal{P}_2$ resulte sempre em uma que segue uma distribuição $\mathcal{P}_3$. Considerando-se o caso especial com apenas uma parcela ou termo, é desejável que, para essas classes de distribuições, ocorra $\{\mathcal{P}_1\} \subseteq \{\mathcal{P}_2\} \subseteq \{\mathcal{P}_3\}$.

3.7.1. Casos particulares

A seguir, serão apresentados alguns casos particulares das transformadas de Laplace e de Mellin. As funções Meijer G e Fox H serão definidas na seção seguinte.

$$\gamma_b^a(x) = \frac{a^b x^{b-1}}{\Gamma(b) e^{ax}} = \frac{a}{\Gamma(b)} G_{0,1}^{1,0} \left(ax \left| \begin{matrix} - \\ b-1 \end{matrix} \right. \right) \xrightarrow{\mathcal{L}} a^b (a+s)^{-b}, \quad (3.43)$$

$$x \geq 0, a > 0, b > 0, \operatorname{Re}(s) \geq 0$$

$$\gamma_{b,c}^a(x) = \frac{ca^b x^{bc-1}}{\Gamma(b) e^{ax^c}} = \frac{a^{\frac{1}{c}}}{\Gamma(b)} H_{0,1}^{1,0} \left(a^{\frac{1}{c}} x \left| \begin{matrix} - \\ b-\frac{1}{c}, \frac{1}{c} \end{matrix} \right. \right) \xrightarrow{\mathcal{L}} \frac{1}{\Gamma(b)} H_{1,1}^{1,1} \left(\frac{s}{a^{\frac{1}{c}}} \left| \begin{matrix} 1-b, \frac{1}{c} \\ (0,1) \end{matrix} \right. \right), \quad (3.44)$$

$$x \geq 0, a > 0, b > 0, c > 0, \operatorname{Re}(s) \geq 0$$

$$\delta(x-c) \xrightarrow{\mathcal{L}} e^{-cs}, \quad x \geq 0, c \geq 0, \operatorname{Re}(s) \geq 0 \quad (3.45)$$

$$\gamma_b^a(x) = \frac{a^b x^{b-1}}{\Gamma(b) e^{ax}} = \frac{a}{\Gamma(b)} G_{0,1}^{1,0} \left(ax \left| \begin{matrix} - \\ b-1 \end{matrix} \right. \right) \xrightarrow{\mathcal{M}} \frac{\Gamma(b-1+s)}{a^{s-1} \Gamma(b)}, \quad (3.46)$$

$$x \geq 0, a > 0, b > 0, \operatorname{Re}(b-1+s) > 0$$

$$\gamma_{b,c}^a(x) = \frac{ca^b x^{bc-1}}{\Gamma(b) e^{ax^c}} = \frac{a^{\frac{1}{c}}}{\Gamma(b)} H_{0,1}^{1,0} \left(a^{\frac{1}{c}} x \left| \begin{matrix} - \\ b-\frac{1}{c}, \frac{1}{c} \end{matrix} \right. \right) \xrightarrow{\mathcal{M}} \frac{\Gamma\left(b-\frac{1}{c}+\frac{s}{c}\right)}{a^{\frac{s-1}{c}} \Gamma(b)}, \quad (3.47)$$

$$x \geq 0, a > 0, b > 0, c > 0, \operatorname{Re}\left(b-\frac{1}{c}+\frac{s}{c}\right) > 0$$

$$\delta(x - e^\nu) \xrightarrow{\mathcal{M}} e^{\nu(s-1)}, \nu \in \mathbb{R}, \operatorname{Re}(s-1) \geq 0 \quad (3.48)$$

$$\operatorname{Unif}_{a,b}(x) = \frac{1}{b-a} \xrightarrow{\mathcal{M}} \frac{b^s - a^s}{s(b-a)}, 0 \leq a \leq x \leq b \quad (3.49)$$

$$\operatorname{LogN}_{\nu,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma x}} e^{-\frac{(\nu - \ln x)^2}{2\sigma^2}} \xrightarrow{\mathcal{M}} e^{\nu(s-1) + \frac{\sigma^2}{2}(s-1)^2}, \quad (3.50)$$

$$x \geq 0, \nu \in \mathbb{R}, \sigma > 0, \operatorname{Re}(s-1) \geq 0$$

Comentário 3.7.2. As densidades (3.45) e (3.48) são Deltas de Dirac. A densidade (3.49) é a uniforme. A densidade (3.50) é a Lognormal.

3.7.2. Propriedades da transformada de Mellin

Algumas propriedades da transformada de Mellin serão dadas a seguir.

$$f(x) \xrightarrow{\mathcal{M}} F(s) \quad (3.51)$$

$$(f_1 \otimes f_2)(x) \xrightarrow{\mathcal{M}} F_1(s) F_2(s) \quad (3.52)$$

$$f^{\otimes \frac{1}{T}}(x) = T x^{T-1} f(x^T) \xrightarrow{\mathcal{M}} F\left(1 + \frac{s-1}{T}\right), T > 0 \quad (3.53)$$

$$f(x)(\ln x)^n \xrightarrow{\mathcal{M}} F^{(n)}(s) = \frac{d^n F(s)}{ds^n} \quad (3.54)$$

$$x^k f(x) \xrightarrow{\mathcal{M}} F(s+k) \quad (3.55)$$

As transformadas de Laplace e de Mellin estão relacionadas pela expressão

$$\mathcal{M}\{\mathcal{L}\{f(x); r\}; s\} = \Gamma(s) \mathcal{M}\{f(x); 1-s\}. \quad (3.56)$$

Comentário 3.7.3. A propriedade (3.52) é o Teorema da Convolução no caso multiplicativo e vale para as densidades de variáveis aleatórias independentes. A propriedade (3.53) é a densidade de uma variável aleatória elevada à potência $1/T$ e sua transformada de Mellin. A propriedade (3.56) pode ser justificada pelo teorema de Fubini; veja Oberhettinger (1974).

A seguir, será provada uma propriedade envolvendo a potência e a multiplicação de cópias de uma mesma variável aleatória.

Teorema 3.7.3. Multiplicação da potência e potência da multiplicação de cópias de uma variável aleatória. Seja uma variável aleatória multiplicativa X i.i.d. tal que

$$X \sim f(x) \xrightarrow{\mathcal{M}} F(s),$$

e $T \in \{1, 2, \dots\}$ cópias (independentes) dessa variável. Então

$$(X^{\times T})^{1/T} = (X^{1/T})^{\times T} \sim \left((f^{\otimes T})^{\otimes \frac{1}{T}} \right)(x) = \left((f^{\otimes \frac{1}{T}})^{\otimes T} \right)(x) \xrightarrow{\mathcal{M}} \left(F \left(1 + \frac{s-1}{T} \right) \right)^T. \quad (3.57)$$

Prova. Aplicando a propriedade (3.52) e depois a (3.53), tem-se

$$X^{\times T} = \underbrace{X \cdot \dots \cdot X}_T \sim \underbrace{(f \otimes \dots \otimes f)}_T(x) = (f^{\otimes T})(x) \xrightarrow{\mathcal{M}} \underbrace{F(s) \dots F(s)}_T = (F(s))^T$$

e

$$(X^{\times T})^{1/T} = (\underbrace{X \cdot \dots \cdot X}_T)^{1/T} \sim \left(\underbrace{(f \otimes \dots \otimes f)}_T^{\otimes \frac{1}{T}} \right)(x) = \left((f^{\otimes T})^{\otimes \frac{1}{T}} \right)(x) \xrightarrow{\mathcal{M}} \left(F \left(1 + \frac{s-1}{T} \right) \right)^T.$$

Aplicando a propriedade (3.53) e depois a (3.52), tem-se

$$X^{1/T} \sim f^{\otimes \frac{1}{T}}(x) = T x^{T-1} f(x^T) \xrightarrow{\mathcal{M}} F \left(1 + \frac{s-1}{T} \right)$$

e

$$\begin{aligned} \left(X^{\frac{1}{T}} \right)^{\times T} &= \underbrace{\left(X^{\frac{1}{T}} \cdot \dots \cdot X^{\frac{1}{T}} \right)}_T \sim \underbrace{\left(f^{\otimes \frac{1}{T}} \otimes \dots \otimes f^{\otimes \frac{1}{T}} \right)}_T(x) = \left((f^{\otimes \frac{1}{T}})^{\otimes T} \right)(x) \xrightarrow{\mathcal{M}} \\ &\quad \underbrace{F \left(1 + \frac{s-1}{T} \right) \dots F \left(1 + \frac{s-1}{T} \right)}_T = \left(F \left(1 + \frac{s-1}{T} \right) \right)^T. \end{aligned}$$

□

3.8. As funções Meijer G e Fox H

A função Fox H é a função especial mais genérica conhecida atualmente. Ela tem como casos particulares várias funções famosas tais como Meijer G, Bessel, Hipergeométrica, Legendre e Gauss. A Fox H se reduz à Meijer G com a aplicação da fórmula da multiplicação de Gauss, sob hipóteses fracas; veja Prudnikov *et al.* (1990, p.629). Os programas Mathematica e Maple usam a Meijer G internamente para realizar alguns tipos de integração algébrica definida e indefinida. Eles transformam funções mais elementares na sua “imagem” Meijer G, aplicam as propriedades necessárias e expressam o resultado novamente em funções mais elementares.

Uma propriedade importante para esta dissertação é que a transformada de Mellin da Meijer G é uma razão de produtos de funções Gama. Com o teorema da convolução multiplicativa, é possível determinar facilmente o produto de duas variáveis aleatórias que seguem uma distribuição Meijer G.

As funções Meijer G e Fox H serão definidas a seguir, com a integral de Mellin-Barnes.

Definição 3.8.1. Função Meijer G. A função Meijer G $G: \mathbb{C} \rightarrow \mathbb{C}$ é definida como

$$G_{p,q}^{m,n} \left(z \left| \begin{matrix} a_1, \dots, a_p \\ b_1, \dots, b_q \end{matrix} \right. \right) = G_{p,q}^{m,n} \left(z \left| \begin{matrix} a_1, \dots, a_n; a_{n+1}, \dots, a_p \\ b_1, \dots, b_m; b_{m+1}, \dots, b_q \end{matrix} \right. \right) = G_{p,q}^{m,n} \left(z \left| \begin{matrix} (a_p) \\ (b_q) \end{matrix} \right. \right) = \frac{1}{2\pi i} \int_L h(s) z^{-s} ds, \quad (3.58)$$

onde

$$h(s) = \frac{\left[\prod_{i=1}^m \Gamma(b_i + s) \right] \left[\prod_{j=1}^n \Gamma(1 - a_j - s) \right]}{\left[\prod_{j=n+1}^p \Gamma(a_j + s) \right] \left[\prod_{i=m+1}^q \Gamma(1 - b_i - s) \right]};$$

m, n, p, q são inteiros tais que $0 \leq n \leq p$ e $0 \leq m \leq q$; a_j , para $1 \leq j \leq p$ e b_i , para $1 \leq i \leq q$, são números complexos tais que $-b_i + n_1 \neq 1 - a_j + n_2$ para $n_1 \in \{0, -1, -2, \dots\}$, $n_2 \in \{0, 1, 2, \dots\}$, $1 \leq i \leq m$ e $1 \leq j \leq n$; e $i = \sqrt{-1}$. Um produtório vazio é interpretado como 1. O contorno L no plano complexo é tal que separa os pontos

$$s = -b_i + n_1, \text{ para } 1 \leq i \leq m \text{ e } n_1 = 0, -1, -2, \dots, \quad (3.59)$$

dos pontos

$$s = 1 - a_j + n_2, \text{ para } 1 \leq j \leq n \text{ e } n_2 = 0, 1, 2, \dots. \quad (3.60)$$

A função (3.58) converge nos casos a seguir, nos quais serão usadas as expressões auxiliares c^* , μ e γ , que serão definidas na **Seção 3.8.1** mais adiante.

1. $L = L_{i\infty}$; $c^* > 0$, $|\arg z| < c^* \pi$.
2. $L = L_{i\infty}$; $c^* \geq 0$, $|\arg z| = c^* \pi$, $(q - p)\gamma < -\operatorname{Re}(\mu)$.
3. $L = L_{-\infty}$; $p < q$, $0 < |z| < \infty$; ou $p = q$, $0 < |z| < 1$; ou $p = q$, $c^* \geq 0$, $|z| = 1$, $\operatorname{Re}(\mu) < 0$.
4. $L = L_{+\infty}$; $p > q$, $0 < |z| < \infty$; ou $p = q$, $|z| > 1$; ou $p = q$, $c^* \geq 0$, $|z| = 1$, $\operatorname{Re}(\mu) < 0$.

Os possíveis contornos L são:

- a. $L_{-\infty}$: um laço sobre uma fita horizontal, orientado para a esquerda, começando no ponto $-\infty + i\varphi_1$, contornando o pólo de maior parte real, terminando no ponto $-\infty + i\varphi_2$, onde $\varphi_1 < \varphi_2$, e tal que deixa todos os pólos dados em (3.59) à esquerda e todos os pólos dados em (3.60) à direita do contorno.
- b. $L_{+\infty}$: um laço sobre uma fita horizontal, orientado para a direita, começando no ponto $+\infty + i\varphi_1$, contornando o pólo de menor parte real, terminando no ponto $+\infty + i\varphi_2$, onde $\varphi_1 < \varphi_2$, e tal que deixa todos os pólos dados em (3.59) à esquerda e todos os pólos dados em (3.60) à direita do contorno.
- c. $L_{i\infty}$: um contorno que começa no ponto $\gamma - i\infty$, termina no ponto $\gamma + i\infty$ e deixa todos os pólos dados em (3.59) à esquerda e todos os pólos dados em (3.60) à direita. O valor γ é um número real tal que o contorno separe os pólos.

Comentário 3.8.1. A definição (3.58) começou a ser usada no final da década de 1960. Antes dessa data, e algum tempo após, alguns autores costumavam defini-la com sinal oposto para a variável s . Pequenos detalhes na definição variam de autor para autor, e isto é devido aos diferentes contornos L . Foi adotada a

definição dada em Prudnikov *et al.* (1990, p.617). Para lembrar da notação

$$G_{p,q}^{m,n} \left(z \left| \begin{matrix} a_1, \dots, a_p \\ b_1, \dots, b_q \end{matrix} \right. \right) = G_{p,q}^{m,n} \left(z \left| \begin{matrix} a_1, \dots, a_n; a_{n+1}, \dots, a_p \\ b_1, \dots, b_m; b_{m+1}, \dots, b_q \end{matrix} \right. \right),$$

é conveniente rotacionar mentalmente para a esquerda as letras m, n e p que estão do lado direito da letra G .

Definição 3.8.2. Função Fox H A função Fox $H: \mathbb{C} \rightarrow \mathbb{C}$ é definida como

$$H_{p,q}^{m,n} \left(z \left| \begin{matrix} [a_p, A_p] \\ [b_q, B_q] \end{matrix} \right. \right) = H_{p,q}^{m,n} \left(z \left| \begin{matrix} (a_1, A_1), \dots, (a_p, A_p) \\ (b_1, B_1), \dots, (b_q, B_q) \end{matrix} \right. \right) = \frac{1}{2\pi i} \int_L h(s) z^{-s} ds, \quad (3.61)$$

onde

$$h(s) = \frac{\left[\prod_{i=1}^m \Gamma(b_i + B_i s) \right] \left[\prod_{j=1}^n \Gamma(1 - a_j - A_j s) \right]}{\left[\prod_{j=n+1}^p \Gamma(a_j + A_j s) \right] \left[\prod_{i=m+1}^q \Gamma(1 - b_i - B_i s) \right]};$$

m, n, p, q são inteiros tais que $0 \leq n \leq p$ e $0 \leq m \leq q$; a_j , para $1 \leq j \leq p$, e b_i , para $1 \leq i \leq q$, são números complexos; $A_j > 0$, para $1 \leq j \leq p$; $B_i > 0$, para $1 \leq i \leq q$; e $i = \sqrt{-1}$. Um produtório vazio é interpretado como 1. O contorno L no plano complexo é tal que separa os pólos de

$$\Gamma(b_i + B_i s), \text{ para } 1 \leq i \leq m, \quad (3.62)$$

dos pólos de

$$\Gamma(1 - a_j - A_j s), \text{ para } 1 \leq j \leq n. \quad (3.63)$$

A função (3.61) converge nos casos a seguir, nos quais serão usadas as expressões auxiliares a^* , c^* , β , μ , γ e Δ , que serão definidas na **Seção 3.8.1** mais adiante.

1. $L = L_{i\infty}$; $a^* > 0$, $|\arg z| < a^* \pi / 2$.
2. $L = L_{i\infty}$; $a^* \geq 0$, $|\arg z| = a^* \pi / 2$, $\gamma \Delta < -\text{Re}(\mu)$.
3. $L = L_{-\infty}$; $\Delta > 0$, $0 < |z| < \infty$; ou $\Delta = 0$, $0 < |z| < \beta$; ou $\Delta = 0$, $a^* \geq 0$, $|z| = \beta$, $\text{Re}(\mu) < 0$.
4. $L = L_{+\infty}$; $\Delta < 0$, $0 < |z| < \infty$; ou $\Delta = 0$, $|z| > \beta$; ou $\Delta = 0$, $a^* \geq 0$, $|z| = \beta$, $\text{Re}(\mu) < 0$.

Os possíveis contornos L são:

- a. $L_{-\infty}$: um laço sobre uma fita horizontal, orientado para a esquerda, começando no ponto $-\infty + i\varphi_1$, contornando o pólo de maior parte real, terminando no ponto $-\infty + i\varphi_2$, onde $\varphi_1 < \varphi_2$, e tal que deixa todos os pólos dados em (3.62) à esquerda e todos os pólos dados em (3.63) à direita do contorno.
- b. $L_{+\infty}$: um laço sobre uma fita horizontal, orientado para a direita, começando no ponto $+\infty + i\varphi_1$, contornando o pólo de menor parte real, terminando no ponto $+\infty + i\varphi_2$, onde $\varphi_1 < \varphi_2$, e tal que deixa todos os pólos dados em (3.62) à esquerda e todos os pólos dados em (3.63) à direita do contorno.
- c. $L_{i\infty}$: um contorno que começa no ponto $\gamma - i\infty$, termina no ponto $\gamma + i\infty$ e deixa todos os

pólos dados em (3.62) à esquerda e todos os pólos dados em (3.63) à direita. O valor γ é um número real tal que o contorno separe os pólos.

Comentário 3.8.2. Foi adotada a definição dada em Prudnikov *et al.* (1990, p.626).

3.8.1. Notações auxiliares para as funções Meijer G e Fox H

A notação e os parâmetros auxiliares usados pelas funções Meijer G e Fox H serão dados a seguir.

$$[a_p, A_p] = (a_1, A_1), \dots, (a_p, A_p) \quad (3.64)$$

$$k \pm (a_p) = k \pm a_1, \dots, k \pm a_p \quad (3.65)$$

$$a^* = \sum_{j=1}^n A_j - \sum_{j=n+1}^p A_j + \sum_{i=1}^m B_i - \sum_{i=m+1}^q B_i \quad (3.66)$$

$$c^* = m + n - \frac{p+q}{2} \quad (3.67)$$

$$\mu = \sum_{i=1}^q b_i - \sum_{j=1}^p a_j + \frac{p-q}{2} + 1 \quad (3.68)$$

$$\beta = \prod_{j=1}^p A_j^{-A_j} \prod_{i=1}^q B_i^{B_i} \quad (3.69)$$

$$\Delta = \sum_{i=1}^q B_i - \sum_{j=1}^p A_j \quad (3.70)$$

$$\gamma = \lim_{\substack{s \rightarrow \infty \\ s \in L_{4\infty}}} \operatorname{Re}(s) \quad (3.71)$$

$$\Delta(k, a) = \frac{a}{k}, \frac{a+1}{k}, \dots, \frac{a+k-1}{k} \quad (3.72)$$

$$\Delta(k, (a_p)) = \frac{(a_p)}{k}, \frac{(a_p)+1}{k}, \dots, \frac{(a_p)+k-1}{k} \quad (3.73)$$

$$\left[\begin{array}{l} \text{Condições suficientes para a} \\ \text{validade de uma expressão} \end{array} \right] \quad (3.74)$$

3.8.2. Casos particulares

Alguns casos particulares da função Meijer G serão dados a seguir.

$$\gamma_b^a(x) = \frac{a^b x^{b-1}}{\Gamma(b) e^{ax}} = \frac{1}{\Gamma(b)x} G_{0,1}^{1,0} \left(ax \left| \begin{array}{c} - \\ b \end{array} \right. \right) = \frac{a}{\Gamma(b)} G_{0,1}^{1,0} \left(ax \left| \begin{array}{c} - \\ b-1 \end{array} \right. \right) \quad (3.75)$$

$$\gamma_{b,c}^a(x) = \frac{ca^b x^{bc-1}}{\Gamma(b)e^{ax^c}} = \frac{c}{\Gamma(b)x} G_{0,1}^{1,0} \left(ax^c \middle| \begin{matrix} - \\ b \end{matrix} \right) = \frac{ca^{1/c}}{\Gamma(b)} G_{0,1}^{1,0} \left(ax^c \middle| \begin{matrix} - \\ b - \frac{1}{c} \end{matrix} \right) \quad (3.76)$$

$$\Gamma(b, x) = G_{1,2}^{2,0} \left(x \middle| \begin{matrix} 1 \\ 0, b \end{matrix} \right) \quad (3.77)$$

$$\gamma(b, x) = G_{1,2}^{1,1} \left(x \middle| \begin{matrix} 1 \\ b; 0 \end{matrix} \right) \quad (3.78)$$

$$\frac{\prod_{j=1}^p \Gamma(a_j)}{\prod_{j=1}^q \Gamma(b_j)} {}_pF_q \left(\begin{matrix} a_1, \dots, a_p \\ b_1, \dots, b_q \end{matrix} \middle| -x \right) = G_{p,q+1}^{1,p} \left(x \middle| \begin{matrix} 1 - (a_p) \\ 0, 1 - (b_q) \end{matrix} \right), \quad (3.79)$$

$$\left\| \begin{matrix} p \leq q \text{ ou} \\ p = q + 1, |x| \leq 1 \end{matrix} \right\|$$

Comentário 3.8.3. A função ${}_pF_q$ em (3.79) é a hipergeométrica.

3.8.3. Propriedades, integrais e transformadas

Algumas propriedades, integrais e transformadas das funções Meijer G e Fox H serão dadas a seguir.

$$G_{p,q}^{m,n} \left(x \middle| \begin{matrix} (a_p) \\ (b_q) \end{matrix} \right) = G_{q,p}^{n,m} \left(\frac{1}{x} \middle| \begin{matrix} 1 - (b_q) \\ 1 - (a_p) \end{matrix} \right) \quad (3.80)$$

$$x^c G_{p,q}^{m,n} \left(x \middle| \begin{matrix} (a_p) \\ (b_q) \end{matrix} \right) = G_{p,q}^{m,n} \left(x \middle| \begin{matrix} c + (a_p) \\ c + (b_q) \end{matrix} \right) \quad (3.81)$$

$$H_{p,q}^{m,n} \left(z \middle| \begin{matrix} (a_p, k) \\ (b_q, k) \end{matrix} \right) = \frac{1}{k} G_{p,q}^{m,n} \left(z^{1/k} \middle| \begin{matrix} (a_p) \\ (b_q) \end{matrix} \right), \quad k > 0 \quad (3.82)$$

$$H_{p,q}^{m,n} \left(\frac{1}{x} \middle| \begin{matrix} (a_p, A_p) \\ (b_q, B_q) \end{matrix} \right) = H_{q,p}^{n,m} \left(x \middle| \begin{matrix} (1 - b_q, B_q) \\ (1 - a_p, A_p) \end{matrix} \right) \quad (3.83)$$

$$x^c H_{p,q}^{m,n} \left(x \middle| \begin{matrix} (a_p, A_p) \\ (b_q, B_q) \end{matrix} \right) = H_{p,q}^{m,n} \left(x \middle| \begin{matrix} (a_p + cA_p, A_p) \\ (b_q + cB_q, B_q) \end{matrix} \right) \quad (3.84)$$

$$H_{p,q}^{m,n} \left(z \middle| \begin{matrix} (a_p, A_p) \\ (b_q, B_q) \end{matrix} \right) = k H_{p,q}^{m,n} \left(z^k \middle| \begin{matrix} (a_p, kA_p) \\ (b_q, kB_q) \end{matrix} \right), \quad k > 0 \quad (3.85)$$

$$G_{p,q}^{m,n} \left(\lambda x \left| \begin{matrix} (a_p) \\ (b_q) \end{matrix} \right. \right) \xrightarrow{\mathcal{M}} \lambda^{-s} \frac{\prod_{i=1}^m \Gamma(b_i + s) \prod_{j=1}^n \Gamma(1 - a_j - s)}{\prod_{i=m+1}^q \Gamma(1 - b_i - s) \prod_{j=n+1}^p \Gamma(a_j + s)}, \quad (3.86)$$

$$\left[\begin{array}{l} 0 \leq n \leq p; 0 \leq m \leq q; \lambda \neq 0; c^* > 0; |\arg \lambda| < c^* \pi; \\ -\min \operatorname{Re}(b_i) < \operatorname{Re}(s) < 1 - \max \operatorname{Re}(a_j), 1 \leq i \leq m, 1 \leq j \leq n \end{array} \right]$$

$$H_{p,q}^{m,n} \left(\lambda x \left| \begin{matrix} [a_p, A_p] \\ [b_q, B_q] \end{matrix} \right. \right) \xrightarrow{\mathcal{M}} \lambda^{-s} \frac{\prod_{i=1}^m \Gamma(b_i + B_i s) \prod_{j=1}^n \Gamma(1 - a_j - A_j s)}{\prod_{i=m+1}^q \Gamma(1 - b_i - B_i s) \prod_{j=n+1}^p \Gamma(a_j + A_j s)}, \quad (3.87)$$

$$\left[\begin{array}{l} mn \neq 0; \lambda \neq 0; a^* > 0; |\arg \lambda| < a^* \pi / 2; \\ -\min \operatorname{Re} \left(\frac{b_i}{B_i} \right) < \operatorname{Re}(s) < 1 - \max \operatorname{Re} \left(\frac{1 - a_j}{A_j} \right), 1 \leq i \leq m, 1 \leq j \leq n \end{array} \right]$$

$$\int_0^y x^{\alpha-1} (y-x)^{\beta-1} G_{p,q}^{m,n} \left(\omega x^{l/k} \left| \begin{matrix} (a_p) \\ (b_q) \end{matrix} \right. \right) dx = \frac{k^\mu y^{\alpha+\beta-1} \Gamma(\beta)}{(2\pi)^{c^*(k-1)} l^\beta} G_{kp+l, kq+l}^{km, kn+l} \left(\frac{\omega^k y^l}{k^{k(q-p)}} \left| \begin{matrix} \Delta(l, 1-\alpha), \Delta(k, (a_p)) \\ \Delta(k, (b_q)), \Delta(l, 1-\alpha-\beta) \end{matrix} \right. \right), \quad (3.88)$$

$$\left[\begin{array}{l} y > 0; \beta > 0; \operatorname{Re}(\alpha + b_j) > 0, 1 \leq j \leq m; \\ a_i - b_j \neq 1, 2, \dots, 1 \leq i \leq n, 1 \leq j \leq m; c^* > 0; |\arg \omega| < c^* \pi \end{array} \right]$$

$$\int_0^y x^{\alpha-1} (y-x)^{\beta-1} H_{p,q}^{m,n} \left(\omega x^r \left| \begin{matrix} [a_p, A_p] \\ [b_q, B_q] \end{matrix} \right. \right) dx = y^{\alpha+\beta-1} \Gamma(\beta) H_{p+1, q+1}^{m, n+1} \left(\omega y^r \left| \begin{matrix} (1-\alpha, r), [a_p, A_p] \\ [b_q, B_q], (1-\alpha-\beta, r) \end{matrix} \right. \right), \quad (3.89)$$

$$\left[\begin{array}{l} mn \neq 0; r \geq 0; y > 0; a^* > 0; \beta > 0; |\arg \omega| < a^* \pi / 2; \\ \operatorname{Re}(\alpha) > 0; \operatorname{Re}(\alpha) + r \min \operatorname{Re} \left(\frac{b_j}{B_j} \right) > 0, 1 \leq j \leq m \end{array} \right]$$

Comentário 3.8.4. A propriedade (3.86) está em Prudnikov *et al.* (1990, p.347, 728). A propriedade (3.87) está em Prudnikov *et al.* (1990, p.355, 729). A propriedade (3.88) está em Prudnikov *et al.* (1990, p.348). A propriedade (3.89) está em Prudnikov *et al.* (1990, p.355).

Exemplo 3.8.1. Relação entre as transformadas de Laplace e Mellin. Aplicando a propriedade (3.56) para a Gama Generalizada com parâmetro $c > 0$, obtém-se

$$f(x) = \gamma_{b,c}^a(x) = \frac{ca^b x^{bc-1}}{\Gamma(b)e^{ax^c}} = \frac{a^{\frac{1}{c}}}{\Gamma(b)} H_{0,1}^{1,0} \left(a^{\frac{1}{c}} x \left| \begin{matrix} ; \\ b - \frac{1}{c}, \frac{1}{c} \end{matrix} \right. \right),$$

$$\mathcal{M}\{\mathcal{L}\{f(x);r\};s\} = \mathcal{M}\left\{\frac{1}{\Gamma(b)}H_{1,1}^{1,1}\left(\frac{r}{a^{\frac{1}{c}}}\left|1-b,\frac{1}{c}\right.\right);s\right\} = \frac{a^{\frac{s}{c}}}{\Gamma(b)}\Gamma(s)\Gamma\left(b-\frac{s}{c}\right)$$

e

$$\Gamma(s)\mathcal{M}\{f(x);1-s\} = \Gamma(s)\left[\frac{\Gamma\left(b-\frac{1}{c}+\frac{r}{c}\right)}{a^{\frac{r-1}{c}}\Gamma(b)}\right]_{r=1-s} = \Gamma(s)\frac{\Gamma\left(b-\frac{s}{c}\right)}{a^{\frac{-s}{c}}\Gamma(b)} = \frac{a^{\frac{s}{c}}}{\Gamma(b)}\Gamma(s)\Gamma\left(b-\frac{s}{c}\right).$$

■

3.8.4. Distribuições Meijer G e Fox H

As distribuições Meijer G e Fox H serão definidas a seguir; veja Carter e Springer (1977).

Definição 3.8.3. Distribuição Meijer G. A distribuição Meijer G tem sua densidade $g : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ definida por

$$g(x) = kG_{p,q}^{m,n}\left(cx \left| \begin{matrix} a_1, \dots, a_p \\ b_1, \dots, b_q \end{matrix} \right. \right), \quad (3.90)$$

onde m, n, p, q, a_j, b_i , para $1 \leq j \leq p$ e $1 \leq i \leq q$, são parâmetros que satisfazem a definição da função Meijer G e que, juntamente com as constantes k e c , fazem $g(x) \geq 0$ para $x \geq 0$, e

$$\int_0^\infty g(x)dx = 1.$$

Definição 3.8.4. Distribuição Fox H. A distribuição Fox H tem sua densidade $h : \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$ definida por

$$h(x) = kH_{p,q}^{m,n}\left(cx \left| \begin{matrix} (a_1, A_1), \dots, (a_p, A_p) \\ (b_1, B_1), \dots, (b_q, B_q) \end{matrix} \right. \right), \quad (3.91)$$

onde $m, n, p, q, a_j, b_i, A_j, B_i$, para $1 \leq j \leq p$ e $1 \leq i \leq q$, são parâmetros que satisfazem a definição da função Fox H e que, juntamente com as constantes k e c , fazem $h(x) \geq 0$ para $x \geq 0$, e

$$\int_0^\infty h(x)dx = 1.$$

3.9. Avaliação numérica da função Meijer G

A avaliação numérica da função Meijer G é feita com a aplicação do teorema dos resíduos. Por esse teorema, a integral de linha pode ser calculada pela soma dos resíduos nos pólos da função. Como pode ser visto na **Figura 3.2.1**, a função $\Gamma(x)$ tem pólos nos pontos $x \in \{0, -1, -2, \dots\}$.

Se todos os pólos são simples, então a avaliação numérica da função Meijer G é relativamente simples e pode ser feita com um somatório de produtos de funções Gama e Hipergeométrica; veja Prudnikov,

Brychkov e Marichev (1990, p.618) ou Mathai e Saxena (1973, p.3). Este caso não será usado neste trabalho.

O caso em que há pólos com multiplicidade maior que 1 é conhecido como caso logarítmico. A avaliação numérica da função Meijer G no caso logarítmico foi obtida por Mathai (1970). Como a expressão é longa, não será apresentada aqui. Mathai também obteve expressões para casos particulares da função Meijer G no caso logarítmico; veja Mathai (1971) e Anderson e Mathai (1984). Aqui serão apresentados resultados que são simplificações dos de Mathai e, em particular, um caso especial da função Meijer G que será usado neste trabalho. É necessário começar com alguns lemas envolvendo limites.

Lema 3.9.1. Se $v - b_i \in \{0, 1, 2, \dots\}$ para todos os i , $1 \leq i \leq k$, então

$$\lim_{s \rightarrow -v} (s+v)^k \Gamma(s+b_1) \Gamma(s+b_2) \dots \Gamma(s+b_k) = \frac{(-1)^{kv-b_1-b_2-\dots-b_k}}{(v-b_1)!(v-b_2)! \dots (v-b_k)!}.$$

Prova. Será apresentada a prova para $k = 2$. Para outros valores, a prova é similar. Considere que $b_1 \leq b_2$. Então

$$\begin{aligned} & \lim_{s \rightarrow -v} (s+v)^2 \Gamma(s+b_1) \Gamma(s+b_2) = \\ &= \lim_{s \rightarrow -v} \frac{(s+v)^2 \Gamma^2(s+v+1) \Gamma(s+b_1) \Gamma(s+b_2)}{(s+v)^2 (s+v-1)^2 \dots (s+b_1)^2 \Gamma(s+b_1) (s+b_1+1) \dots (s+b_2) \Gamma(s+b_2)} \\ &= \frac{\Gamma^2(1)}{(-1)^2 \dots (-v+b_1)^2 (-v+b_1+1) \dots (-v+b_2)} = \frac{(-1)^{v-b_1} (-1)^{v-b_2}}{(v-b_1)!(v-b_2)!} = \frac{(-1)^{2v-b_1-b_2}}{(v-b_1)!(v-b_2)!}. \end{aligned}$$

□

Lema 3.9.2. Sejam

$$\Theta(s) = \frac{\Gamma(b_{k+1}+s) \dots \Gamma(b_m+s) \Gamma(1-a_1-s) \dots \Gamma(1-a_n-s)}{\Gamma(a_{n+1}+s) \dots \Gamma(a_p+s) \Gamma(1-b_{m+1}-s) \dots \Gamma(1-b_q-s)}; \quad b_{\min} = \min_{1 \leq i \leq m} b_i; \text{ e } a_{\max} = \max_{1 \leq j \leq n} a_j.$$

Se $-v + b_i \in \{0, -1, -2, \dots\}$ para todos os i , $1 \leq i \leq k$; $-v + b_i \notin \{0, -1, -2, \dots\}$ para todos os i , $k+1 \leq i \leq m$; $1 - b_i + v \notin \{0, -1, -2, \dots\}$ para todos os i , $m+1 \leq i \leq q$; $-b_{\min} < 1 - a_{\max}$; $a_j - v \notin \{0, -1, -2, \dots\}$, para todos os j , $n+1 \leq j \leq p$; então

$$\lim_{s \rightarrow -v} (s+v)^k \Gamma(s+b_1) \dots \Gamma(s+b_k) \Theta(s) = \frac{(-1)^{kv-b_1-\dots-b_k}}{(v-b_1)! \dots (v-b_k)!} \Theta(-v).$$

Prova. As hipóteses condicionam os argumentos de todas as funções Gama de $\Theta(s)$ de tal forma que eles não mudam a ordem (multiplicidade) do pólo do produto $\Gamma(s+b_1) \dots \Gamma(s+b_k)$ no ponto $s = -v$. Aplicando o lema anterior, o resultado é obtido.

□

Lema 3.9.3. Se $1 \leq t \leq k$ e $v - b_i \in \{0, 1, 2, \dots\}$ para todos os i , $1 \leq i \leq k$, então

$$\lim_{s \rightarrow -v} \frac{d^t}{ds^t} \ln \left[(s+v)^k \Gamma(s+b_1) \Gamma(s+b_2) \dots \Gamma(s+b_k) \right] = \sum_{i=1}^k \psi_G(t-1, v-b_i),$$

onde

$$\psi_G(d, n) = \psi^{(d)}(1) + d! \sum_{i=1}^n i^{-d-1} = (-1)^{d+1} d! \zeta(d+1, 1) + d! \sum_{i=1}^n i^{-d-1}.$$

Prova. Será apresentada a prova para $t = k = 2$. Para outros valores, a prova é similar. Considere que $b_1 \leq b_2$. Então

$$\begin{aligned} & \lim_{s \rightarrow -v} \frac{d^2}{ds^2} \ln[(s+v)^2 \Gamma(s+b_1) \Gamma(s+b_2)] = \\ &= \lim_{s \rightarrow -v} \frac{d^2}{ds^2} \ln \left[(s+v)^2 \frac{\Gamma^2(s+v+1) \Gamma(s+b_1) \Gamma(s+b_2)}{(s+v)^2 (s+v-1)^2 \dots (s+b_1)^2 \Gamma(s+b_1) (s+b_1-1) \dots (s+b_2) \Gamma(s+b_2)} \right] \\ &= \lim_{s \rightarrow -v} \frac{d^2}{ds^2} \left[\ln \Gamma^2(s+v+1) + \ln(s+v-1)^{-2} + \dots + \ln(s+b_1)^{-2} + \ln(s+b_1-1)^{-1} + \dots + \ln(s+b_2)^{-1} \right] \\ &= \lim_{s \rightarrow -v} \frac{d^2}{ds^2} [2 \ln \Gamma(s+v+1) - 2 \ln(s+v-1) - \dots - 2 \ln(s+b_1) - \ln(s+b_1-1) - \dots - \ln(s+b_2)] \\ &= \lim_{s \rightarrow -v} \frac{d}{ds} \left[2\psi(s+v+1) - \frac{2}{(s+v-1)} - \dots - \frac{2}{(s+b_1)} - \frac{1}{(s+b_1-1)} - \dots - \frac{1}{(s+b_2)} \right] \\ &= \lim_{s \rightarrow -v} \left[2\psi^{(1)}(s+v+1) + \frac{2.1!}{(s+v-1)^2} + \dots + \frac{2.1!}{(s+b_1)^2} + \frac{1.1!}{(s+b_1-1)^2} + \dots + \frac{1.1!}{(s+b_2)^2} \right] \\ &= 2\psi^{(1)}(1) + \frac{2.1!}{(-1)^2} + \dots + \frac{2.1!}{(-v+b_1)^2} + \frac{1.1!}{(-v+b_1-1)^2} + \dots + \frac{1.1!}{(-v+b_2)^2} \\ &= \psi^{(1)}(1) + \frac{1!}{(1)^2} + \dots + \frac{1!}{(v-b_1)^2} + \psi^{(1)}(1) + \frac{1!}{(1)^2} + \dots + \frac{1!}{(v-b_1)^2} + \frac{1!}{(v-b_1+1)^2} + \dots + \frac{1!}{(v-b_2)^2} \\ &= \psi_G(1, v-b_1) + \psi_G(1, v-b_2). \end{aligned}$$

□

Lema 3.9.4. Sejam

$$\Theta(s) = \frac{\Gamma(b_{k+1}+s) \dots \Gamma(b_m+s) \Gamma(1-a_1-s) \dots \Gamma(1-a_n-s)}{\Gamma(a_{n+1}+s) \dots \Gamma(a_p+s) \Gamma(1-b_{m+1}-s) \dots \Gamma(1-b_q-s)}; \quad b_{\min} = \min_{1 \leq i \leq m} b_i; \quad e \quad a_{\max} = \max_{1 \leq j \leq n} a_j.$$

Se $1 \leq t \leq k$; $-v+b_i \in \{0, -1, -2, \dots\}$ para todos os i , $1 \leq i \leq k$; $-v+b_i \notin \{0, -1, -2, \dots\}$ para todos os i , $k+1 \leq i \leq m$; $1-b_i+v \notin \{0, -1, -2, \dots\}$ para todos os i , $m+1 \leq i \leq q$; $-b_{\min} < 1-a_{\max}$; e $a_j-v \notin \{0, -1, -2, \dots\}$, para todos os j , $n+1 \leq j \leq p$; então

$$\begin{aligned} \lim_{s \rightarrow -v} \frac{d^t}{ds^t} \ln[(s+v)^k \Gamma(s+b_1) \dots \Gamma(s+b_k) \Theta(s)] &= \sum_{i=1}^k \psi_G(t-1, v-b_i) + \sum_{i=k+1}^m \psi^{(t-1)}(b_i-v) + \\ &+ (-1)^{t-1} \sum_{i=m+1}^q \psi^{(t-1)}(1-b_i+v) + (-1)^t \sum_{j=1}^n \psi^{(t-1)}(1-a_j+v) - \sum_{j=n+1}^p \psi^{(t-1)}(a_j-v), \end{aligned}$$

onde $\psi_G(d, n)$ foi dado no lema anterior.

Prova. As hipóteses condicionam os argumentos de todas as funções Gama de $\Theta(s)$ de tal forma que eles não mudam a ordem (multiplicidade) do pólo do produto $\Gamma(s+b_1)\dots\Gamma(s+b_k)$ no ponto $s=-v$. Notando que

$$\begin{aligned}\lim_{s \rightarrow -v} \frac{d^t}{ds^t} \ln \Theta(s) &= \lim_{s \rightarrow -v} \frac{d^t}{ds^t} \left[\frac{\ln \Gamma(b_{k+1}+s) + \dots + \ln \Gamma(b_m+s) - \ln \Gamma(1-b_{m+1}-s) - \dots - \ln \Gamma(1-b_q-s) +}{\ln \Gamma(1-a_1-s) + \dots + \ln \Gamma(1-a_n-s) - \ln \Gamma(a_{n+1}+s) - \dots - \ln \Gamma(a_p+s)} \right] \\ &= \lim_{s \rightarrow -v} \frac{d^{t-1}}{ds^{t-1}} \left[\frac{\psi(b_{k+1}+s) + \dots + \psi(b_m+s) + \psi(1-b_{m+1}-s) + \dots + \psi(1-b_q-s) +}{-\psi(1-a_1-s) - \dots - \psi(1-a_n-s) - \psi(a_{n+1}+s) - \dots - \psi(a_p+s)} \right] \\ &= \psi^{(t-1)}(b_{k+1}-v) + \dots + \psi^{(t-1)}(b_m-v) + \\ &\quad (-1)^{t-1} [\psi^{(t-1)}(1-b_{m+1}+v) + \dots + \psi^{(t-1)}(1-b_q+v)] + \\ &\quad (-1)^t [\psi^{(t-1)}(1-a_1+v) + \dots + \psi^{(t-1)}(1-a_n+v)] + \\ &\quad -\psi^{(t-1)}(a_{n+1}-v) - \dots - \psi^{(t-1)}(a_p-v) \quad ,\end{aligned}$$

e inserindo este resultado na prova do lema anterior, o resultado é obtido. □

Lema 3.9.5. Se $\Delta(s)$ é uma razão de produtos de funções Gama com um pólo de ordem $k \geq 1$ em $s=-v$, então o resíduo de $\Delta(s)x^{-s}$ neste ponto é

$$Res(\Delta(s)x^{-s}, -v) = x^v \left\{ \lim_{s \rightarrow -v} [(s+v)^k \Delta(s)] \right\} \sum_{t=0}^{k-1} \frac{[-\ln x]^{k-t-1}}{(k-t-1)!} M(t) \quad ,$$

onde

$$M(t) = \begin{cases} 1 & , \quad t=0 \\ \frac{1}{t} \sum_{u=0}^{t-1} \frac{\lim_{s \rightarrow -v} [\ln[(s+v)^k \Delta(s)]]^{(t-u)}}{(t-u-1)!} M(u) & , \quad t>0 \quad . \end{cases}$$

Prova. O resíduo é dado por

$$\begin{aligned}Res(\Delta(s)x^{-s}, -v) &= \lim_{s \rightarrow -v} \frac{1}{(k-1)!} \frac{d^{k-1}}{ds^{k-1}} [(s+v)^k \Delta(s) x^{-s}] \\ &= \frac{1}{(k-1)!} \lim_{s \rightarrow -v} \sum_{t=0}^{k-1} \binom{k-1}{t} [(s+v)^k \Delta(s)]^{(t)} [e^{-s \ln x}]^{(k-1-t)} \\ &= \frac{x^v}{(k-1)!} \lim_{s \rightarrow -v} \sum_{t=0}^{k-1} \binom{k-1}{t} [-\ln x]^{k-1-t} [(s+v)^k \Delta(s)]^{(t)} \\ &= \frac{x^v}{(k-1)!} \sum_{t=0}^{k-1} \binom{k-1}{t} [-\ln x]^{k-1-t} \lim_{s \rightarrow -v} [(s+v)^k \Delta(s)]^{(t)} \\ &= x^v \sum_{t=0}^{k-1} \frac{[-\ln x]^{k-1-t}}{(k-1-t)!} \frac{1}{t!} \lim_{s \rightarrow -v} [(s+v)^k \Delta(s)]^{(t)} \quad .\end{aligned}$$

Definindo $L(t)$ como

$$\begin{aligned}
L(t) &= \frac{1}{t!} \lim_{s \rightarrow -v} [(s+v)^k \Delta(s)]^{(t)} \\
&= \begin{cases} \frac{1}{0!} \lim_{s \rightarrow -v} [(s+v)^k \Delta(s)] & , \quad t=0 \\ \frac{1}{t!} \lim_{s \rightarrow -v} \left\{ \left[\frac{d}{ds} \ln[(s+v)^k \Delta(s)] \right] [(s+v)^k \Delta(s)] \right\}^{(t-1)} & , \quad t>0 \end{cases} \\
&= \begin{cases} \lim_{s \rightarrow -v} [(s+v)^k \Delta(s)] & , \quad t=0 \\ \frac{1}{t!} \lim_{s \rightarrow -v} \left\{ \sum_{u=0}^{t-1} \binom{t-1}{u} \left[\frac{d}{ds} \ln[(s+v)^k \Delta(s)] \right]^{(t-1-u)} [(s+v)^k \Delta(s)]^{(u)} \right\} & , \quad t>0 \end{cases} \\
&= \begin{cases} \lim_{s \rightarrow -v} [(s+v)^k \Delta(s)] & , \quad t=0 \\ \frac{1}{t!} \left\{ \sum_{u=0}^{t-1} \binom{t-1}{u} \left[\lim_{s \rightarrow -v} \left[\ln[(s+v)^k \Delta(s)] \right]^{(t-u)} \right] \left[\lim_{s \rightarrow -v} [(s+v)^k \Delta(s)]^{(u)} \right] \right\} & , \quad t>0 \end{cases} \\
&= \begin{cases} \lim_{s \rightarrow -v} [(s+v)^k \Delta(s)] & , \quad t=0 \\ \frac{1}{t!} \left\{ \sum_{u=0}^{t-1} \frac{(t-1)!}{(t-1-u)!} \left[\lim_{s \rightarrow -v} \left[\ln[(s+v)^k \Delta(s)] \right]^{(t-u)} \right] \frac{1}{u!} \left[\lim_{s \rightarrow -v} [(s+v)^k \Delta(s)]^{(u)} \right] \right\} & , \quad t>0 \end{cases} \\
&= \begin{cases} \lim_{s \rightarrow -v} [(s+v)^k \Delta(s)] & , \quad t=0 \\ \frac{1}{t!} \left\{ \sum_{u=0}^{t-1} \frac{1}{(t-u-1)!} \left[\lim_{s \rightarrow -v} \left[\ln[(s+v)^k \Delta(s)] \right]^{(t-u)} \right] L(u) \right\} & , \quad t>0 \end{cases} ,
\end{aligned}$$

então

$$\begin{aligned}
Res(\Delta(s)x^{-s}, -v) &= x^v \sum_{t=0}^{k-1} \frac{[-\ln x]^{k-1-t}}{(k-1-t)!} L(t) \\
&= x^v \sum_{t=0}^{k-1} \frac{[-\ln x]^{k-t-1}}{(k-t-1)!} \lim_{s \rightarrow -v} [(s+v)^k \Delta(s)] M(t) \\
&= x^v \left\{ \lim_{s \rightarrow -v} [(s+v)^k \Delta(s)] \right\} \sum_{t=0}^{k-1} \frac{[-\ln x]^{k-t-1}}{(k-t-1)!} M(t) ,
\end{aligned}$$

onde

$$M(t) = \begin{cases} 1 & , \quad t=0 \\ \frac{1}{t!} \sum_{u=0}^{t-1} \frac{\lim_{s \rightarrow -v} \left[\ln[(s+v)^k \Delta(s)] \right]^{(t-u)}}{(t-u-1)!} M(u) & , \quad t>0 \end{cases} .$$

□

No lema anterior, a razão $\Delta(s)$ de produtos de funções Gama pode, em geral, ser simplificada. Em alguns pólos, algumas das funções Gama (com argumento zero ou inteiro negativo no pólo) no numerador podem simplificar algumas funções Gama no denominador (com argumento zero ou inteiro negativo no pólo), introduzindo monômios. Em um caso especial no qual esta simplificação não ocorre em nenhum dos infinitos pólos, o **Lema 3.9.2** e o **Lema 3.9.4** podem ser empregados e a avaliação numérica da função Meijer G pode ser feita mais facilmente. Este caso especial é dado a seguir.

Lema 3.9.6. Caso logarítmico da função Meijer G com pólos no numerador e no denominador separados. Se

$$\Delta(s) = \frac{\Gamma(b_1 + s) \dots \Gamma(b_m + s) \Gamma(1 - a_1 - s) \dots \Gamma(1 - a_n - s)}{\Gamma(a_{n+1} + s) \dots \Gamma(a_p + s) \Gamma(1 - b_{m+1} - s) \dots \Gamma(1 - b_q - s)}; \quad b_{\min} = \min_{1 \leq i \leq m} b_i, \quad a_{\max} = \max_{1 \leq j \leq n} a_j;$$

$1 - b_j + b_i \notin \{0, -1, -2, \dots\}$ para todos os j e i , $m + 1 \leq j \leq q$, $1 \leq i \leq m$; $-b_{\min} < 1 - a_{\max}$; e $a_j - b_i \notin \{0, \pm 1, \pm 2, \dots\}$ para todos os j e i , $n + 1 \leq j \leq p$, $1 \leq i \leq m$; então

$$\Theta(s) = \frac{\Gamma(b_{k+1} + s) \dots \Gamma(b_m + s) \Gamma(1 - a_1 - s) \dots \Gamma(1 - a_n - s)}{\Gamma(a_{n+1} + s) \dots \Gamma(a_p + s) \Gamma(1 - b_{m+1} - s) \dots \Gamma(1 - b_q - s)}$$

não muda a ordem do pólos do produto $\Gamma(s + b_1) \dots \Gamma(s + b_k)$.

Prova. Um pólo de ordem $k \geq 1$ ocorre em $s = -b_i + n_i$, $n_i \in \{0, -1, -2, \dots\}$, $1 \leq i \leq m$. A ordem do pólo é determinada pelo número de termos do produto $\Gamma(b_1 + s) \dots \Gamma(b_k + s)$; as funções Gama do produto $\Gamma(b_{k+1} + s) \dots \Gamma(b_m + s)$ não contribuem para ele. Como $-b_{\min} < 1 - a_{\max}$, as funções $\Gamma(1 - a_j - s)$, $1 \leq j \leq n$, não afetarão a ordem do pólo. Se $a_j - b_i \notin \{0, \pm 1, \pm 2, \dots\}$ para todos os j e i , $n + 1 \leq j \leq p$, $1 \leq i \leq m$, então $a_j - b_i + n_i \notin \{0, -1, -2, \dots\}$ e as funções $\Gamma(a_j + s)$, para $n + 1 \leq j \leq p$, não afetarão a ordem do pólo. Se $1 - b_j + b_i \notin \{0, -1, -2, \dots\}$ para todos os j e i , $m + 1 \leq j \leq q$, $1 \leq i \leq m$, então $1 - b_j + b_i - n_i \notin \{0, -1, -2, \dots\}$ e as funções Gama $\Gamma(1 - b_j - s)$, para $m + 1 \leq j \leq q$, não afetarão a ordem do pólo.

□

Exemplo 3.9.1. Avaliação numérica de um caso especial da função Meijer G. As fórmulas desta seção foram implementadas no Mathematica. O código fonte está no **Apêndice A**. Uma função Meijer G com parâmetros arbitrários satisfazendo as condições do lema anterior foi calculada.

A função Meijer G escolhida e seu valor são

$$G_{3,6}^{4,2} \left(30 \left| \begin{matrix} -1, -3; 5/2 \\ 1, 2, 2, 5; -1, -2 \end{matrix} \right. \right) = 2.900754807997581278090100786 \dots$$

Este resultado, com os dígitos mostrados, coincide com os calculados internamente pelo Mathematica, versões 4.2 e 5.0.

■

3.10. Convolução aditiva

A convolução aditiva de distribuições Gama com parâmetros gerais foi obtida por Mathai (1982), Moschopoulos (1985), Provost (1989), Sim (1992) e Hürliman (2001). O resultado é uma mistura de infinitas distribuições Gama.

Como um número infinito de distribuições é indesejável, nesta dissertação será usada a distribuição Erlang. A convolução aditiva de distribuições Erlang foi obtida por Mathai (1982). Ele usou a função

geradora de momentos, expansão em frações parciais e derivadas recursivas, obtendo uma mistura finita de distribuições Erlang. Veja também Sneddon (1972, p.244). Aqui será apresentado um resultado alternativo com uma demonstração similar à de Mathai, mas com base na transformada de Laplace.

Teorema 3.10.1. Convolução aditiva de distribuições Erlang com diferentes parâmetros de escala.

Sejam $X_1, X_2, \dots, X_D$, para $D > 1$, variáveis aleatórias independentes, que seguem, respectivamente, as densidades

$$\gamma_{b_1}^{a_1}(x_1), \gamma_{b_2}^{a_2}(x_2), \dots, \gamma_{b_D}^{a_D}(x_D),$$

onde $a_i > 0$, $b_i \in \{1, 2, \dots\}$ e $x_i \geq 0$ para todos os i , $1 \leq i \leq D$ e $a_i \neq a_j$ para $i \neq j$. Então $X_1 + X_2 + \dots + X_D$ segue a densidade

$$(\gamma_{b_1}^{a_1} \oplus \gamma_{b_2}^{a_2} \oplus \dots \oplus \gamma_{b_D}^{a_D})(x) = \sum_{d=1}^D \prod_{\substack{j=1 \\ j \neq d}}^D \left(1 - \frac{a_d}{a_j}\right)^{-b_j} \sum_{r=1}^{b_d} a_d^{b_d-r} M_d^{b_d-r} \gamma_r^{a_d}(x), \quad (3.92)$$

onde

$$M_d^u = \begin{cases} 1 & , \quad u = 0 \\ \frac{1}{u} \sum_{v=0}^{u-1} \left[\sum_{\substack{j=1 \\ j \neq d}}^D b_j (a_d - a_j)^{v-u} \right] M_d^v & , \quad u > 0 \end{cases}.$$

Prova. A transformada de Laplace da densidade Erlang é

$$\gamma_b^a(x) \xrightarrow{\mathcal{L}} a^b (a+s)^{-b}.$$

Pelo teorema da convolução,

$$(\gamma_{b_1}^{a_1} \oplus \dots \oplus \gamma_{b_D}^{a_D})(x) \xrightarrow{\mathcal{L}} a_1^{b_1} (a_1 + s)^{-b_1} \dots a_D^{b_D} (a_D + s)^{-b_D} = \left[\prod_{j=1}^D a_j^{b_j} \right] \prod_{d=1}^D (a_d + s)^{-b_d}. \quad (3.93)$$

Pela teoria de expansão em frações parciais,

$$\prod_{d=1}^D (a_d + s)^{-b_d} = \sum_{d=1}^D \sum_{r=1}^{b_d} c_{d,r} (s + a_d)^{-r},$$

onde $c_{d,r}$ são números reais que serão calculados. Eles são dados por

$$c_{d,r} = \lim_{s \rightarrow -a_d} \left\{ \frac{1}{(b_d - r)!} \frac{\partial^{b_d-r}}{\partial s^{b_d-r}} \left[(s + a_d)^{b_d} \prod_{j=1}^D (s + a_j)^{-b_j} \right] \right\} = \frac{1}{(b_d - r)!} \lim_{s \rightarrow -a_d} \frac{\partial^{b_d-r}}{\partial s^{b_d-r}} \left[\prod_{\substack{j=1 \\ j \neq d}}^D (s + a_j)^{-b_j} \right].$$

As funções auxiliares $\Delta_d(s)$ e $A_d(s)$ são definidas como

$$\Delta_d(s) = \Delta_d^{(0)}(s) = \prod_{\substack{j=1 \\ j \neq d}}^D (s + a_j)^{-b_j}$$

e

$$\begin{aligned}
A_d(s) &= A_d^{(0)}(s) = \frac{\partial}{\partial s} \ln(\Delta_d(s)) = \frac{1}{\Delta_d(s)} \frac{\partial}{\partial s} \Delta_d(s) \\
&= \frac{\partial}{\partial s} \ln \left(\prod_{\substack{j=1 \\ j \neq d}}^D (s + a_j)^{-b_j} \right) = \frac{\partial}{\partial s} \left(\sum_{\substack{j=1 \\ j \neq d}}^D -b_j \ln(s + a_j) \right) = - \sum_{\substack{j=1 \\ j \neq d}}^D \frac{b_j}{s + a_j} .
\end{aligned}$$

As derivadas de ordem superior dessas funções são

$$\begin{aligned}
\Delta_d^{(u)}(s) &= \frac{\partial^{u-1}}{\partial s^{u-1}} \frac{\partial}{\partial s} \Delta_d(s) = \frac{\partial^{u-1}}{\partial s^{u-1}} [A_d(s) \Delta_d(s)] = \sum_{v=0}^{u-1} \binom{u-1}{v} A_d^{(u-1-v)}(s) \Delta_d^{(v)}(s) \\
&= \begin{cases} \prod_{\substack{j=1 \\ j \neq d}}^D (s + a_j)^{-b_j} & , \quad u = 0 \\ \sum_{v=0}^{u-1} \binom{u-1}{v} A_d^{(u-1-v)}(s) \Delta_d^{(v)}(s) & , \quad u > 0 \end{cases}
\end{aligned}$$

e

$$A_d^{(k)}(s) = \frac{\partial^k}{\partial s^k} \left[- \sum_{\substack{j=1 \\ j \neq d}}^D \frac{b_j}{s + a_j} \right] = (-1)^{k+1} k! \sum_{\substack{j=1 \\ j \neq d}}^D b_j (s + a_j)^{-k-1} .$$

Como $\Delta_d^{(u)}(s)$ é uma função recursiva, é possível fazer

$$\Delta_d^{(u)}(s) = \prod_{\substack{i=1 \\ i \neq d}}^D (s + a_i)^{-b_i} L_d^u(s) ,$$

onde

$$\begin{aligned}
L_d^u(s) &= \begin{cases} 1 & , \quad u = 0 \\ \sum_{v=0}^{u-1} \binom{u-1}{v} A_d^{(u-1-v)}(s) L_d^v(s) & , \quad u > 0 \end{cases} \\
&= \begin{cases} 1 & , \quad u = 0 \\ \sum_{v=0}^{u-1} \binom{u-1}{v} (-1)^{u-v} (u-1-v)! \sum_{\substack{j=1 \\ j \neq d}}^D b_j (s + a_j)^{v-u} L_d^v(s) & , \quad u > 0 \end{cases} \\
&= \begin{cases} 1 & , \quad u = 0 \\ \sum_{v=0}^{u-1} (u-1)! \sum_{\substack{j=1 \\ j \neq d}}^D b_j (-s - a_j)^{v-u} \frac{L_d^v(s)}{v!} & , \quad u > 0 \end{cases} .
\end{aligned}$$

Definindo

$$\begin{aligned}
M_d^u &= \lim_{s \rightarrow -a_d} \frac{L_d^u(s)}{u!} \\
&= \begin{cases} 1/u! & , \quad u = 0 \\ \sum_{v=0}^{u-1} \frac{1}{u} \sum_{\substack{j=1 \\ j \neq d}}^D b_j (a_d - a_j)^{v-u} M_d^v & , \quad u > 0 \end{cases} \\
&= \begin{cases} 1 & , \quad u = 0 \\ \frac{1}{u} \sum_{v=0}^{u-1} \left[\sum_{\substack{j=1 \\ j \neq d}}^D b_j (a_d - a_j)^{v-u} \right] M_d^v & , \quad u > 0 \end{cases} ,
\end{aligned}$$

então

$$\begin{aligned}
c_{d,r} &= \frac{1}{(b_d - r)!} \lim_{s \rightarrow -a_d} \Delta_d^{(b_d-r)}(s) = \frac{1}{(b_d - r)!} \lim_{s \rightarrow -a_d} \prod_{\substack{j=1 \\ j \neq d}}^D (s + a_j)^{-b_j} L_d^{b_d-r}(s) \\
&= \frac{1}{(b_d - r)!} \prod_{\substack{j=1 \\ j \neq d}}^D (-a_d + a_j)^{-b_j} \lim_{s \rightarrow -a_d} L_d^{b_d-r}(s) = \left[\prod_{\substack{j=1 \\ j \neq d}}^D (a_j - a_d)^{-b_j} \right] M_d^{b_d-r} .
\end{aligned}$$

O lado direito de (3.93) se torna

$$\left[\prod_{j=1}^D a_j^{b_j} \right] \prod_{d=1}^D (a_d + s)^{-b_d} = \left[\prod_{j=1}^D a_j^{b_j} \right] \sum_{d=1}^D \sum_{r=1}^{b_d} c_{d,r} a_d^{-r} a_d^r (s + a_d)^{-r} .$$

Aplicando a transformada inversa de Laplace,

$$\begin{aligned}
(\gamma_{b_1}^{a_1} \oplus \dots \oplus \gamma_{b_D}^{a_D})(x) &= \left[\prod_{j=1}^D a_j^{b_j} \right] \sum_{d=1}^D \sum_{r=1}^{b_d} c_{d,r} a_d^{-r} \gamma_r^{a_d}(x) = \left[\prod_{j=1}^D a_j^{b_j} \right] \sum_{d=1}^D \sum_{r=1}^{b_d} \left[\prod_{\substack{j=1 \\ j \neq d}}^D (a_j - a_d)^{-b_j} \right] M_d^{b_d-r} a_d^{-r} \gamma_r^{a_d}(x) \\
&= \left[\prod_{j=1}^D a_j^{b_j} \right] \sum_{d=1}^D \left[\prod_{\substack{j=1 \\ j \neq d}}^D (a_j - a_d)^{-b_j} \right] \sum_{r=1}^{b_d} a_d^{-r} M_d^{b_d-r} \gamma_r^{a_d}(x) \\
&= \sum_{d=1}^D \left[\prod_{\substack{j=1 \\ j \neq d}}^D \left(1 - \frac{a_d}{a_j} \right)^{-b_j} \right] \sum_{r=1}^{b_d} a_d^{b_d-r} M_d^{b_d-r} \gamma_r^{a_d}(x) .
\end{aligned}$$

□

O resultado é uma mistura finita com no máximo $b_1 + b_2 + \dots + b_D$ distribuições Erlang. Os pesos da mistura podem assumir valores nulos ou negativos, mas sua soma deve ser 1. Os pesos, em módulo, podem ser números muito grandes, e a subtração de resultados intermediários aproximadamente iguais causa perda de precisão numérica em números de ponto flutuante nativos dos processadores. Por isso, é necessário cuidado na implementação computacional.

Exemplo 3.10.1. Convolução aditiva de Erlangs. Foi feita uma implementação do teorema anterior no Mathematica (**Apêndice A**) e foi calculada a convolução aditiva de uma distribuição Erlang com parâmetros $a_1 = 4$, $b_1 = 2$ e outra com parâmetros $a_2 = 5$, $b_2 = 3$. O resultado foi

$$(\gamma_2^4 \oplus \gamma_3^5)(x) = \int_0^x \gamma_2^4(t) \gamma_3^5(x-t) dt = -1500\gamma_1^4(x) + 125\gamma_2^4(x) + 1200\gamma_1^5(x) + 160\gamma_2^5(x) + 16\gamma_3^5(x).$$

■

O teorema anterior exclui o caso de distribuições Erlang com o mesmo parâmetro de escala. Este caso será tratado no teorema seguinte.

Teorema 3.10.2. Convolução aditiva de distribuições Erlang com parâmetros de escala iguais. Sejam $X_1, X_2, \dots, X_D$, para $D > 1$, variáveis aleatórias independentes que seguem, respectivamente, as densidades

$$\gamma_{b_1}^a(x_1), \gamma_{b_2}^a(x_2), \dots, \gamma_{b_D}^a(x_D),$$

onde $a > 0$, $b_d \in \{1, 2, \dots\}$ e $x_d \geq 0$ para todos os d , $1 \leq d \leq D$. Então $X_1 + X_2 + \dots + X_D$ segue a densidade

$$(\gamma_{b_1}^a \oplus \gamma_{b_2}^a \oplus \dots \oplus \gamma_{b_D}^a)(x) = \gamma_{b_1+b_2+\dots+b_D}^a(x) \quad (3.94)$$

Prova. A transformada de Laplace da densidade Erlang é

$$\gamma_b^a(x) \xrightarrow{\mathcal{L}} a^b (a+s)^{-b}.$$

Pelo teorema da convolução,

$$(\gamma_{b_1}^a \oplus \dots \oplus \gamma_{b_D}^a)(x) \xrightarrow{\mathcal{L}} a^{b_1} (a+s)^{-b_1} \dots a^{b_D} (a+s)^{-b_D} = a^{b_1+\dots+b_D} (a+s)^{-b_1-\dots-b_D}.$$

Mas

$$\gamma_{b_1+\dots+b_D}^a(x) \xrightarrow{\mathcal{L}} a^{b_1+\dots+b_D} (a+s)^{-b_1-\dots-b_D}.$$

□

Comentário 3.10.1. O teorema anterior também vale para a distribuição Gama.

3.11. Combinação linear positiva

Se variáveis aleatórias independentes seguem, cada uma, uma mistura finita de Erlangs, então sua densidade conjunta é dada pelo seu produto, gerando um produto tensorial finito de Erlangs. Além disso, como foi mostrado no **Exemplo 3.5.2**, o produto tensorial finito de Erlangs também pode representar a densidade conjunta de variáveis aleatórias dependentes.

Se a densidade multivariada de variáveis aleatórias (dependentes ou independentes) é um produto tensorial de Erlangs, então a combinação linear positiva dessas variáveis aleatórias será uma mistura finita de Erlangs. Note que na combinação linear, a soma dos pesos não precisa ser igual a 1. Considere inicialmente uma parcela do produto tensorial. Então a densidade de uma combinação linear positiva pode ser calculada com o teorema a seguir.

Teorema 3.11.1. Combinação linear positiva de variáveis aleatórias seguindo uma parcela do produto tensorial de Erlangs. Sejam $[X_1, X_2, \dots, X_D]^T$, onde $D > 1$, variáveis aleatórias com uma distribuição conjunta dada pela densidade

$$\gamma_{b_1}^{a_1}(x_1) \gamma_{b_2}^{a_2}(x_2) \dots \gamma_{b_D}^{a_D}(x_D),$$

onde $a_d > 0$, $b_d \in \{1, 2, \dots\}$ e $x_d \geq 0$ para todos os d , $1 \leq d \leq D$. Então a combinação linear positiva $w_1 X_1 + w_2 X_2 + \dots + w_D X_D$, onde $w_1 > 0, w_2 > 0, \dots, w_D > 0$, segue a densidade dada por

$$\left(\gamma_{b_1}^{a_1/w_1} \oplus \gamma_{b_2}^{a_2/w_2} \oplus \dots \oplus \gamma_{b_D}^{a_D/w_D} \right)(x).$$

Prova. Olhando para a figura a seguir, para o caso $D = 2$, a massa abaixo da linha $w_1 x_1 + w_2 x_2 = t$ é dada por

$$\int_0^{\frac{t}{w_1}} \int_0^{\frac{t-w_1 x_1}{w_2}} \gamma_{b_1}^{a_1}(x_1) \gamma_{b_2}^{a_2}(x_2) dx_2 dx_1.$$

Fazendo a transformação $w_1 x_1 = u$ e $w_2 x_2 = v$, então $w_1 dx_1 = du$ e $w_2 dx_2 = dv$. A integral se torna

$$\int_0^t \int_0^{t-u} \gamma_{b_1}^{a_1} \left(\frac{u}{w_1} \right) \gamma_{b_2}^{a_2} \left(\frac{v}{w_2} \right) \frac{dv du}{w_2 w_1}.$$

Aplicando a propriedade (3.24), ela se torna

$$\int_0^t \int_0^{t-u} \gamma_{b_1}^{a_1/w_1}(u) \gamma_{b_2}^{a_2/w_2}(v) dv du,$$

que é a função acumulada da densidade

$$\left(\gamma_{b_1}^{a_1/w_1} \oplus \gamma_{b_2}^{a_2/w_2} \right)(t).$$

Para $D > 2$, a prova é similar.

Figura 3.11.1. Combinação linear positiva de duas variáveis aleatórias.

□

Corolário 3.11.1. Combinação linear positiva de variáveis aleatórias seguindo um produto tensorial finito de Erlangs. Sejam $[X_1, X_2, \dots, X_D]^T$, onde $D > 1$, variáveis aleatórias com uma densidade conjunta dada por

$$\sum_{i=1}^n p_i \gamma_{b_{i,1}}^{a_{i,1}}(x_1) \gamma_{b_{i,2}}^{a_{i,2}}(x_2) \dots \gamma_{b_{i,D}}^{a_{i,D}}(x_D),$$

onde $a_{i,d} > 0$, $b_{i,d} \in \{1, 2, \dots\}$ e $x_d \geq 0$ para todos os i e d , $1 \leq i \leq n$, $1 \leq d \leq D$, $p_i \in \mathbb{R}$ e $\sum_{i=1}^n p_i = 1$. Então a combinação linear positiva $w_1 \mathbf{X}_1 + w_2 \mathbf{X}_2 + \dots + w_D \mathbf{X}_D$, onde $w_1 > 0, w_2 > 0, \dots, w_D > 0$, segue a densidade dada por

$$\sum_{i=1}^n p_i \left(\gamma_{b_{i,1}}^{a_{i,1}/w_1} \oplus \gamma_{b_{i,2}}^{a_{i,2}/w_2} \oplus \dots \oplus \gamma_{b_{i,D}}^{a_{i,D}/w_D} \right)(x).$$

Prova. Para o caso $D = 2$, a função acumulada é

$$\int_0^{\frac{t}{w_1}} \int_0^{\frac{t-w_1 x_1}{w_2}} \sum_{i=1}^n p_i \gamma_{b_{i,1}}^{a_{i,1}}(x_1) \gamma_{b_{i,2}}^{a_{i,2}}(x_2) dx_2 dx_1 = \sum_{i=1}^n p_i \int_0^{\frac{t}{w_1}} \int_0^{\frac{t-w_1 x_1}{w_2}} \gamma_{b_{i,1}}^{a_{i,1}}(x_1) \gamma_{b_{i,2}}^{a_{i,2}}(x_2) dx_2 dx_1,$$

e que, pelo teorema anterior, é também a função acumulada de

$$\sum_{i=1}^n p_i \left(\gamma_{b_{i,1}}^{a_{i,1}/w_1} \oplus \gamma_{b_{i,2}}^{a_{i,2}/w_2} \right)(t).$$

Para $D > 2$, a prova é similar. □

O teorema e o corolário anteriores também valem para o produto tensorial finito de Gamas. Entretanto, a convolução aditiva de Gamas gera uma mistura com infinitas densidades Gama, enquanto a convolução aditiva de Erlangs gera uma mistura finita de Erlangs. Este foi o motivo para a adoção da distribuição Erlang e não da Gama.

Exemplo 3.11.1. Combinação linear positiva de variáveis aleatórias seguindo um produto tensorial finito de Erlangs. Para um caso particular de produto tensorial bivariado de Erlangs, foi calculada a densidade de uma combinação linear positiva. Se

$$[\mathbf{X}_1, \mathbf{X}_2]^T \sim m(x_1, x_2) = 0.4 \gamma_2^4(x_1) \gamma_3^5(x_2) + 0.6 \gamma_3^6(x_1) \gamma_2^3(x_2),$$

então

$$\begin{aligned} w_1 \mathbf{X}_1 + w_2 \mathbf{X}_2 \sim f(t) &= 0.4 \left(\gamma_2^{4/w_1} \oplus \gamma_3^{5/w_2} \right)(t) + 0.6 \left(\gamma_3^{6/w_1} \oplus \gamma_2^{3/w_2} \right)(t) = \\ &= \frac{d}{dt} \left[\int_0^{\frac{t}{w_1}} \left[\int_0^{\frac{t-w_1 x_1}{w_2}} m(x_1, x_2) dx_2 \right] dx_1 \right]. \end{aligned}$$

Para $w_1 = 0.2$, $w_2 = 0.5$,

$$\begin{aligned} 0.2 \mathbf{X}_1 + 0.5 \mathbf{X}_2 \sim f(t) &= -\frac{225}{256} \gamma_1^6(t) + \frac{75}{64} \gamma_2^6(t) + \frac{24}{5} \gamma_1^{10}(t) - \frac{16}{5} \gamma_2^{10}(t) + \frac{8}{5} \gamma_3^{10}(t) \\ &\quad - \frac{12}{5} \gamma_1^{20}(t) - \frac{2}{5} \gamma_2^{20}(t) + \frac{45}{256} \gamma_1^{30}(t) + \frac{3}{32} \gamma_2^{30}(t) + \frac{3}{80} \gamma_3^{30}(t). \end{aligned}$$

A implementação em Mathematica está no **Apêndice A**. ■

3.12. Convolução multiplicativa

A convolução multiplicativa de duas variáveis aleatórias independentes que seguem densidades Gama pode ser encontrada em Johnson, Kotz e Balakrishnan (1994, p.385).

A convolução multiplicativa de duas variáveis aleatórias independentes que seguem densidades Erlang foi obtida por Mathai (1972), e será dada a seguir.

Teorema 3.12.1. Convolução multiplicativa de distribuições Erlang. Sejam $X_1, X_2, \dots, X_D$, para $D > 1$, variáveis aleatórias independentes que seguem, respectivamente, as densidades

$$\gamma_{b_1}^{a_1}(x_1), \gamma_{b_2}^{a_2}(x_2), \dots, \gamma_{b_D}^{a_D}(x_D),$$

onde $a_d > 0$, $b_d \in \{1, 2, \dots\}$ e $x_d \geq 0$ para todos os d , $1 \leq d \leq D$. Então $X_1 X_2 \dots X_D$ segue a densidade

$$(\gamma_{b_1}^{a_1} \otimes \gamma_{b_2}^{a_2} \otimes \dots \otimes \gamma_{b_D}^{a_D})(x) = \frac{a_1 a_2 \dots a_D}{\Gamma(b_1) \Gamma(b_2) \dots \Gamma(b_D)} G_{0,D}^{D,0} \left(a_1 a_2 \dots a_D x \middle| \begin{matrix} - \\ b_1 - 1 \quad b_2 - 1 \quad \dots \quad b_D - 1 \end{matrix} \right). \quad (3.95)$$

Prova. A transformada de Mellin da densidade Erlang é

$$\gamma_b^a(x) = \frac{a^b x^{b-1}}{\Gamma(b) e^{ax}} = \frac{a}{\Gamma(b)} G_{0,1}^{1,0} \left(ax \middle| \begin{matrix} - \\ b-1 \end{matrix} \right) \xrightarrow{\mathcal{M}} \frac{\Gamma(b-1+s)}{a^{s-1} \Gamma(b)}.$$

Pelo teorema da convolução,

$$(\gamma_{b_1}^{a_1} \otimes \dots \otimes \gamma_{b_D}^{a_D})(x) \xrightarrow{\mathcal{M}} \frac{\Gamma(b_1-1+s)}{a_1^{s-1} \Gamma(b_1)} \dots \frac{\Gamma(b_D-1+s)}{a_D^{s-1} \Gamma(b_D)} = \frac{\Gamma(b_1-1+s) \dots \Gamma(b_D-1+s)}{(a_1 \dots a_D)^{s-1} \Gamma(b_1) \dots \Gamma(b_D)}.$$

Mas

$$G_{0,D}^{D,0} \left(\lambda x \middle| \begin{matrix} - \\ b_1 \quad \dots \quad b_D \end{matrix} \right) \xrightarrow{\mathcal{M}} \frac{1}{\lambda^s} \prod_{i=1}^D \Gamma(b_i + s),$$

ou

$$\begin{aligned} \frac{a_1 \dots a_D}{\Gamma(b_1) \dots \Gamma(b_D)} G_{0,D}^{D,0} \left(a_1 \dots a_D x \middle| \begin{matrix} - \\ b_1 - 1 \quad \dots \quad b_D - 1 \end{matrix} \right) &\xrightarrow{\mathcal{M}} \frac{a_1 \dots a_D (a_1 \dots a_D)^{-s}}{\Gamma(b_1) \dots \Gamma(b_D)} \prod_{i=1}^D \Gamma(b_i - 1 + s) = \\ &= \frac{\Gamma(b_1 - 1 + s) \dots \Gamma(b_D - 1 + s)}{(a_1 \dots a_D)^{s-1} \Gamma(b_1) \dots \Gamma(b_D)}. \end{aligned}$$

□

Comentário 3.12.1. A classe de distribuições Meijer G é fechada em relação à convolução multiplicativa.

3.13. Potenciação

Se uma variável aleatória segue uma distribuição Fox H, então sua potência pode ser calculada com

o teorema a seguir. Note que a elevação de uma variável aleatória a uma potência n gera um resultado (em geral) diferente da multiplicação de n variáveis, pois a potenciação é feita sempre sobre a mesma variável, enquanto a multiplicação é feita com variáveis aleatórias diferentes (dependentes ou independentes).

Teorema 3.13.1. Potenciação de distribuições Fox H. Se X é uma variável aleatória que segue a densidade $f(x)$ dada por

$$f(x) = kH_{p,q}^{m,n} \left(\lambda x \left[\begin{matrix} a_p, A_p \\ b_q, B_q \end{matrix} \right] \right),$$

onde $\lambda > 0$, então sua potência $X^{1/t}$, $t > 0$, segue a distribuição dada pela densidade

$$f^{\otimes \frac{1}{t}}(x) = tx^{t-1} f(x^t) = tx^{t-1} kH_{p,q}^{m,n} \left(\lambda x^t \left[\begin{matrix} a_p, A_p \\ b_q, B_q \end{matrix} \right] \right) = \lambda^{\frac{1}{t}-1} kH_{p,q}^{m,n} \left(\lambda^{\frac{1}{t}} x \left[\begin{matrix} a_p + A_p - \frac{A_p}{t}, \frac{A_p}{t} \\ b_q + B_q - \frac{B_q}{t}, \frac{B_q}{t} \end{matrix} \right] \right). \quad (3.96)$$

Prova 1. Se X segue $f(x)$, então

$$X \sim f(x) = kH_{p,q}^{m,n} \left(\lambda x \left[\begin{matrix} a_p, A_p \\ b_q, B_q \end{matrix} \right] \right) \xrightarrow{\mathcal{M}} F(s) = k\lambda^{-s} \frac{\prod_{i=1}^m \Gamma(b_i + B_i s) \prod_{j=1}^n \Gamma(1 - a_j - A_j s)}{\prod_{i=m+1}^q \Gamma(1 - b_i - B_i s) \prod_{j=n+1}^p \Gamma(a_j + A_j s)}$$

e

$$\begin{aligned} X^{\frac{1}{t}} \sim f^{\otimes \frac{1}{t}}(x) &= tx^{t-1} f(x^t) \xrightarrow{\mathcal{M}} F\left(1 - \frac{1}{t} + \frac{s}{t}\right) = \\ &= k\lambda^{-1 + \frac{1}{t} - \frac{s}{t}} \frac{\prod_{i=1}^m \Gamma\left(b_i + B_i - \frac{B_i}{t} + \frac{B_i s}{t}\right) \prod_{j=1}^n \Gamma\left(1 - a_j - A_j + \frac{A_j}{t} - \frac{A_j s}{t}\right)}{\prod_{i=m+1}^q \Gamma\left(1 - b_i - B_i + \frac{B_i}{t} - \frac{B_i s}{t}\right) \prod_{j=n+1}^p \Gamma\left(a_j + A_j - \frac{A_j}{t} + \frac{A_j s}{t}\right)}. \end{aligned}$$

Aplicando a transformada inversa de Mellin,

$$X^{1/t} \sim f^{\otimes \frac{1}{t}}(x) = tx^{t-1} f(x^t) = tx^{t-1} kH_{p,q}^{m,n} \left(\lambda x^t \left[\begin{matrix} a_p, A_p \\ b_q, B_q \end{matrix} \right] \right) = \lambda^{\frac{1}{t}-1} kH_{p,q}^{m,n} \left(\lambda^{\frac{1}{t}} x \left[\begin{matrix} a_p + A_p - \frac{A_p}{t}, \frac{A_p}{t} \\ b_q + B_q - \frac{B_q}{t}, \frac{B_q}{t} \end{matrix} \right] \right).$$

□

Prova 2. Usando as propriedades (3.84) e (3.85),

$$\begin{aligned}
X^{1/t} &\sim f^{\otimes \frac{1}{t}}(x) = tx^{t-1}f(x^t) = tx^{t-1}kH_{p,q}^{m,n}\left(\lambda x^t \begin{bmatrix} a_p, A_p \\ b_q, B_q \end{bmatrix}\right) = \frac{tx^t(\lambda x^t)^{-1/t}}{\lambda^{-1/t}}kH_{p,q}^{m,n}\left(\lambda x^t \begin{bmatrix} a_p, A_p \\ b_q, B_q \end{bmatrix}\right) = \\
&= \frac{tx^t}{\lambda^{-1/t}}kH_{p,q}^{m,n}\left(\lambda x^t \begin{bmatrix} a_p - \frac{A_p}{t}, A_p \\ b_q - \frac{B_q}{t}, B_q \end{bmatrix}\right) = \frac{t(\lambda x^t)}{\lambda^{-1/t}\lambda}kH_{p,q}^{m,n}\left(\lambda x^t \begin{bmatrix} a_p - \frac{A_p}{t}, A_p \\ b_q - \frac{B_q}{t}, B_q \end{bmatrix}\right) = \\
&= \frac{t}{\lambda^{1-1/t}}kH_{p,q}^{m,n}\left(\lambda x^t \begin{bmatrix} a_p + A_p - \frac{A_p}{t}, \frac{A_p t}{t} \\ b_q + B_q - \frac{B_q}{t}, \frac{B_q t}{t} \end{bmatrix}\right) = \lambda^{\frac{1}{t}-1}kH_{p,q}^{m,n}\left(\lambda^{\frac{1}{t}}x \begin{bmatrix} a_p + A_p - \frac{A_p}{t}, \frac{A_p}{t} \\ b_q + B_q - \frac{B_q}{t}, \frac{B_q}{t} \end{bmatrix}\right).
\end{aligned}$$

□

Comentário 3.13.1. A classe de distribuições Fox H é fechada em relação à convolução multiplicativa e à potenciação; veja Carter e Springer (1977).

3.14. Valor esperado

Nesta seção, será calculado o valor esperado de algumas distribuições particulares que serão usadas neste trabalho.

Teorema 3.14.1. Valor esperado de uma variável aleatória que segue uma distribuição Erlang. Se X segue a densidade

$$\gamma_b^a(x),$$

onde $a > 0$, $b \in \{1, 2, \dots\}$ e $x \geq 0$, então o valor esperado de X é

$$E[X] = \frac{b}{a}.$$

Prova. A transformada de Mellin da densidade Erlang é

$$\gamma_b^a(x) \xrightarrow{\mathcal{M}} \int_0^\infty x^{s-1} \gamma_b^a(x) dx = \frac{\Gamma(b-1+s)}{a^{s-1}\Gamma(b)}.$$

Fazendo $s = 2$, então

$$\int_0^\infty x^{2-1} \gamma_b^a(x) dx = \int_0^\infty x \gamma_b^a(x) dx = \frac{\Gamma(b-1+2)}{a^{2-1}\Gamma(b)} = \frac{\Gamma(b+1)}{a\Gamma(b)} = \frac{b}{a}.$$

□

Corolário 3.14.1. Valor esperado de uma variável aleatória seguindo uma mistura finita de Erlangs. Se X segue a densidade

$$p_1 \gamma_{b_1}^{a_1}(x) + p_2 \gamma_{b_2}^{a_2}(x) + \dots + p_n \gamma_{b_n}^{a_n}(x),$$

onde $p_i \in \mathbb{R}$, para $1 \leq i \leq n$, $\sum_{i=1}^n p_i = 1$; $a_i > 0$, $b_i \in \{1, 2, \dots\}$ para $1 \leq i \leq n$; e $x \geq 0$; então o valor esperado de X é

$$E[X] = p_1 \frac{b_1}{a_1} + p_2 \frac{b_2}{a_2} + \dots + p_n \frac{b_n}{a_n}.$$

Prova. Trivial. □

Comentário 3.14.1. O corolário e o teorema anteriores também valem para a distribuição Gama.

Corolário 3.14.2. Valor esperado de uma variável aleatória seguindo uma convolução aditiva de Erlangs. Se X segue a densidade

$$(\gamma_{b_1}^{a_1} \oplus \gamma_{b_2}^{a_2} \oplus \dots \oplus \gamma_{b_D}^{a_D})(x),$$

onde $x \geq 0$; e $a_i > 0$, $b_i \in \{1, 2, \dots\}$ para todos os i , $1 \leq i \leq n$; então o valor esperado de X é

$$E[X] = \frac{b_1}{a_1} + \frac{b_2}{a_2} + \dots + \frac{b_D}{a_D}.$$

Prova. Trivial. □

Corolário 3.14.3. Valor esperado de uma variável aleatória seguindo uma mistura finita de convoluções aditivas de Erlangs com parâmetros de escala constantes. Se X segue a densidade

$$\sum_{i=1}^n p_i (\gamma_{b_{i,1}}^{a_1} \oplus \gamma_{b_{i,2}}^{a_2} \oplus \dots \oplus \gamma_{b_{i,D}}^{a_D})(x),$$

onde $p_i \in \mathbb{R}$, para $1 \leq i \leq n$, $\sum_{i=1}^n p_i = 1$; $x \geq 0$; e $a_i > 0$, $b_i \in \{1, 2, \dots\}$ para todos os i , $1 \leq i \leq n$; então o valor esperado de X é

$$E[X] = \sum_{i=1}^n p_i \left(\frac{b_{i,1}}{a_1} + \frac{b_{i,2}}{a_2} + \dots + \frac{b_{i,D}}{a_D} \right).$$

Prova. Trivial. □

Corolário 3.14.4. Valor esperado de uma variável aleatória que segue um caso particular de distribuição com a função Meijer G. Se X segue a densidade

$$kG_{0,D}^{D,0} \left(cx \left| \begin{matrix} - \\ b_1 \ b_2 \ \dots \ b_D \end{matrix} \right. \right),$$

onde $k > 0$, $c > 0$, $x \geq 0$ e $b_1, \dots, b_D \in \{1, 2, \dots\}$, para $D \geq 1$; então o valor esperado de X é

$$E[X] = \frac{k}{c^2} \prod_{i=1}^D \Gamma(b_i + 1). \quad (3.97)$$

Prova. A transformada de Mellin da densidade é

$$kG_{0,D}^{D,0}\left(cx \middle| \begin{matrix} - \\ b_1 \ b_2 \ \dots \ b_D \end{matrix} \right) \xrightarrow{\mathcal{M}} kc^{-s} \prod_{i=1}^D \Gamma(b_i - 1 + s).$$

Fazendo $s = 2$, o resultado é obtido. □

Corolário 3.14.5. Valor esperado de uma variável aleatória que segue uma mistura de um caso particular de distribuição com a função Meijer G. Se X segue a densidade

$$\sum_{i=1}^n p_i k_i G_{0,D_i}^{D_i,0}\left(c_i x \middle| \begin{matrix} - \\ b_{i,1} \ b_{i,2} \ \dots \ b_{i,D_i} \end{matrix} \right),$$

onde $p_i \in \mathbb{R}$, para $1 \leq i \leq n$, $\sum_{i=1}^n p_i = 1$; $x \geq 0$; $k_i > 0$, $c_i > 0$, $b_{i,1}, \dots, b_{i,D_i} \in \{1, 2, \dots\}$ e $D_i \geq 1$, para todos os i , $1 \leq i \leq n$; então o valor esperado de X é

$$E[X] = \sum_{i=1}^n p_i \frac{k_i}{c_i^2} \prod_{j=1}^{D_i} \Gamma(b_{i,j} + 1). \quad (3.98)$$

Prova. Trivial. □

3.15. *Downside Risk*

O *Downside Risk* é uma medida de risco usada em Finanças e que será definida e vista com detalhes no capítulo seguinte. Ela será definida a seguir para o contexto corrente.

Definição 3.15.1. *Downside Risk.* O *Downside Risk*, com um retorno de referência $R_B > 0$ e um coeficiente de aversão a retornos indesejáveis $\alpha > 0$ de uma variável aleatória X que segue a densidade $f(x)$, $x \in \mathbb{R}$, é definido como

$$DR_{\alpha, R_B}[X] = \left(\int_{-\infty}^{R_B} (R_B - x)^\alpha f(x) dx \right)^{1/\alpha}. \quad (3.99)$$

Comentário 3.15.1. O limite inferior da integral é definido como $-\infty$ para permitir carteiras alavancadas, mas se este não for o caso, ele pode ser substituído por 0. Neste trabalho, o limite inferior sempre será 0. Para haver sentido financeiro, é necessário que $1 \leq 1 + r_f \leq R_B$ e $\alpha \geq 1$, onde $r_f \geq 0$ é a taxa de retorno do ativo livre de risco (com variância nula).

Comentário 3.15.2. A definição de *Downside Risk* tem semelhança com a definição da integral fracionária; veja Mainardi, Paridisi e Gorenflo (1998, p.20).

A seguir, será calculado o *Downside Risk* de algumas distribuições particulares que serão usadas neste trabalho.

Corolário 3.15.1. *Downside Risk* de uma variável aleatória que segue uma distribuição Erlang. Se X segue a densidade $\gamma_b^a(x)$, onde $a > 0$, $b \in \{1, 2, \dots\}$ e $x \geq 0$, então o *Downside Risk* de X com parâmetros α e R_B é

$$DR_{\alpha, R_B}[X] = R_B \left(\frac{\Gamma(1+\alpha)}{\Gamma(b)} G_{1,2}^{1,1} \left(aR_B \left| \begin{matrix} 1; \\ b; -\alpha \end{matrix} \right. \right) \right)^{1/\alpha}. \quad (3.100)$$

Prova. Considerando-se que

$$\gamma_b^a(x) = \frac{a^b x^{b-1}}{\Gamma(b) e^{ax}} = \frac{a}{\Gamma(b)} G_{0,1}^{1,0} \left(ax \left| \begin{matrix} - \\ b-1 \end{matrix} \right. \right)$$

e aplicando as propriedades (3.88) e (3.81), obtém-se

$$\begin{aligned} DR_{\alpha, R_B}^\alpha[X] &= \int_0^{R_B} (R_B - x)^\alpha \gamma_b^a(x) dx = \frac{a}{\Gamma(b)} \int_0^{R_B} (R_B - x)^\alpha G_{0,1}^{1,0} \left(ax \left| \begin{matrix} - \\ b-1 \end{matrix} \right. \right) dx = \\ &= \frac{a}{\Gamma(b)} \Gamma(1+\alpha) R_B^{1+\alpha} G_{1,2}^{1,1} \left(aR_B \left| \begin{matrix} 0; \\ b-1; -1-\alpha \end{matrix} \right. \right) = \frac{\Gamma(1+\alpha) R_B^\alpha}{\Gamma(b)} G_{1,2}^{1,1} \left(aR_B \left| \begin{matrix} 1; \\ b; -\alpha \end{matrix} \right. \right). \end{aligned}$$

□

Casos particulares do *Downside Risk* ocorrem quando $\alpha = 1$ ou $\alpha = 2$.

Corolário 3.15.2. *Downside Risk* com $\alpha = 1$ de uma variável aleatória que segue uma distribuição Erlang. Se X segue a densidade $\gamma_b^a(x)$, onde $a > 0$, $b \in \{1, 2, \dots\}$ e $x \geq 0$, então o *Downside Risk* de X com parâmetros $\alpha = 1$ e $R_B > 0$ é

$$DR_{1, R_B}[X] = \left(R_B - \frac{b}{a} \right) \left(1 - \sum_{i=1}^b \gamma_i^1(aR_B) \right) + \frac{b}{a} \gamma_{b+1}^1(aR_B). \quad (3.101)$$

Prova. Usando o Mathematica e a propriedade (3.32),

$$\begin{aligned} DR_{1, R_B}[X] &= \int_0^{R_B} (R_B - x) \gamma_b^a(x) dx = -\frac{b}{a} + R_B - R_B \frac{\Gamma(b, aR_B)}{\Gamma(b)} + \frac{\Gamma(b+1, aR_B)}{a\Gamma(b)} \\ &= R_B - \frac{b}{a} - R_B \sum_{i=1}^b \gamma_i^1(aR_B) + \frac{b}{a} \sum_{i=1}^{b+1} \gamma_i^1(aR_B) \\ &= \left(R_B - \frac{b}{a} \right) \left(1 - \sum_{i=1}^b \gamma_i^1(aR_B) \right) + \frac{b}{a} \gamma_{b+1}^1(aR_B). \end{aligned}$$

□

Corolário 3.15.3. *Downside Risk* com $\alpha = 2$ de uma variável aleatória que segue uma distribuição Erlang. Se X segue a densidade $\gamma_b^a(x)$, onde $a > 0$, $b \in \{1, 2, \dots\}$ e $x \geq 0$, então o *Downside Risk* de X com parâmetros $\alpha = 2$ e $R_B > 0$ é

$$DR_{2, R_B}[X] = \sqrt{\frac{b}{a^2} + \left(R_B - \frac{b}{a} \right)^2} \left[1 - \sum_{i=1}^{b+1} \gamma_i^1(aR_B) \right] + \left(\frac{b+1}{a} \right) \left(R_B - \frac{b}{a} \right) \gamma_{b+2}^1(aR_B). \quad (3.102)$$

Prova. Usando o Mathematica e a propriedade (3.32),

$$\begin{aligned}
DR_{2,R_B}^2[X] &= \int_0^{R_B} (R_B - x)^2 \gamma_b^a(x) dx \\
&= \left[\frac{b}{a^2} + \left(R_B - \frac{b}{a} \right)^2 \right] + \left(\frac{R_B}{a} - \frac{bR_B}{a} + R_B^2 \right) \gamma_{b+2}^1(aR_B) - \left[\frac{b}{a^2} + \left(R_B - \frac{b}{a} \right)^2 \right] \frac{\Gamma(b+2, aR_B)}{\Gamma(b+2)} \\
&= \left[\frac{b}{a^2} + \left(R_B - \frac{b}{a} \right)^2 \right] \left[1 - \frac{\Gamma(b+2, aR_B)}{\Gamma(b+2)} \right] + R_B \left(\frac{1-b}{a} + R_B \right) \gamma_{b+2}^1(aR_B) \\
&= \left[\frac{b}{a^2} + \left(R_B - \frac{b}{a} \right)^2 \right] \left[1 - \sum_{i=1}^{b+2} \gamma_i^1(aR_B) \right] + R_B \left(\frac{1-b}{a} + R_B \right) \gamma_{b+2}^1(aR_B) \\
&= \left[\frac{b}{a^2} + \left(R_B - \frac{b}{a} \right)^2 \right] \left[1 - \sum_{i=1}^{b+1} \gamma_i^1(aR_B) \right] + \left(\frac{b+1}{a} \right) \left(R_B - \frac{b}{a} \right) \gamma_{b+2}^1(aR_B) .
\end{aligned}$$

□

O *Downside Risk* de uma variável aleatória que segue uma distribuição Meijer G pode ser calculado usando fórmulas encontradas em Prudnikov *et al.* (1990, p.348) ou Erdélyi *et al.* (1954, p.417). Alguns casos serão dados a seguir.

Corolário 3.15.4. *Downside Risk* de uma variável aleatória que segue um caso particular da distribuição Meijer G. Se X segue a densidade

$$\frac{k_A}{k_B} G_{0,D}^{D,0} \left(k_A x \left| \begin{matrix} - \\ b_1 - 1 \ b_2 - 1 \ \dots \ b_D - 1 \end{matrix} \right. \right),$$

onde $k_A > 0$, $k_B > 0$; $b_1, \dots, b_D \in \{1, 2, \dots\}$, $D \geq 1$; e $x \geq 0$; então o *Downside Risk* de X com parâmetros $\alpha > 0$ e $R_B > 0$ é

$$DR_{\alpha, R_B}[X] = R_B \left(\frac{\Gamma(\alpha+1)}{k_B} G_{1,D+1}^{D,1} \left(k_A R_B \left| \begin{matrix} 1; \\ b_1 \ b_2 \ \dots \ b_D; -\alpha \end{matrix} \right. \right) \right)^{1/\alpha}. \quad (3.103)$$

Prova. Aplicando a fórmula (3.88) e a propriedade (3.81), obtém-se

$$\begin{aligned}
DR_{\alpha, R_B}^\alpha[X] &= \frac{k_A}{k_B} R_B^{\alpha+1} \Gamma(\alpha+1) G_{1,D+1}^{D,1} \left(k_A R_B \left| \begin{matrix} 0; \\ b_1 - 1 \ b_2 - 1 \ \dots \ b_D - 1; -\alpha - 1 \end{matrix} \right. \right) \\
&= \frac{R_B^\alpha \Gamma(\alpha+1)}{k_B} G_{1,D+1}^{D,1} \left(k_A R_B \left| \begin{matrix} 1; \\ b_1 \ b_2 \ \dots \ b_D; -\alpha \end{matrix} \right. \right).
\end{aligned}$$

□

Corolário 3.15.5. *Downside Risk* da potência $1/t$ de uma variável aleatória que segue uma função Meijer G. Se $t \in \{1, 2, \dots\}$, $k, \lambda \in \mathbb{R}$ e

$$X \sim k G_{p,q}^{m,n} \left(\lambda x \left| \begin{matrix} (a_p) \\ (b_q) \end{matrix} \right. \right),$$

então

$$DR_{\alpha, R_B}^{\alpha} [X^{1/t}] = kR_B^t \Gamma(1+\alpha) \left(\frac{R_B}{t} \right)^{\alpha} G_{p+t, q+t}^{m, n+t} \left(\lambda R_B^t \left| \begin{array}{c} \frac{1-t}{t}, \frac{2-t}{t}, \dots, \frac{0}{t}, (a_p) \\ (b_q), \frac{-t-\alpha}{t}, \frac{1-t-\alpha}{t}, \dots, \frac{-1-\alpha}{t} \end{array} \right. \right). \quad (3.104)$$

Prova. Se

$$X \sim f(x) = kG_{p,q}^{m,n} \left(\lambda x \left| \begin{array}{c} (a_p) \\ (b_q) \end{array} \right. \right),$$

então

$$X^{1/t} \sim tx^{t-1} f(x^t) = ktx^{t-1} G_{p,q}^{m,n} \left(\lambda x^t \left| \begin{array}{c} (a_p) \\ (b_q) \end{array} \right. \right),$$

e aplicando a propriedade (3.88), obtém-se

$$\begin{aligned} DR_{\alpha, R_B}^{\alpha} [X^{1/t}] &= \int_0^{R_B} (R_B - x)^{\alpha} ktx^{t-1} G_{p,q}^{m,n} \left(\lambda x^t \left| \begin{array}{c} (a_p) \\ (b_q) \end{array} \right. \right) dx \\ &= kt \frac{\Gamma(\alpha+1)}{t^{\alpha+1} R_B^{-t-\alpha}} G_{p+t, q+t}^{m, n+t} \left(\lambda R_B^t \left| \begin{array}{c} \Delta(t, 1-t), (a_p) \\ (b_q), \Delta(t, -\alpha-t) \end{array} \right. \right) \\ &= kR_B^t \Gamma(1+\alpha) \left(\frac{R_B}{t} \right)^{\alpha} G_{p+t, q+t}^{m, n+t} \left(\lambda R_B^t \left| \begin{array}{c} \frac{1-t}{t}, \frac{2-t}{t}, \dots, \frac{0}{t}, (a_p) \\ (b_q), \frac{-t-\alpha}{t}, \frac{1-t-\alpha}{t}, \dots, \frac{-1-\alpha}{t} \end{array} \right. \right). \end{aligned}$$

□

Corolário 3.15.6. Downside Risk da potência $1/t$ do produto de t variáveis aleatórias independentes seguindo densidades Erlang. Se $t \in \{1, 2, \dots\}$ e $X_1, X_2, \dots, X_t$ são variáveis aleatórias independentes que seguem, respectivamente, as densidades

$$\gamma_{b_1}^{a_1}(x_1), \gamma_{b_2}^{a_2}(x_2), \dots, \gamma_{b_t}^{a_t}(x_t),$$

onde $a_i > 0$ e $b_i \in \{1, 2, \dots\}$ para $1 \leq i \leq t$, então a potência α do *Downside Risk* da variável aleatória $X = (X_1 \cdot \dots \cdot X_t)^{1/t}$ é dado por

$$DR_{\alpha, R_B}^{\alpha} [X] = \frac{\Gamma(1+\alpha)}{\Gamma(b_1) \dots \Gamma(b_t)} \left(\frac{R_B}{t} \right)^{\alpha} G_{t, 2t}^{t, t} \left(a_1 \dots a_t R_B^t \left| \begin{array}{c} \frac{1}{t}, \frac{2}{t}, \dots, \frac{t}{t} \\ b_1, \dots, b_t; \frac{-\alpha}{t}, \frac{1-\alpha}{t}, \dots, \frac{t-1-\alpha}{t} \end{array} \right. \right). \quad (3.105)$$

Prova. Como

$$(X_1 \cdot \dots \cdot X_t) \sim (\gamma_{b_1}^{a_1} \otimes \dots \otimes \gamma_{b_t}^{a_t})(x) = \frac{a_1 \dots a_t}{\Gamma(b_1) \dots \Gamma(b_t)} G_{0,t}^{t,0} \left(a_1 \dots a_t x \left| \begin{array}{c} - \\ b_1-1 \dots b_t-1 \end{array} \right. \right)$$

e

$$X^{1/t} \sim f^{\otimes \frac{1}{t}}(x) = tx^{t-1} f(x^t)$$

então

$$X = (X_1 \cdot \dots \cdot X_t)^{1/t} \sim tx^{t-1} (\gamma_{b_1}^{a_1} \otimes \dots \otimes \gamma_{b_t}^{a_t})(x^t) = \frac{a_1 \dots a_t}{\Gamma(b_1) \dots \Gamma(b_t)} tx^{t-1} G_{0,t}^{t,0} \left(a_1 \dots a_t x^t \middle| b_1 - 1 \dots b_t - 1 \right).$$

Aplicando (3.88) e (3.81), obtém-se

$$\begin{aligned} DR_{\alpha, R_B}^\alpha [X] &= DR_{\alpha, R_B}^\alpha \left[(X_1 \cdot \dots \cdot X_t)^{1/t} \right] = \int_0^{R_B} (R_B - x)^\alpha \frac{a_1 \dots a_t}{\Gamma(b_1) \dots \Gamma(b_t)} tx^{t-1} G_{0,t}^{t,0} \left(a_1 \dots a_t x^t \middle| b_1 - 1 \dots b_t - 1 \right) dx \\ &= \frac{a_1 \dots a_t t}{\Gamma(b_1) \dots \Gamma(b_t)} \frac{\Gamma(\alpha + 1)}{t^{\alpha+1} R_B^{-t-\alpha}} G_{t,2t}^{t,t} \left(a_1 \dots a_t R_B^t \middle| \Delta(t, 1-t); (a_p) \right) \\ &= \frac{a_1 \dots a_t R_B^t \Gamma(1+\alpha)}{\Gamma(b_1) \dots \Gamma(b_t)} \left(\frac{R_B}{t} \right)^\alpha G_{t,2t}^{t,t} \left(a_1 \dots a_t R_B^t \middle| \frac{1-t}{t}, \frac{2-t}{t}, \dots, \frac{0}{t} \right) \\ &= \frac{\Gamma(1+\alpha)}{\Gamma(b_1) \dots \Gamma(b_t)} \left(\frac{R_B}{t} \right)^\alpha G_{t,2t}^{t,t} \left(a_1 \dots a_t R_B^t \middle| b_1 - 1, \dots, b_t - 1; \frac{-t-\alpha}{t}, \frac{1-t-\alpha}{t}, \dots, \frac{t-1-t-\alpha}{t} \right) \\ &= \frac{\Gamma(1+\alpha)}{\Gamma(b_1) \dots \Gamma(b_t)} \left(\frac{R_B}{t} \right)^\alpha G_{t,2t}^{t,t} \left(a_1 \dots a_t R_B^t \middle| \frac{1}{t}, \frac{2}{t}, \dots, \frac{t}{t} \right). \end{aligned}$$

□

3.16. Lemas auxiliares

Nesta seção, serão apresentados alguns lemas para uso posterior.

Lema 3.16.1. Convexidade e concavidade com a composição de funções. Sejam $g: \mathbb{D}_1 \rightarrow \mathbb{D}_2$ e $h: \mathbb{D}_2 \rightarrow \mathbb{D}_3$ funções com primeira e segunda derivadas e $f = h \circ g: \mathbb{D}_1 \rightarrow \mathbb{D}_3$.

Se h é convexa e não-decrescente e g é convexa, então f é convexa. (3.106)

Se h é convexa e não-crescente e g é côncava, então f é convexa. (3.107)

Se h é côncava e não-decrescente e g é côncava, então f é côncava. (3.108)

Se h é côncava e não-crescente e g é convexa, então f é côncava. (3.109)

Prova. Basta verificar o sinal da segunda derivada de $f(x) = h(g(x))$; veja Boyd e Vandenberghe (2004, p.83).

□

Lema 3.16.2. Desigualdades de Minkowski. Se $a_i \geq 0$ e $b_i \geq 0$ para todos os i , $1 \leq i \leq n$, então

$$\left[(a_1 + b_1)^\alpha + \dots + (a_n + b_n)^\alpha \right]^{1/\alpha} \leq \left[a_1^\alpha + \dots + a_n^\alpha \right]^{1/\alpha} + \left[b_1^\alpha + \dots + b_n^\alpha \right]^{1/\alpha}, \alpha \geq 1 \quad (3.110)$$

e

$$\left[(a_1 + b_1)^\alpha + \dots + (a_n + b_n)^\alpha \right]^{\frac{1}{\alpha}} \geq \left[a_1^\alpha + \dots + a_n^\alpha \right]^{\frac{1}{\alpha}} + \left[b_1^\alpha + \dots + b_n^\alpha \right]^{\frac{1}{\alpha}}, \quad 0 \leq \alpha \leq 1. \quad (3.111)$$

Prova. Veja Widder (1989, p.315) e Hardy, Littlewood e Pólya (1952, p.32).

□

Lema 3.16.3. Generalização de um limite fundamental. Seja $f : \mathbb{D}_1 \rightarrow \mathbb{R}$ uma função contínua com a primeira derivada contínua na vizinhança de $x = 1$, e tal que

$$\lim_{x \rightarrow 1} f(x) = 1;$$

$k \in \mathbb{R}$; e $g : \mathbb{D}_2 \rightarrow \mathbb{R}$ e $h : \mathbb{D}_3 \rightarrow \mathbb{R}$ funções tais que

$$\lim_{t \rightarrow \infty} g(t) = \lim_{t \rightarrow \infty} h(t) = \infty \text{ e } \lim_{t \rightarrow \infty} \frac{g(t)}{h(t)} = 1.$$

Então

$$\lim_{t \rightarrow \infty} \left(f \left(1 + \frac{k}{g(t)} \right) \right)^{h(t)} = \exp[kf'(1)]. \quad (3.112)$$

Prova. Considerando-se que

$$\lim_{x \rightarrow 1} \ln(f(x)) = 0 \text{ e que } \lim_{t \rightarrow \infty} \frac{g(t)}{h(t)} = \lim_{t \rightarrow \infty} \frac{\frac{dg(t)}{dt}}{\frac{dh(t)}{dt}} = 1,$$

então, aplicando L'Hospital e a regra da cadeia, obtém-se

$$\begin{aligned} \lim_{t \rightarrow \infty} \left(f \left(1 + \frac{k}{g(t)} \right) \right)^{h(t)} &= \exp \left[\lim_{t \rightarrow \infty} \frac{\ln \left(f \left(1 + \frac{k}{g(t)} \right) \right)}{\frac{1}{h(t)}} \right] = \exp \left[\lim_{t \rightarrow \infty} \frac{\frac{d}{dt} f \left(1 + \frac{k}{g(t)} \right)}{f \left(1 + \frac{k}{g(t)} \right) \left(-h(t)^{-2} \right) \frac{dh(t)}{dt}} \right] = \\ &= \exp \left[\lim_{t \rightarrow \infty} \frac{f' \left(1 + \frac{k}{g(t)} \right) k \left(-g(t)^{-2} \right) \frac{dg(t)}{dt}}{f \left(1 + \frac{k}{g(t)} \right) \left(-h(t)^{-2} \right) \frac{dh(t)}{dt}} \right] = \exp \left[\frac{kf'(1)}{f(1)} \right] = \exp[kf'(1)]. \end{aligned}$$

□

Exemplo 3.16.1. Limite fundamental. Para o caso particular $f(x) = x$ e $g(t) = h(t) = t$, o lema anterior torna-se o bem conhecido limite

$$\lim_{t \rightarrow \infty} \left(1 + \frac{k}{t} \right)^t = e^k. \quad (3.113)$$

■

3.17. Conclusões

O produto tensorial finito possibilita a representação ou aproximação, com precisão arbitrária, da densidade multivariada de variáveis aleatórias dependentes sobre um domínio fechado, pois o número de parcelas pode crescer arbitrariamente. A densidade multivariada pode ser multimodal. Se houver apenas uma variável aleatória, então o produto tensorial se reduz a uma mistura finita de densidades. Dado um conjunto de amostras, os parâmetros da mistura ou do produto tensorial devem ser calibrados com um critério de ajuste. O escolhido para ser usado neste trabalho foi o da máxima verossimilhança.

O oposto da verossimilhança é uma função convexa sobre os pesos da mistura e do produto tensorial e é, em geral, não-convexa sobre os parâmetros das distribuições componentes. Por isso, é desejável usar o menor número de parâmetros possível. A quantidade de pesos, que são variáveis de otimização deste problema, cresce muito rapidamente com o número de variáveis aleatórias. Este efeito é conhecido como “maldição da dimensionalidade”. Mas a maioria dos pesos serão iguais a zero porque o problema tem uma estrutura especial. Por isso, é necessário um algoritmo feito sob medida para tirar proveito desta propriedade.

Embora o número de amostras seja muito menor que o número de variáveis de otimização, os exemplos indicam que a qualidade da aproximação é muito boa.

A geração de pontos aleatórios de variáveis aleatórias dependentes é muito simples se a densidade multivariada for representada por um produto tensorial finito de distribuições arbitrárias.

As transformadas de Laplace e de Mellin são úteis para se determinar, respectivamente, a densidade da soma e do produto de variáveis aleatórias independentes. Isto é feito com os teoremas da convolução aditiva e multiplicativa.

A função Fox H é a função especial mais genérica conhecida atualmente. As funções Meijer G e Erlang são alguns dos seus casos particulares. Uma função Fox H também pode ser representada por uma função Meijer G, sob hipóteses fracas.

A mistura finita e o produto tensorial finito de distribuições Erlang possuem propriedades muito convenientes para a teoria de carteiras. A convolução aditiva de misturas finitas de Erlangs gera uma mistura finita de Erlangs. Entretanto, os coeficientes podem ser muito grandes em módulo, e a subtração de resultados intermediários aproximadamente iguais pode causar perda de precisão numérica durante os cálculos. A combinação linear positiva de variáveis aleatórias dependentes que seguem conjuntamente um produto tensorial finito de Erlangs gera uma variável aleatória que segue uma mistura finita de Erlangs. A convolução multiplicativa de misturas finitas de Erlangs gera uma mistura finita de um caso particular da distribuição Meijer G. Considerando-se que a Erlang é um caso especial dessa distribuição, então esta classe é fechada em relação à convolução multiplicativa. A distribuição Fox H é fechada em relação à potenciação.

O cálculo da esperança de uma mistura finita de Erlangs é simples. Mas o cálculo do *Downside Risk* requer a avaliação numérica de um caso particular da função Meijer G, que é o caso logarítmico com pólos no numerador e no denominador separados. Os lemas necessários foram demonstrados e foi feita uma implementação no Mathematica. Para alguns casos particulares do *Downside Risk*, a função Meijer G se

reduz a um somatório envolvendo funções elementares.

Foram enunciados três lemas que serão usados posteriormente. Eles tratam da preservação da convexidade e da concavidade na composição de funções, das desigualdades de Minkowski, e de uma generalização de um limite fundamental.

4. RISCO

4.1. Introdução

Risco é um conceito presente em diferentes áreas da ciência e que possui várias operacionalizações ou medidas ou definições matemáticas. Os conceitos relacionados usados para explicar o que é risco incluem variação, dispersão, perda, incerteza e informação. Em Finanças, a medida mais conhecida é a variância ou a sua forma na escala original, o desvio-padrão. Markowitz usou a variância em seu artigo clássico (1952) sobre seleção de carteiras. Um investidor racional, por definição, prefere menos risco. No **Capítulo 2**, foram mencionadas algumas medidas que os autores definiram ao longo do tempo. Neste capítulo, serão apresentados alguns tópicos mais importantes para este trabalho.

4.2. Problemas de medidas de risco com distribuições genéricas

Neste trabalho, serão usadas misturas finitas de Erlangs para representar as densidades das carteiras. Estas densidades são assimétricas e podem ser multimodais. Para esta distribuição, e para as não-gaussianas e assimétricas em geral, algumas medidas de risco muito conhecidas podem causar problemas em uma análise retorno *versus* risco.

As medidas de risco calculadas com base na dispersão dos retornos e as calculadas sobre o domínio de um dado quantil podem levar um investidor a tomar decisões financeiras incorretas do ponto de vista de um investidor racional. Estes problemas serão mostrados a seguir com exemplos simples.

O problema com medidas calculadas com base na dispersão dos retornos, como, por exemplo, a variância, é que elas penalizam os retornos altos ou “desejáveis” de distribuições assimétricas. Por exemplo, na **Figura 4.2.1**, uma carteira com densidade assimétrica foi escolhida e teve seu gráfico rebatido ao redor da sua média para representar uma segunda carteira. As duas densidades têm a mesma média e a mesma variância, porém uma tem cauda longa para as perdas e a outra tem cauda longa para os ganhos. Um investidor que usar a variância (ou desvio-padrão) como medida de risco será indiferente a essas duas carteiras, mas um investidor racional claramente escolherá a segunda. Uma situação ainda pior ocorreria se as duas carteiras tivessem a mesma média e a com cauda longa para a direita tivesse uma variância maior.

As medidas de risco baseadas em quantil ou percentil calculam o risco considerando somente a parte da cauda esquerda da densidade cuja área é igual ao quantil α especificado. O problema com medidas calculadas sobre o domínio de um dado quantil, como o *Conditional Value-at-Risk*, é que elas ou

subestimam ou superestimam os retornos “ruins” ou “indesejáveis”. Na **Figura 4.2.2**, foram traçadas as densidades de duas carteiras com mesmo retorno esperado e mesmo risco para um dado quantil (por exemplo, $\alpha = 10\%$). A primeira carteira possui uma probabilidade de ocorrência de retornos “ruins” (taxas de retorno negativas, à direita da abscissa do quantil) que não foi contabilizada pela medida de risco, enquanto a segunda não possui, isto é, todos os retornos “ruins” foram considerados. Um investidor que usasse esta medida seria indiferente às duas carteiras, mas um investidor racional claramente iria preferir a segunda. Veja também Yamai e Yoshida (2002).

Figura 4.2.1. Problema da variância como medida de risco. As densidades 1) e 2) têm a mesma média e a mesma variância. Mas um investidor racional escolherá a carteira 2), pois ela tem possibilidade de retornos altos e não tem possibilidade de perda.

Figura 4.2.2. Problema de medidas de risco baseadas em quantis. As densidades de duas carteiras têm a mesma área α na região cinza, a mesma média e o mesmo risco baseado em quantil. Mas um investidor

racional escolherá a carteira 2) porque parte dos retornos “ruins” da carteira 1) não foram contabilizados pela medida de risco.

A entropia como medida de risco possui o mesmo problema da variância, isto é, os retornos altos (com probabilidade baixa) são penalizados.

A probabilidade de perda, que é a área sob a função densidade com taxas de retorno negativas, possui o problema de considerar as perdas pequenas e as grandes igualmente importantes. Por exemplo, para esta medida, uma Delta de Dirac teria a mesma importância nos pontos de retorno 0.01 e 0.99, o que não faz sentido financeiro.

4.3. Medidas coerentes de risco

Na seção anterior, foi visto que algumas definições matemáticas para o conceito de risco financeiro podem levar a decisões financeiras incorretas, ou seja, não têm sentido financeiro. É importante identificar as propriedades desejáveis nas medidas de risco para evitar possíveis “incoerências”.

Artzner, Delbaen, Eber e Heath (1999) propuseram as seguintes propriedades (ou axiomas) para uma medida de risco $\mathcal{R}$ ser definida como “coerente”.

Definição 4.3.1. Medida coerente de risco. *Dados o espaço de variáveis aleatórias $L(\Omega, \mathcal{A}, P)$ e um ativo com retorno livre de risco R_f , uma medida de risco $\mathcal{R}[X]$ é dita coerente se possui as propriedades a seguir.*

1. *Subaditividade:* Para todos os X e Y , $\mathcal{R}[X + Y] \leq \mathcal{R}[X] + \mathcal{R}[Y]$.
2. *Homogeneidade positiva:* Para todo X e $\lambda \geq 0$, $\mathcal{R}[\lambda X] = \lambda \mathcal{R}[X]$.
3. *Invariância à translação:* Para todo X e $\alpha \in \mathbb{R}$, $\mathcal{R}[X + \alpha R_f] = \mathcal{R}[X] - \alpha$.
4. *Monotonicidade:* Para todos os X e Y tais que $X \leq Y$, $\mathcal{R}[X] \geq \mathcal{R}[Y]$.

A propriedade da subaditividade vem da diversificação das carteiras; uma fusão não cria risco. Pela propriedade da homogeneidade positiva, se a variável aleatória X for multiplicada por um escalar não-negativo, então o risco também é multiplicado. A invariância à translação torna o risco do ativo livre de risco negativo, isto é, ele não tem risco. Pela propriedade da monotonicidade, se uma carteira X tem sempre retornos menores do que Y , então o risco de X deve ser maior que o de Y .

Vários autores propuseram um conjunto de propriedades desejáveis para medidas de risco financeiro, tais como Kijima e Ohnishi (1993), Pedersen e Satchell (1998), Pflug (2002), Rockafellar, Uryasev e Zabarankin (2003). Veja também Albrecht (2004).

Pedersen e Satchell (1998) deram exemplos de várias medidas de risco definidas ao longo do tempo, e separaram aquelas que satisfazem as propriedades desejáveis das que não satisfazem.

Schulmerich e Trautmann (2003) disseram que a medida *Expected Shortfall* (também conhecida por *Conditional Value-at-Risk*) não satisfaz a invariância à translação e propuseram uma modificação nessa propriedade.

As propriedades a seguir parecem ser importantes para uma medida de risco calculada para uma variável aleatória do tipo retorno.

1. O domínio deve ser o dos retornos “indesejáveis”, isto é, de $-\infty$ até um retorno padrão R_B , para evitar a sub e superestimação dos retornos indesejáveis.
2. A medida deve assumir um valor não-negativo em cada ponto do domínio, de forma que o conjunto imagem será o dos reais não-negativos, e permitirá a potenciação do valor calculado.
3. A medida deve resultar em valores na mesma escala da esperança, pois, caso contrário, as propriedades da convexidade e da dominância estocástica podem deixar de ser válidas nessa escala.
4. A medida deve ser convexa sobre os pesos não-negativos dos ativos nas carteiras, porque a convexidade representa o princípio de que a diversificação não deve aumentar o risco.
5. O conceito de dominância estocástica deve ser satisfeito pela medida.

Um maior detalhamento dessas propriedades está fora do escopo deste trabalho. É possível ajustar a densidade dos retornos (isto será feito no **Capítulo 6**) antes de se aplicar a medida de risco.

4.4. *Downside Risk*

A medida *Downside Risk* foi a escolhida para ser usada neste trabalho. Ela teve seus primórdios com a semivariância, definida pelo próprio Markowitz, em 1959, e evoluiu ao longo do tempo. Veja Markowitz (1991a, p.374).

Markowitz (1991a, p.188) definiu semivariância para amostras discretas como a soma dos quadrados dos retornos abaixo de um retorno padrão menos o retorno padrão, sobre o número total de amostras. Bawa (1975) definiu a medida *Lower Partial Variance*. Bawa e Lindenberg (1977) definiram o *Lower Partial Moment*. Fishburn (1977) definiu a “função dos desvios abaixo de um retorno alvo especificado ponderada pela probabilidade” ou modelo alfa-t. Para alfa=2, ele chamou de “variância abaixo do retorno alvo”. Stone (1973) definiu uma classe de medidas de risco bem ampla que inclui as citadas. Estas medidas serão definidas a seguir, para $X \sim p(x)$.

Semivariância (Markowitz, 1959).

$$S_b[X] = \frac{1}{n} \sum_{i=1}^n (\max\{b - x_i, 0\})^2 \quad (4.1)$$

Medida de risco com três parâmetros (Stone, 1973).

$$L_{x_0,k,A}[X] = \int_{-\infty}^A |x - x_0|^k dP(x), \quad k \geq 0 \quad (4.2)$$

$$R_{x_0,k,A}[X] = (L_{x_0,k,A}[X])^{1/k} = \left(\int_{-\infty}^A |x - x_0|^k dP(x) \right)^{1/k}, \quad k \geq 0 \quad (4.3)$$

Lower Partial Variance (Bawa, 1975).

$$LPV_t[X] = \int_a^t (t-x)^2 dP(x), \quad a = \inf \{x\} \quad (4.4)$$

Lower Partial Moment (Bawa e Lindenberg, 1977).

$$LPM_{n,t}[X] = \int_a^t (t-x)^n dP(x), \quad a = \inf \{x\}, \quad n = 0, 1, 2, \dots \quad (4.5)$$

Modelo Alfa-t (Fishburn, 1977).

$$F_{\alpha,t}[X] = \int_{-\infty}^t (t-x)^\alpha dP(x), \quad \alpha > 0 \quad (4.6)$$

Em alguns artigos mais recentes, os autores definem o *Lower Partial Moment* considerando o parâmetro n como um número real positivo, ou seja, usam a mesma definição do modelo Alfa-t. Esta definição também será usada neste trabalho. Veja também Nawrocki (1999), que fez uma pesquisa histórica e Pedersen e Satchell (1998), que apresentaram várias medidas de risco.

A medida de risco *Downside Risk* será definida a seguir. Ela é a potência $1/\alpha$ do modelo Alfa-t. Embora seja uma simples mudança de escala, para casos particulares de seus parâmetros, ela torna linear a fronteira eficiente com um ativo livre de risco (separação em um fundo); veja Bawa e Lindenberg (1977). Esta é uma propriedade desejável, pois os cálculos são pesados computacionalmente. Esta diferença é a mesma que ocorre para as fronteiras eficientes calculadas com o desvio-padrão ou com a variância.

Definição 4.4.1. Downside Risk. Dado um retorno de referência $R_B > 0$ e um coeficiente de aversão a retornos indesejáveis $\alpha > 0$, o *Downside Risk* de uma variável aleatória R (do tipo retorno) que segue uma distribuição $F(R)$ é definido como

$$DR_{\alpha,R_B}[R] = \left(\int_{-\infty}^{R_B} (R_B - R)^\alpha dF(R) \right)^{1/\alpha}. \quad (4.7)$$

O termo $(R_B - R)$ é sempre não-negativo. Se a variável aleatória for do tipo riqueza, basta fazer uma transformação de variáveis. O limite inferior da integral é $-\infty$ para permitir carteiras alavancadas. Se este não for o caso, então ele pode ser substituído por 0 ou pelo limite inferior do domínio da variável aleatória. O limite superior da integral é R_B , mas ele pode ser interpretado (ou definido) como R_{B-} porque uma possível distribuição Delta de Dirac nesse ponto adicionaria à integral sua massa (um número entre 0 e 1) vezes 0. A notação $dF(R)$ é a usada pela integral de Lebesgue-Stieltjes.

O parâmetro R_B foi chamado de “retorno alvo” por Fishburn (1977) e de “mínimo retorno aceitável” por Sortino (2001). Sortino argumentou que o termo “retorno alvo” induz o investidor a especificar um retorno muito alto, isto é, o retorno “desejado” por ele e, por causa desta confusão, o termo “mínimo retorno

aceitável” é preferível. Os termos “máximo retorno ruim” ou “maior retorno indesejado” descrevem R_B de forma um pouco mais precisa. Note que R_B pode ser uma variável, tal como $R_B = E[R]$.

A medida *Downside Risk* tem sentido financeiro para investidores racionais somente se $\alpha \geq 1$ e $1 \leq 1 + r_f = R_f \leq R_B$, onde r_f é a taxa de retorno do ativo “livre de risco” (com variância nula). Se $0 < \alpha < 1$, então a medida de risco não é convexa e, portanto, investidores com aversão ao risco evitarão estes valores. A condição de que $R_f \leq R_B$ vem da consideração de que retornos abaixo de R_f devem ser penalizados, pois mesmo que não haja perda absoluta, haverá a perda de oportunidade de se investir no ativo livre de risco. Alguns investidores podem querer fazer $R_B = 1$ porque o conceito de perda absoluta é mais facilmente entendido, porque retornos acima de 1 são contabilizados pela esperança, e porque facilita os cálculos. Esta questão se reduz a uma discussão filosófica sobre o que é um “retorno indesejável”.

O *Downside Risk* sempre será um escalar não-negativo e, portanto, não satisfaz a propriedade da invariância à translação, segundo a qual uma medida de risco pode ter valores negativos. Também possui importantes relações com os conceitos de dominância estocástica, utilidade, finanças comportamentais, separação em um e dois fundos e CAPM. Estes tópicos não serão abordados neste trabalho.

O *Downside Risk* pode ter também as notações a seguir.

$$\begin{aligned} DR_{\alpha, R_B}[R] &= (LPM_{\alpha, R_B}[R])^{1/\alpha} = LPM_{\alpha, R_B}^{1/\alpha}[R] = \\ &= \left(\int_{-\infty}^{\infty} (\max\{R_B - R, 0\})^\alpha dF(R) \right)^{1/\alpha} = \left(\int_{-\infty}^{\infty} (R_B - R)_+^\alpha dF(R) \right)^{1/\alpha} \end{aligned} \quad (4.8)$$

4.4.1. Convexidade

Nesta seção, será provado que o *Downside Risk* e o *Lower Partial Moment* são medidas de risco convexas para casos particulares de α , R_B e pesos dos ativos. A convexidade é a principal propriedade das medidas de risco porque representa o princípio de que a diversificação não deve aumentar o risco. A convexidade também é muito desejável em problemas de minimização.

Embora o *Lower Partial Moment* e o *Downside Risk* sejam funções relacionadas, a função $f(x) = x^{1/\alpha}$, para $\alpha > 1$, é côncava e não-decrescente; por isso, a convexidade de $LPM_{\alpha, R_B}[X]$ não implica diretamente que $LPM_{\alpha, R_B}^{1/\alpha}[X] = DR_{\alpha, R_B}[X]$ seja convexa. É necessário deixar claro quais os valores de α e R_B que estão sendo usados, pois as provas dos teoremas são diferentes. Há uma grande confusão na literatura por causa destes detalhes.

Bawa (1978, p.267, teorema 4) disse que $LPM_{\alpha, R_B}[X]$ é uma função convexa para $\alpha = 1, 2, \dots$ e R_B fixo. Ele apresentou uma prova, mas com pouco rigor matemático; veja também Hogan e Warren (1972). Grootveld e Hallerbach (1999, p.309) citaram o teorema de Bawa de forma incorreta. Pedersen e Satchell (1998, p.114) apresentaram uma prova da subaditividade de $DR_{\alpha, E[X]}[X]$, para $\alpha > 1$, com uma inserção incorreta de módulos.

Bawa e Lindenberg (1977, p.194) disseram que a convexidade de $LPM_{2, R_f}[X]$ não é suficiente para garantir a convexidade de $DR_{2, R_f}[X] = LPM_{2, R_f}^{1/2}[X]$ e que, para isso, é necessário que $\ln(LPM_{2, R_f}[X])$

também seja convexa. Também disseram que não foram capazes de prover condições simples para distribuições em geral. Pedersen e Satchell (2002, p.222) provaram que para um valor fixo $R_B < R_f$, a função $DR_{2,R_B}[X]$ é convexa no caso em que há um ativo livre de risco na carteira e todos os pesos são positivos; e côncava no caso em que o peso do ativo livre de risco é negativo. Veja também Brogan e Stidham (c2003a).

A seguir, será provada a propriedade da convexidade das medidas *Downside Risk* e *Lower Partial Moment*.

Teorema 4.4.1. Convexidade do *Downside Risk*. Se $\alpha \geq 1$, R_B é uma constante, $0 \leq u \leq 1$ e $v = 1 - u$, então $DR_{\alpha,R_B}[R]$ é uma função convexa sobre os pesos dos ativos, isto é,

$$DR_{\alpha,R_B}[uX + vY] \leq uDR_{\alpha,R_B}[X] + vDR_{\alpha,R_B}[Y].$$

Prova. Suponha que $[X, Y]^T \sim m(x, y)$. No gráfico bivariado da figura a seguir, a densidade $m(x, y)$ deve ser imaginada na terceira dimensão, para cima do papel. O *Downside Risk* de X é calculado sobre a área à esquerda da reta vertical $R_x = R_B$, o *Downside Risk* de Y é calculado sobre a área abaixo da reta horizontal $R_y = R_B$ e o *Downside Risk* de $Z = uX + vY$ é calculado sobre a área à esquerda e abaixo da reta $uR_x + vR_y = R_B$.

Figura 4.4.1. Convexidade do *Downside Risk*.

Há cinco regiões de interesse (4 octantes e 1 quadrante) e um pequeno paralelepípedo foi desenhado em cada uma delas. A altura h_i de cada paralelepípedo i , $1 \leq i \leq 5$, representa a probabilidade de as variáveis aleatórias assumirem valores dentro de sua base, dividida pela área da base. Dependendo da região, cada paralelepípedo pode contribuir ou não para $DR[uX + vY]$, $DR[X]$ e $DR[Y]$.

Nas linhas L_1 a L_5 da tabela seguinte, é mostrado o quanto a altura h_i de cada paralelepípedo contribui para $(DR[uX + vY])^\alpha$, $(DR[X])^\alpha$ e $(DR[Y])^\alpha$. Somando uma quantidade não-negativa a cada valor da coluna C_1 , obtém-se, na coluna C_2 , um valor majorado para $(DR[uX + vY])^\alpha$. A igualdade ocorrerá se $h_i = 0$.

A linha L_5 contém as somas das colunas. Nesta linha, a coluna C_3 foi multiplicada por u^α , e a coluna C_4 , por v^α . Elevando as somas à potência $1/\alpha$ e aplicando a desigualdade de Minkowski (3.110), a linha L_M é obtida. Particionando as cinco regiões de interesse com infinitos paralelepípedos de modo que as áreas de suas bases tendam a zero (veja Widder, 1989, p.198) e repetindo o raciocínio, a linha L_C é obtida.

Tabela 4.4.1. Convexidade do *Downside Risk*.

	C_1		C_2		C_3		C_4
	$(DR[uX + vY])_i^\alpha$		$(DR[uX + vY])_i^\alpha$ majorado		$a_i^\alpha = (DR[X])_i^\alpha$		$b_i^\alpha = (DR[Y])_i^\alpha$
L_1	$[R_B - (uR_{1x} + vR_{1y})]^\alpha h_1$	=	$[u(R_B - R_{1x}) + v(R_B - R_{1y})]^\alpha h_1^{\frac{\alpha}{\alpha}}$		$(R_B - R_{1x})^\alpha h_1^{\frac{\alpha}{\alpha}}$		$(R_B - R_{1y})^\alpha h_1^{\frac{\alpha}{\alpha}}$
L_2	$[R_B - (uR_{2x} + vR_{2y})]^\alpha h_2$	≤	$[u(R_B - R_{2x}) + v0]^\alpha h_2^{\frac{\alpha}{\alpha}}$		$(R_B - R_{2x})^\alpha h_2^{\frac{\alpha}{\alpha}}$		$0^\alpha h_2^{\frac{\alpha}{\alpha}}$
L_3	$[R_B - (uR_{3x} + vR_{3y})]^\alpha h_3$	≤	$[u0 + v(R_B - R_{3y})]^\alpha h_3^{\frac{\alpha}{\alpha}}$		$0^\alpha h_3^{\frac{\alpha}{\alpha}}$		$(R_B - R_{3y})^\alpha h_3^{\frac{\alpha}{\alpha}}$
L_4	0	≤	$[u(R_B - R_{4x}) + v0]^\alpha h_4^{\frac{\alpha}{\alpha}}$		$(R_B - R_{4x})^\alpha h_4^{\frac{\alpha}{\alpha}}$		$0^\alpha h_4^{\frac{\alpha}{\alpha}}$
L_5	0	≤	$[u0 + v(R_B - R_{5y})]^\alpha h_5^{\frac{\alpha}{\alpha}}$		$0^\alpha h_5^{\frac{\alpha}{\alpha}}$		$(R_B - R_{5y})^\alpha h_5^{\frac{\alpha}{\alpha}}$
L_S	$\sum_i (DR[uX + vY])_i^\alpha$	≤	$\sum_i (ua_i + vb_i)^\alpha$		$\sum_i (ua_i)^\alpha$		$\sum_i (vb_i)^\alpha$
L_M	$\left(\sum_i (DR[uX + vY])_i^\alpha \right)^{\frac{1}{\alpha}}$	≤	$\left(\sum_i (ua_i + vb_i)^\alpha \right)^{\frac{1}{\alpha}}$	≤	$\left(\sum_i (ua_i)^\alpha \right)^{\frac{1}{\alpha}}$	+	$\left(\sum_i (vb_i)^\alpha \right)^{\frac{1}{\alpha}}$
L_C	$DR[uX + vY]$	≤		≤	$uDR[X]$	+	$vDR[Y]$

□

Note que o teorema anterior vale para $\alpha \geq 1$, R_B fixo e pesos não-negativos. A seguir, será provada a convexidade do *Lower Partial Moment*.

Corolário 4.4.1. Convexidade do *Lower Partial Moment*. Se $\alpha \geq 1$, R_B é uma constante, $0 \leq u \leq 1$ e $v = 1 - u$, então $LPM_{\alpha, R_B}[R]$ é uma função convexa sobre os pesos dos ativos, isto é,

$$LPM_{\alpha, R_B}[uX + vY] \leq uLPM_{\alpha, R_B}[X] + vLPM_{\alpha, R_B}[Y].$$

Prova. Como a função $r(x) = x^\alpha$, para $\alpha \geq 1$, é convexa e não-decrescente e $DR_{\alpha, R_B}[R]$ é uma função convexa para as condições dadas, então, pela propriedade (3.106), $LPM_{\alpha, R_B}[R] = r(DR_{\alpha, R_B}[R]) = DR_{\alpha, R_B}^\alpha[R]$ é uma função convexa.

□

4.5. Falência

A possibilidade da ocorrência de falência ou ruína pode afetar as decisões financeiras dos investidores.

Yamai e Yoshida (2001) deram um exemplo de uma empresa que vende seguros contra terremotos

em Tóquio e em Los Angeles. A probabilidade de acontecer um desastre é de 1% por período em cada cidade, e os eventos são independentes. A empresa tem ativos de 100 milhões de ienes e patrimônio líquido de 49 milhões de ienes. A empresa tem duas escolhas: vender uma apólice de 100 milhões de ienes para a ocorrência de um terremoto em Tóquio; ou vender uma apólice de 50 milhões de ienes para a ocorrência de um terremoto em Tóquio e outra de 50 milhões de ienes para a ocorrência de um terremoto em Los Angeles. Se o aspecto relevante de risco é a falência da empresa, então a empresa decidirá pela primeira alternativa, isto é, não diversificar os investimentos, porque a probabilidade de falência é 1% no primeiro caso e 2% no segundo. Claro que este nível de alavancagem não deveria ser permitido. Yamai e Yoshiba (2001) disseram que a subaditividade pode não ser relevante se a perda for maior que o patrimônio líquido da empresa. Na realidade, a explicação financeira é um pouco mais sutil.

No mundo real, as carteiras alavancadas são permitidas e, em geral, os investidores tais como bancos podem ter mais de uma carteira. Se uma carteira tiver um retorno de -1% (isto é, uma taxa de retorno de -101%), as demais carteiras provavelmente iriam ser suficientes para cobrir a diferença. Assumindo que o investidor aplicou toda sua riqueza e considerando-se uma carteira que contém todas as demais, um retorno de -1% significa a sua falência. Atribuir um risco ∞ ou utilidade $-\infty$ aos retornos menores ou iguais a zero (taxas de retorno menores ou iguais a -100%) faz a medida de risco ou a utilidade divergirem para todos os investimentos que possam causar falência, isto é, todas as carteiras possuindo retornos menores ou iguais a zero teriam risco ∞ ou utilidade $-\infty$. Atribuir um risco finito ou utilidade finita aos retornos menores ou iguais a zero causaria apenas uma mudança na medida de risco ou na curva de utilidade, de forma que seria possível haver duas carteiras com mesmo retorno esperado e mesmo risco, ou mesma utilidade, mas com diferentes probabilidades de causar a falência do investidor. Deixar de diferenciar essas duas carteiras não faz sentido financeiro, pois todos os investidores racionais escolherão a que tiver menor probabilidade de falência. Por causa da condição de responsabilidade limitada do investidor, não é necessário atribuir pesos, risco ou utilidade aos retornos negativos; a probabilidade acumulada dos retornos de $-\infty$ a 0 é suficiente para uma decisão financeira.

Este exemplo sugere fortemente que a escolha de carteiras capazes de causar a falência do investidor deve ser baseada em três variáveis: retorno esperado (ou recompensa, para ser mais genérico), risco e probabilidade de falência. Essas carteiras incluem as alavancadas (nas quais o peso é negativo para um ativo com possibilidade de retornos altos) e as que possuem ativos com possibilidade de falência (os pesos são positivos, mas existe possibilidade de retornos menores ou iguais a zero). Na abordagem paretiana, uma superfície eficiente será gerada, e o investidor deve fornecer valores aceitáveis para duas das três variáveis. Isto é uma generalização do modelo de Markowitz. Na abordagem da utilidade, a curva deve ser finita para todos os retornos possíveis e deve ser incluída uma restrição indicando a máxima probabilidade de falência tolerada.

Nesta dissertação, por hipótese, os pesos dos ativos são não-negativos, e a probabilidade de retornos menores ou iguais a zero é nula; portanto, o evento falência tem probabilidade zero. No mundo real, a

probabilidade de uma empresa falir e causar a falência de um investidor, mesmo não-alavancado, é maior que zero. O relaxamento destas duas hipóteses não é simples e está fora do escopo deste trabalho.

Zeleny (1981) argumentou que um conceito mais generalizado de risco deve incluir as várias dimensões ou tipos de risco. Entretanto, a forma multidimensional do risco não recebeu muita atenção dos autores, provavelmente porque forçaria o investidor a especificar valores aceitáveis para todos os tipos de risco, e porque uma única medida de risco pode conter os diferentes aspectos de risco por meio de, por exemplo, uma média ponderada.

4.6. Conclusões

Ainda não há um consenso entre os autores sobre qual é a medida de risco mais adequada para o risco financeiro. O reconhecimento de que a medida de risco deve satisfazer algumas propriedades para que haja sentido ou coerência financeira está estimulando a produção de artigos restringindo as possíveis medidas e dando uma maior formalização matemática para esta área.

Diferentes autores citam diferentes propriedades consideradas coerentes para uma medida de risco financeiro. Uma propriedade, para ser considerada “coerente”, deve ser estritamente necessária, isto é, sua falta causará resultados sem sentido financeiro ou situações indesejáveis. A identificação de tais propriedades seria uma contribuição valiosa para Finanças.

As medidas de risco baseadas em quantis (por exemplo, VaR e CVaR) podem causar uma situação financeira incoerente, na qual parte dos retornos indesejáveis de uma carteira não é contabilizada. Isto causa a subestimação do risco. A raiz do problema remete a uma discussão filosófica sobre o que é um retorno indesejável ou arriscado.

A medida de risco *Downside Risk* é uma evolução da semivariância, proposta por Markowitz em 1959. Ela não causa a sub ou superestimação dos retornos indesejáveis. Ela não satisfaz a propriedade da invariância à translação do conjunto de medidas coerentes proposto por Artzner *et al.* (1999). Outros autores não citam esta propriedade. Ainda é um problema de pesquisa se a falta da invariância à translação causa resultados incoerentes.

A medida de risco *Downside Risk* é uma função convexa sobre os pesos dos ativos quando $\alpha \geq 1$, R_B é constante e os pesos são não-negativos. A convexidade do *Lower Partial Moment* decorre diretamente da convexidade do *Downside Risk*. A convexidade é uma propriedade importante em Finanças porque representa o princípio de que a diversificação não deve aumentar o risco, e em problemas de minimização porque geralmente a solução é única sobre domínios convexos.

A falência é um tipo de risco ou evento indesejável, mas não é conveniente incluí-la na medida de risco nem na curva de utilidade por causar a sua divergência ou a sua não-identificação. Por isso, a seleção de carteiras capazes de causar falência deve ser feita em uma abordagem paretiana com três variáveis: recompensa, risco e probabilidade de falência, e o investidor deve fornecer valores aceitáveis para duas delas.

5. OTIMIZAÇÃO UNIPERIÓDICA DE CARTEIRAS

5.1. Introdução

No **Capítulo 3**, foi visto que a densidade multivariada dos retornos dependentes dos ativos pode ser aproximada por um produto tensorial finito de Erlangs. A partir dessa densidade multivariada, é possível se obter a densidade dos retornos de uma carteira dada pelos pesos dos ativos. Esta densidade é univariada e será uma mistura finita de Erlangs. Os pesos dos ativos nas carteiras devem somar 1 e, neste trabalho, por hipótese, são não-negativos.

Uma mistura pode ser assimétrica e multimodal e, portanto, permite uma representação mais acurada da densidade de probabilidade. Em vários artigos é citado o desvio da normalidade das densidades dos retornos dos ativos, notadamente em mercados emergentes. Além disso, derivativos podem tornar assimétrica a densidade da carteira. A multimodalidade pode ocorrer por causa da possibilidade de um evento financeiro incomum no dia-a-dia da empresa. Por exemplo, se uma empresa farmacêutica está pesquisando um medicamento promissor, a probabilidade de um retorno muito alto pode ser maior que a dos retornos intermediários, de forma que a densidade de probabilidade terá duas modas. Por outro lado, a possibilidade de ocorrência de um desastre pode ser representada por uma moda no lado das perdas.

Neste capítulo, serão determinadas as expressões para a densidade da carteira, esperança e *Downside Risk*. Elas serão usadas na formulação do problema de otimização para a obtenção das carteiras da fronteira eficiente.

5.2. Convolução aditiva

Se duas variáveis aleatórias independentes X e Y têm suas funções densidade de probabilidade conhecidas, então é possível determinar a densidade de $X + Y$ aplicando a integral da convolução ou usando a transformada de Laplace juntamente com o teorema da convolução aditiva; veja a **Seção 3.7**. A aplicação da integral possui uma interpretação gráfica que pode ser encontrada em livros didáticos sobre probabilidade.

A transformada de Laplace é uma das transformadas integrais que podem ser usadas com o teorema da convolução aditiva. Para variáveis aleatórias independentes e definidas sobre o domínio dos reais não-negativos, tem-se a relação a seguir.

$$X \sim f(x) \xrightarrow{\mathcal{L}} F(s) = \int_0^{\infty} e^{-sx} f(x) dx$$

$$\Downarrow$$

$$(5.1)$$

$$X_1 + X_2 \sim (f_1 \oplus f_2)(x) = \int_0^x f_1(t) f_2(x-t) dt \xrightarrow{\mathcal{L}} F_1(s) F_2(s), X_1 \text{ indep. } X_2.$$

5.3. Combinação linear positiva

Se existir dependência entre variáveis aleatórias tais como os retornos dos ativos em um dado instante, então a convolução aditiva não pode ser usada; é necessário usar o conceito de combinação linear.

Se os pesos são não-negativos, então a combinação linear é chamada não-negativa. Se os pesos são positivos, então a combinação linear é chamada positiva. Se a soma dos pesos é unitária, então se obtém uma combinação afim. Se os pesos são não-negativos e têm soma unitária, então se obtém uma combinação convexa.

Uma forma de se entender a combinação linear dos retornos dependentes dos ativos é a técnica de “fatiamento e pesagem”. Considere um exemplo no qual uma carteira P tem 2 ativos, na qual $w_1 = w_2 = 0.5$. Qual é a densidade de probabilidade de o retorno da carteira ser $R_P = 1.1$? Uma possível realização seria $R_1 = 1.2$ e $R_2 = 1.0$. Outra seria $R_1 = 1.1$ e $R_2 = 1.1$. E outra seria $R_1 = 1.0$ e $R_2 = 1.2$. Estas três possibilidades estão sobre uma linha no plano R_1 versus R_2 , cuja inclinação é determinada pelos pesos dos ativos na carteira. Qualquer ponto sobre essa linha fará com que a carteira tenha um retorno igual a 1.1. Em geral, cada ponto terá uma densidade de probabilidade diferente. O gráfico $m(R_1, R_2)$ versus $(R_1 \text{ versus } R_2)$ representa a densidade de probabilidade da ocorrência simultânea de R_1 e R_2 . A densidade de probabilidade de a carteira ter um retorno igual a 1.1, isto é, $f(1.1)$, será a soma das densidades de probabilidade sobre a linha, que é a integral de linha com respeito à densidade de probabilidade sobre a linha. O valor da integral é o “peso” da “fatia”. A densidade de probabilidade de o retorno da carteira ser $R_P = 1.3$ será o “peso” de uma “fatia” paralela à primeira. Idealmente, o gráfico deve ser fatiado um número infinito de vezes.

No exemplo com dois ativos, o gráfico de $m(R_1, R_2)$ versus $(R_1 \text{ versus } R_2)$ tem três dimensões e possui a forma aproximada de um sino. Seu volume e sua massa são iguais a 1. Se ele for fatiado 10, 1000 ou infinitas vezes, a soma dos pesos das fatias ainda será igual a 1. O peso de cada fatia pode ser mapeado para uma barra ou um ponto no gráfico bidimensional da função densidade de probabilidade da carteira $f(R_P)$. Se houver um número finito de fatias, então cada fatia terá uma largura e o gráfico será parecido com um histograma. Se houver infinitas fatias, então a curva será contínua, a área sob o gráfico será unitária, e todos os pontos serão não-negativos, caracterizando uma densidade de probabilidade. Este é o conceito de integral de Riemann. Tecnicamente, é necessário um ajuste, que é a multiplicação pelo Jacobiano.

Com três ativos, é possível representar a densidade de probabilidade multivariada $m(R_1, R_2, R_3)$ como um valor em cada ponto no espaço R_1 versus R_2 versus R_3 . Uma forma de visualizar um gráfico em

quatro dimensões seria imaginar, por exemplo, o Sol, onde as temperaturas de seus pontos centrais são maiores que dos pontos da superfície. A fatia será um plano inclinado em relação aos eixos. O peso do fatia será uma integral de superfície com respeito à densidade.

Outra forma de se obter a densidade de probabilidade da carteira é calcular a função densidade de probabilidade acumulada e, então, aplicar a derivada. Esta forma de cálculo é preferível porque não requer o cálculo do Jacobiano.

Se a densidade multivariada de variáveis aleatórias dependentes é representada por um produto tensorial finito de Erlangs, então a densidade da combinação linear positiva gera uma densidade univariada na qual a convolução aditiva de Erlangs aparece, e seu resultado final é uma mistura finita de Erlangs. Veja o **Exemplo 3.11.1**. Se as variáveis aleatórias forem os retornos discretos dos ativos, então a densidade dos retornos discretos de qualquer carteira não-alavancada será uma mistura finita de Erlangs. Esta é uma propriedade matemática bastante desejável. Ela pode ser justificada rapidamente notando-se que a Erlang é uma função formada por um produto de uma exponencial e um polinômio, e que sua derivada e sua integral também têm este formato.

5.4. Fronteira eficiente

O cálculo do risco e do retorno para diferentes carteiras com diferentes pesos $w_1, \dots, w_D$ e sua plotagem em um gráfico retorno *versus* risco faz com que uma região compacta seja gerada. A fronteira superior esquerda do gráfico é um conjunto de pontos com a esperança e o risco das carteiras eficientes. Dado um certo risco, existe um ponto no gráfico que representa a carteira (ou carteiras) com maior retorno esperado possível. Dado um certo retorno esperado, existe um ponto no gráfico que representa a carteira (ou carteiras) com menor risco possível. Portanto, um investidor racional, que deseja mais retorno e menos risco, deve escolher uma carteira sobre a fronteira eficiente.

Há diferentes formas de se definir e calcular o retorno esperado (ou recompensa, para ser mais genérico) e o risco. Neste capítulo, o retorno esperado será calculado da forma usual e o risco será calculado com a medida de risco *Downside Risk*.

Neste trabalho, a densidade uniperiódica das carteiras é representada por uma mistura finita de Erlangs. Embora possa haver um número grande de parcelas, a densidade Erlang tem uma forma simples para o cálculo de sua média; veja no **Corolário 3.14.1** a expressão da esperança de uma mistura finita de Erlangs. É possível calcular a variância de uma mistura de Erlangs, mas já foi visto no **Capítulo 4** que a variância não é uma boa medida de risco.

O cálculo do *Downside Risk* para a mistura finita de Erlangs é bem mais complicado. Para um coeficiente de aversão a risco $\alpha \geq 1$ genérico, é necessário calcular numericamente a função Meijer G; veja o **Corolário 3.15.1**. Entretanto, para os casos particulares $\alpha = 1$ e $\alpha = 2$, a avaliação numérica da Meijer G não é necessária; veja o **Corolário 3.15.2** e o **Corolário 3.15.3**.

A função especial Meijer G foi definida na **Seção 3.8**. A avaliação numérica de um caso particular

desta função, conhecido como caso logarítmico, mas com pólos no numerador e no denominador separados, foi feita na **Seção 3.9**.

A seguir, será feita a formulação do problema de otimização que leva à fronteira eficiente uniperiódica com densidades genéricas. Nas seções seguintes, os cálculos serão mostrados com mais detalhes para a densidade Erlang.

Problema de otimização 5.4.1. Fronteira eficiente uniperiódica. Existem D ativos dependentes. A densidade multivariada de seus retornos $\mathbf{R}_d$, para $1 \leq d \leq D$, é conhecida. Os pesos w_d são as variáveis de otimização, devem somar 1 e ser não-negativos. A medida de risco $\mathcal{R}$ e o retorno esperado R_E que o investidor deseja são dados. A densidade dos retornos da carteira $\mathbf{R}_P$ é expressa por $f(R; w_1, \dots, w_D)$. O problema de otimização se torna

$$\begin{aligned} \min_{w_d} \mathcal{R}[\mathbf{R}_P] &= \mathcal{R} \left[\sum_{d=1}^D w_d \mathbf{R}_d \right] = \mathcal{R}(f(R; w_1, \dots, w_D)) \\ \text{s.t.} \\ E[\mathbf{R}_P] &= E \left[\sum_{d=1}^D w_d \mathbf{R}_d \right] = E(f(R; w_1, \dots, w_D)) \geq R_E \\ w_d &\geq 0, \quad 1 \leq d \leq D \\ \sum_{d=1}^D w_d &= 1. \end{aligned}$$

Na formulação na qual os pesos são não-negativos, existe um valor máximo para o retorno que uma carteira pode atingir, e a fronteira eficiente é limitada.

5.4.1. Distribuição multivariada dos retornos dos ativos

Neste trabalho, a função densidade de probabilidade multivariada dos retornos dos D ativos dependentes $\mathbf{R}_d$, para $1 \leq d \leq D$, é representada por um produto tensorial finito de Erlangs, isto é,

$$[\mathbf{R}_1, \dots, \mathbf{R}_D]^T \sim m(R_1, \dots, R_D) = \sum_{i=1}^n p_i \gamma_{b_{i,1}}^{a_1}(R_1) \dots \gamma_{b_{i,D}}^{a_D}(R_D) \geq 0, \quad \sum_{i=1}^n p_i = 1, \quad 1 \leq n < \infty. \quad (5.2)$$

Para este produto tensorial, n é o número de parcelas, p_i , para $1 \leq i \leq n$, são os pesos das parcelas, $a_d > 0$, para $1 \leq d \leq D$, são os parâmetros de escala, constantes para cada variável aleatória d , $b_{i,d} \in \{1, 2, \dots, b_{\max,d}\}$, para $1 \leq i \leq n$ e $1 \leq d \leq D$, são parâmetros de forma e $b_{\max,d}$, para $1 \leq d \leq D$, são inteiros fornecidos. Os pesos p_i , para $1 \leq i \leq n$, podem ser negativos desde que $m(R_1, \dots, R_D) \geq 0$ no domínio.

Os parâmetros são calibrados com o critério da máxima verossimilhança, usando dados históricos dos retornos dos ativos. Este é um problema de otimização em larga escala, mas com uma estrutura especial. Ele foi formulado no **Problema de otimização 3.5.5**. Nesta formulação, os pesos são forçados a ser não-negativos, isto é, $p_i \geq 0$, para $1 \leq i \leq n$, porque garantir a não-negatividade da densidade com pesos irrestritos causa uma grande dificuldade computacional. É mais simples forçar os pesos a serem não-

negativos e aumentar o número de parcelas do produto tensorial.

Uma hipótese necessária para a utilização do critério da máxima verossimilhança com dados históricos é que a densidade multivariada tenha sido i.i.d. no passado. Uma hipótese necessária para a utilização, em inferências para o futuro, da densidade multivariada calculada é que ela continue i.i.d. no futuro.

5.4.2. Distribuição dos retornos da carteira

Uma carteira R_P é representada pelos pesos dos ativos w_d , para $1 \leq d \leq D$. Os pesos têm soma unitária e são não-negativos por hipótese.

Pelo **Corolário 3.11.1**, a densidade univariada dos retornos da carteira é dada por

$$R_P \sim f(R; w_1, \dots, w_D) = \sum_{i=1}^n p_i \left(\gamma_{b_{i,1}}^{a_1/w_1} \oplus \dots \oplus \gamma_{b_{i,D}}^{a_D/w_D} \right) (R) \quad (5.3)$$

Se os parâmetros de escala coincidem para duas Erlangs, isto é, $a_u/w_u = a_v/w_v$, para $u \neq v$, então a convolução aditiva é dada pelo **Teorema 3.10.2**.

$$\left(\gamma_{b_{i,u}}^{a_u/w_u} \oplus \gamma_{b_{i,v}}^{a_v/w_v} \right) (x) = \left(\gamma_{b_{i,u}}^{a_u/w_u} \oplus \gamma_{b_{i,v}}^{a_u/w_u} \right) (x) = \gamma_{b_{i,u}+b_{i,v}}^{a_u/w_u} (x). \quad (5.4)$$

Se $a_u/w_u \neq a_v/w_v$ para todos os u e v , $u \neq v$, então a densidade univariada dos retornos da carteira R_P é dada pelo **Teorema 3.10.1** e é expressa por uma mistura finita de Erlangs.

$$\begin{aligned} R_P \sim f(R; w_1, \dots, w_D) &= \sum_{i=1}^n p_i \left(\gamma_{b_{i,1}}^{a_1/w_1} \oplus \dots \oplus \gamma_{b_{i,D}}^{a_D/w_D} \right) (R) = \\ &= \sum_{i=1}^n p_i \left\{ \sum_{d=1}^D \left[\prod_{\substack{j=1 \\ j \neq d}}^D \left(1 - \frac{a_d w_j}{w_d a_j} \right)^{-b_{i,j}} \right] \left[\sum_{r=1}^{b_{i,d}} \left(\frac{a_d}{w_d} \right)^{b_{i,d}-r} M_{i,d}^{b_{i,d}-r} \gamma_r^{a_d/w_d} (R) \right] \right\}, \end{aligned} \quad (5.5)$$

onde

$$M_{i,d}^u = \begin{cases} 1 & , \quad u = 0 \\ \frac{1}{u} \sum_{v=0}^{u-1} \left[\sum_{\substack{j=1 \\ j \neq d}}^D b_{i,j} \left(\frac{a_d}{w_d} - \frac{a_j}{w_j} \right)^{v-u} \right] M_{i,d}^v & , \quad u > 0 \end{cases}.$$

Nas fórmulas seguintes, será conveniente uma expressão em que a densidade da carteira não contenha densidades Erlang repetidas. Para isso, é possível fazer

$$R_P \sim f(R; w_1, \dots, w_D) = \sum_{i=1}^n p_i \left(\gamma_{b_{i,1}}^{a_1/w_1} \oplus \dots \oplus \gamma_{b_{i,D}}^{a_D/w_D} \right) (R) = \sum_{d'=1}^{D'} \sum_{r=1}^{b'_{\max,d'}} k_{r,d'} \gamma_r^{a'_{d'}/w'_{d'}} (R), \quad (5.6)$$

onde

$$D' = \text{número de componentes distintos do vetor } \left[\frac{a_1}{w_1}, \dots, \frac{a_d}{w_d}, \dots, \frac{a_D}{w_D} \right]^T, \quad (5.7)$$

$$\left[\frac{a'_1}{w'_1}, \dots, \frac{a'_{d'}}{w'_{d'}}, \dots, \frac{a'_{D'}}{w'_{D'}} \right]^T = \text{vetor } \left[\frac{a_1}{w_1}, \dots, \frac{a_d}{w_d}, \dots, \frac{a_D}{w_D} \right]^T \text{ sem componentes repetidos}, \quad (5.8)$$

$$b'_{i,d'} = \sum_{d=1}^D b_{i,d}, \quad 1 \leq i \leq n, \quad 1 \leq d' \leq D', \quad (5.9)$$

$$\frac{a_d}{w_d} = \frac{a'_{d'}}{w'_{d'}}$$

$$b'_{\max,d'} = \max_{1 \leq i \leq n} b'_{i,d'}, \quad 1 \leq d' \leq D', \quad (5.10)$$

$$k_{r,d'} = \sum_{i=1}^n p_i \left\{ \prod_{\substack{j=1 \\ j \neq d'}}^{D'} \left(1 - \frac{a'_{d'} w'_j}{w'_{d'} a'_j} \right)^{-b'_{i,j}} \right\} \left(\frac{a'_{d'}}{w'_{d'}} \right)^{b'_{i,d'}-r} M_{i,d'}^{b'_{i,d'}-r}, \quad 1 \leq r \leq b'_{\max,d'}, \quad 1 \leq d' \leq D' \quad (5.11)$$

e

$$M_{i,d'}^u = \begin{cases} 1 & , \quad u = 0 \\ \frac{1}{u} \sum_{v=0}^{u-1} \left[\sum_{\substack{j=1 \\ j \neq d'}}^{D'} b'_{i,j} \left(\frac{a'_{d'}}{w'_{d'}} - \frac{a'_j}{w'_j} \right)^{v-u} \right] M_{i,d'}^v & , \quad u > 0 \end{cases} \quad (5.12)$$

O número máximo de componentes Erlang distintos da carteira será

$$\sum_{d'=1}^{D'} b'_{\max,d'}. \quad (5.13)$$

5.4.3. Retorno esperado

O valor esperado dos retornos da carteira $\mathbf{R}_P$ foi dado no **Corolário 3.14.3**.

$$E[\mathbf{R}_P] = \sum_{i=1}^n p_i \left(\frac{w_1 b_{i,1}}{a_1} + \dots + \frac{w_D b_{i,D}}{a_D} \right) = \frac{w_1}{a_1} \sum_{i=1}^n p_i b_{i,1} + \dots + \frac{w_D}{a_D} \sum_{i=1}^n p_i b_{i,D}. \quad (5.14)$$

5.4.4. Downside Risk

Como a minimização do *Downside Risk* elevado a uma potência $\alpha > 0$ leva ao mesmo resultado que a minimização do *Downside Risk*, é preferível usar a primeira alternativa durante os cálculos. A seguir, serão dadas as expressões para o *Downside Risk* elevado a um escalar α genérico, $\alpha = 1$ e $\alpha = 2$. Nestes dois últimos casos, não é necessária a avaliação numérica da função Meijer G. Serão usados os valores D' , $b'_{\max,d'}$ e $k_{r,d'}$ dados em (5.6).

Pelo **Corolário 3.15.1**, o *Downside Risk* elevado à potência α da carteira é

$$DR_{\alpha, R_B}^\alpha[\mathbf{R}_P] = \sum_{d'=1}^{D'} \sum_{r=1}^{b'_{\max,d'}} k_{r,d'} \left(\frac{\Gamma(1+\alpha)}{\Gamma(r)} R_B^\alpha G_{1,2}^{\alpha} \left(\frac{a'_{d'}}{w'_{d'}} R_B \middle| \begin{matrix} 1; \\ r; -\alpha \end{matrix} \right) \right). \quad (5.15)$$

Pelo **Corolário 3.15.2**, para $\alpha = 1$, o *Downside Risk* é

$$DR_{1, R_B}[\mathbf{R}_P] = \sum_{d'=1}^{D'} \sum_{r=1}^{b'_{\max,d'}} k_{r,d'} \left(\left(R_B - r \frac{w'_{d'}}{a'_{d'}} \right) \left(1 - \sum_{h=1}^r \gamma_h \left(\frac{a'_{d'}}{w'_{d'}} R_B \right) \right) + r \frac{w'_{d'}}{a'_{d'}} \gamma_{r+1}^1 \left(\frac{a'_{d'}}{w'_{d'}} R_B \right) \right). \quad (5.16)$$

Pelo **Corolário 3.15.3**, para $\alpha = 2$, o quadrado do *Downside Risk* é

$$DR_{2,R_B}^2[\mathbf{R}_P] = \sum_{d'=1}^{D'} \sum_{r=1}^{b'_{\max,d'}} k_{r,d'} \left\{ \left[r \left(\frac{w'_{d'}}{a'_{d'}} \right)^2 + \left(R_B - r \frac{w'_{d'}}{a'_{d'}} \right)^2 \right] \left[1 - \sum_{h=1}^{r+1} \gamma_h^1 \left(\frac{a'_{d'}}{w'_{d'}} R_B \right) \right] + \right. \\ \left. (r+1) \frac{w'_{d'}}{a'_{d'}} \left(R_B - r \frac{w'_{d'}}{a'_{d'}} \right) \gamma_{r+2}^1 \left(\frac{a'_{d'}}{w'_{d'}} R_B \right) \right\}. \quad (5.17)$$

5.4.5. Formulação do problema de otimização

A formulação do problema de otimização decorre diretamente das expressões dadas para o retorno esperado e o *Downside Risk*. A formulação seguinte será feita para o caso com α genérico.

Problema de otimização 5.4.2. Fronteira eficiente uniperiódica calculada com produto tensorial finito de Erlangs e *Downside Risk*. A medida de risco é o *Downside Risk* DR_{α,R_B} . Os parâmetros α , R_B e o retorno esperado R_E desejado são dados. Existem D ativos dependentes com uma distribuição multivariada de retornos $\mathbf{R}_1, \dots, \mathbf{R}_D$ i.i.d. ao longo do tempo representada por um produto tensorial finito de Erlangs com parâmetros de escala constantes, isto é,

$$[\mathbf{R}_1, \dots, \mathbf{R}_D]^\top \sim m(R_1, \dots, R_D) = \sum_{i=1}^n p_i \gamma_{b_{i,1}}^{a_{i,1}}(R_1) \dots \gamma_{b_{i,D}}^{a_{i,D}}(R_D) \geq 0, \quad \sum_{i=1}^n p_i = 1, \quad 1 \leq n < \infty.$$

Os pesos dos ativos w_d , para $1 \leq d \leq D$, são as variáveis de otimização, devem somar 1 e ser não-negativos. A densidade de uma carteira $\mathbf{R}_P$ é dada por

$$\mathbf{R}_P \sim f(R; w_1, \dots, w_D) = \sum_{i=1}^n p_i \left(\gamma_{b_{i,1}}^{a_{i,1}/w_1} \oplus \dots \oplus \gamma_{b_{i,D}}^{a_{i,D}/w_D} \right)(R) = \sum_{d'=1}^{D'} \sum_{r=1}^{b'_{\max,d'}} k_{r,d'} \gamma_r^{a'_{d'}/w'_{d'}}(R),$$

onde os valores D' , $b'_{\max,d'}$ e $k_{r,d'}$ foram dados em (5.6).

O problema de otimização se torna

$$\min_{w_d} DR_{\alpha,R_B}^\alpha[\mathbf{R}_P] = \Gamma(1+\alpha) R_B^\alpha \sum_{d'=1}^{D'} \sum_{r=1}^{b'_{\max,d'}} \frac{k_{r,d'}}{\Gamma(r)} G_{1,2}^{1,1} \left(\frac{a'_{d'}}{w'_{d'}} R_B \middle| \begin{matrix} 1; \\ r, -\alpha \end{matrix} \right), \quad 1 \leq d \leq D$$

s.t.

$$E[\mathbf{R}_P] = E \left[\sum_{d=1}^D w_d \mathbf{R}_d \right] = \sum_{i=1}^n p_i \left(\frac{w_1 b_{i,1}}{a_1} + \dots + \frac{w_D b_{i,D}}{a_D} \right) = \frac{w_1}{a_1} \sum_{i=1}^n p_i b_{i,1} + \dots + \frac{w_D}{a_D} \sum_{i=1}^n p_i b_{i,D} \geq R_E$$

$$w_d \geq 0, \quad 1 \leq d \leq D$$

$$\sum_{d=1}^D w_d = 1.$$

No **Teorema 4.4.1**, foi provado que o *Downside Risk* é uma medida de risco convexa sobre os pesos não-negativos dos ativos. Portanto, a minimização do risco sobre um domínio convexo é um problema de otimização convexa. A convexidade é uma propriedade bastante importante porque a otimização pode ser feita com um menor esforço computacional.

5.5. Conclusões

Se duas variáveis aleatórias forem independentes, então sua soma pode ser calculada pelo teorema da convolução, juntamente com a transformada de Laplace, ou pela integral da convolução. Se houver dependência entre as variáveis, então o conceito de combinação linear deve ser usado, o qual pode ser visualizado com a técnica de “fatiamento e pesagem”. Outra forma de se obter a densidade da combinação linear é calcular a densidade acumulada e, então, aplicar sua derivada.

O produto tensorial finito de Erlangs possui propriedades matemáticas muito desejáveis em uma análise retorno *versus* risco de carteiras. A combinação linear positiva de variáveis aleatórias dependentes que seguem conjuntamente um produto tensorial finito de Erlangs resulta em uma variável aleatória que segue uma mistura finita de Erlangs.

O problema de otimização uniperiódico foi formulado para o caso em que a densidade dos retornos das carteiras é representado por uma mistura finita de Erlangs. O número de variáveis de otimização é igual ao número de ativos, e o problema de otimização é convexo.

Embora a quantidade de parcelas da mistura finita de Erlangs possa ser grande, o cálculo da esperança é simples. Como o desvio-padrão não tem sentido financeiro para distribuições genéricas, não será usado neste trabalho. O cálculo do *Downside Risk* é bem mais pesado computacionalmente e requer, em geral, a avaliação numérica da função Meijer G. Entretanto, o cálculo pode ser feito com funções mais elementares para os casos particulares $\alpha = 1$ e $\alpha = 2$.

O *Downside Risk* é uma medida de risco convexa sobre os pesos dos ativos nas carteiras, para o caso particular em que $\alpha \geq 1$, R_B é fixo e os pesos dos ativos são não-negativos. Neste caso, o problema de otimização uniperiódico é convexo. Para investidores racionais, os parâmetros devem satisfazer as condições $\alpha \geq 1$ e $1 \leq 1 + r_f \leq R_B$.

6. OTIMIZAÇÃO MULTIPERÍODICA DE CARTEIRAS

6.1. Introdução

Após o cálculo da fronteira eficiente uniperiódica, uma questão natural que surge é como estendê-lo para mais de um período. Neste trabalho, é adotada a hipótese de que a função densidade de probabilidade multivariada uniperiódica será i.i.d. nos períodos futuros. Ser independente significa que não há dependência entre as densidades multivariadas em diferentes instantes de tempo. Ser identicamente distribuída significa que a densidade multivariada será a mesma em todos os períodos. A melhor estimativa para a densidade multivariada de cada período futuro é usualmente a calculada para o caso uniperiódico.

Considere que o horizonte de tempo tem T períodos, onde $T \in \mathbb{N}$, $1 \leq T < \infty$. O objetivo é obter, no período inicial ($t = 0$), a fronteira eficiente multiperiódica no gráfico retorno *versus* risco, na qual cada ponto representa uma ou mais seqüências ordenadas de T carteiras. Cada carteira de uma dada seqüência é determinada pelos pesos de seus ativos e deve ser mantida durante o período t . Os pesos dos ativos das carteiras de uma seqüência não precisam ser constantes ao longo do tempo. Caso sejam, a carteira é chamada de carteira com pesos constantes.

Um investidor racional pode, então, escolher um ponto (uma seqüência de carteiras) sobre a fronteira eficiente multiperiódica, de acordo com seu retorno desejado ou risco tolerado. Aqui será assumido que o investidor fixa o retorno (composto médio) esperado e não o risco, porque isso é computacionalmente mais conveniente.

Um detalhe importante na análise multiperiódica é que a variável retorno (discreto) é do tipo multiplicativo ao longo do tempo, isto é, os retornos uniperiódicos (em tempo discreto) das carteiras de uma seqüência devem ser multiplicados para se obter o retorno acumulado em todo o período. Uma forma alternativa para os cálculos é a utilização de retornos contínuos.

A distribuição multiperiódica das carteiras será ajustada para uma distribuição “composta média”, de forma que a distribuição fique na escala uniperiódica. Com isso, a medida de risco não precisa ser ajustada. Posteriormente, este ajuste será generalizado, de forma a considerar a longevidade do investidor.

Por causa da hipótese i.i.d., é possível aplicar a técnica de programação dinâmica, desenvolvida por Bellman (1957); veja Markowitz (1991a, p.276). O cálculo deve ser feito do futuro para o presente, pois a carteira do instante t depende da carteira do instante $t + 1$, mas não depende da carteira do instante $t - 1$. No último período ($t = T - 1$), uma carteira deve ser escolhida com o retorno desejado, e a otimização

uniperiódica deve ser feita. No penúltimo período ($t = T - 2$), a carteira será escolhida considerando-se que a carteira do período seguinte já é conhecida, e de tal forma que o risco nos dois últimos períodos seja minimizado para um dado nível de retorno (composto médio) esperado. Este procedimento é repetido para os demais períodos.

No fim de cada período, uma nova amostra de retornos será observada, e esta informação pode ser utilizada para um novo recálculo completo. A estratégia se torna dinâmica.

Comentário 6.1.1. Uma forma de se obter a densidade multivariada para T períodos é usar amostras disjuntas de T períodos. Mas isto requer um longo período de dados históricos, e a conjuntura econômica do passado é provavelmente muito diferente da do presente. Outra forma seria um tipo de “auto-convolução multiplicativa multivariada”, mas isto ainda é um problema de pesquisa.

6.2. Mundos de retornos discretos e contínuos

Se um ativo tiver uma taxa de retorno de 1% no primeiro dia e 3% no segundo, então a taxa de retorno efetivo médio será de $\sqrt{1.01 \cdot 1.03} - 1 = 1.995\%$. No mundo discreto, a taxa de retorno é obtida a partir do retorno, que é uma variável do tipo multiplicativo. Se a adição for aplicada, então o resultado numérico será aproximadamente igual ao efetivo, mas somente se as taxas de retorno forem pequenas. Para o cálculo financeiro da taxa de retorno médio não se pode usar a adição. Isto significa que a operação Adição (no sentido financeiro) de taxas de retorno discreto (para o mesmo ativo, em instantes diferentes) não está definida. Isto implica que a média aritmética, a integral e a esperança no tempo (isto é, de taxas de retorno discreto em diferentes instantes) também não estão definidas, pois dependem da adição.

Um ativo que paga como dividendo anual uma taxa de retorno discreto r_1 pode alternativamente pagar um dividendo r_2 capitalizado semestralmente. A taxa r_2 deve ser menor que r_1 , de forma que os dois rendimentos finais sejam iguais. Algebricamente,

$$1 + r_1 = \left(1 + \frac{r_2}{2}\right)^2 \Rightarrow r_2 = 2(\sqrt{1 + r_1} - 1).$$

A taxa de retorno discreto r_1 pode ser transformada em uma taxa de retorno r_n capitalizada com um número qualquer n de períodos. No limite quando n tende a infinito, a taxa de retorno discreto r_n se transforma no retorno contínuo $\tilde{R}_1$, que deve ser capitalizado e pago continuamente. Aplicando (3.113), tem-se

$$\lim_{n \rightarrow \infty} \left(1 + \frac{r_n}{n}\right)^n = e^{r_\infty} = 1 + r_1 \Rightarrow \tilde{R}_1 = r_\infty = \ln(1 + r_1).$$

No exemplo a seguir, será mostrado que a operação Adição pode ser usada com os retornos contínuos, isto é, a Adição está bem definida no mundo de retornos contínuos (para o mesmo ativo, em instantes diferentes). O retorno contínuo é uma variável do tipo aditivo. É necessário levar os retornos discretos para o mundo de retornos contínuos aplicando a transformada $\ln()$, realizar a adição, e levar o

resultado de volta para o mundo de retornos discretos aplicando a transformada $\exp()$.

Taxa de retorno discreto (r)	Retorno discreto ($R = 1 + r$)	Transformada	Retorno contínuo ($\tilde{R}$)
1%	1.01	$\xrightarrow{\ln()}$	0.995%
3%	1.03	$\xrightarrow{\ln()}$	2.956%
			Soma = 3.951%
1.995%	1.01995	$\xleftarrow{\exp()}$	Média aritmética = 1.975%

Neste exemplo, pode ser visto que voltar para o mundo discreto é importante.

Agora suponha que um investidor tenha uma carteira com dois ativos e aloque 50% de seu valor no ativo 1 e 50% no ativo 2, isto é, $w_1 = w_2 = 50\%$. Se a taxa de retorno discreto do ativo 1 for 2% e a do ativo 2 for 6%, então a taxa de retorno discreto da carteira será $0.5 \cdot 2\% + 0.5 \cdot 6\% = 4\%$. Isto significa que a operação Adição (no sentido financeiro) de taxas de retorno discreto (do mesmo ou de diferentes ativos) em um mesmo instante está bem definida.

Como deve ser calculado o retorno contínuo de uma carteira? A seguir, serão mostradas algumas possíveis formas de cálculo.

Alternativa	Possível transformada para o mundo de retornos contínuos	Retorno contínuo ($\tilde{R}$)	Taxa de retorno discreto ($r = \exp(\tilde{R}) - 1$)
A	$0.5 \cdot \ln(1.02) + 0.5 \cdot \ln(1.06)$	3.904%	3.981%
B	$\ln(0.5 \cdot 1.02) + \ln(0.5 \cdot 1.06)$	-130.822%	-72.970%
C	$\ln(1 + 0.5 \cdot 0.02) + \ln(1 + 0.5 \cdot 0.06)$	3.951%	4.030%
D	$\ln(0.5 \cdot \ln(1.02) + \ln(0.5 \cdot \ln(1.06))$	-5.412%	-5.268%
E	$\ln(0.5 \cdot 1.02 + 0.5 \cdot 1.06)$	3.922%	4.000%
F	$\ln(1 + 0.5 \cdot 0.02 + 0.5 \cdot 0.06)$	3.922%	4.000%

As alternativas com o resultado correto são E e F, e são iguais, pois

$$\sum w_i R_i = \sum w_i (1 + r_i) = \sum (w_i + w_i r_i) = \sum w_i + \sum w_i r_i = 1 + \sum w_i r_i.$$

Quem deve ser levado para o mundo dos retornos contínuos é o retorno discreto da carteira, que é a média ponderada dos retornos discretos em um certo instante. Os pesos w_i do mundo discreto não são diretamente mapeáveis para o mundo de retornos contínuos. As operações de adição de retornos discretos e de taxas de retorno discreto estarão bem definidas se as parcelas estiverem no mesmo instante.

O mundo aditivo ou de retornos contínuos é um homeomorfismo (um caso especial de isomorfismo) do mundo multiplicativo ou de retornos discretos; veja Lima (1987, p.276). O mundo de retornos contínuos é usado nos casos em que os cálculos matemáticos são facilitados. Nele, pode-se empregar as ferramentas

usuais do Cálculo e da Estatística, pois a Adição (no tempo) está bem definida.

Note que “retorno discreto” é um termo técnico que deve ser entendido como retorno no mundo (de tempo) discreto. Isto não implica que os retornos sejam variáveis discretas. Neste trabalho, os retornos discretos e contínuos são variáveis contínuas em cada instante de tempo, mas o tempo é uma variável discreta.

Se o intervalo de tempo entre dois períodos tender a zero (estritamente maior que zero), os retornos do mundo discreto tenderão a ficar ao redor do valor 1, e os retornos contínuos, ao redor do valor 0. Note que os domínios desses mundos não vão se alterar. É possível subtrair o valor 1 dos retornos discretos e transformar essa variável na taxa de retorno discreto, com valores ao redor de 0. Então surge uma questão que não é muito discutida pelos autores: o mundo das taxas de retorno discreto converge para o mundo dos retornos contínuos? Se for aceitável uma resposta aproximada, então não há necessidade de muita discussão. Entretanto, em geral, esses mundos estão separados, isto é, eles não irão convergir para o mesmo mundo quando o intervalo de tempo tender a zero. O fato de os domínios serem diferentes indica que, em situações nas quais o domínio é importante, os dois mundos não podem ser considerados o mesmo. O fato de uma variável ser do tipo multiplicativo e outra ser do tipo aditivo pode causar inconsistências em demonstrações matemáticas. O fato de os pesos dos ativos das carteiras não serem diretamente mapeáveis para o mundo contínuo também pode causar problemas. Uma distribuição Delta de Dirac no ponto de retorno discreto 0 deve ser mapeada em uma Delta de Dirac no ponto de retorno contínuo $-\infty$, de forma que ambas estarão bem distantes do ponto de taxa de retorno zero, qualquer que seja o intervalo de tempo. Por causa desses problemas, em um estudo matematicamente mais rigoroso, é preferível fazer todos os cálculos no mundo dos retornos discretos, e só então aplicar o limite quando o intervalo de tempo tende a zero. Veja também Merton (1971, p.412), Merton (1975) e Rubinstein (1976, p.553).

6.3. Variável aleatória multiplicativa

Uma “variável aleatória” é uma variável cuja realização futura é incerta. Um caso limite é a distribuição Delta de Dirac, para a qual o valor futuro é certo. A definição precisa de “variável aleatória” não será apresentada neste trabalho porque requer vários conceitos matemáticos prévios. Ela pode ser encontrada, por exemplo, em Shafer e Vovk (2001, p.41).

Nesta seção, será definido informalmente que “variável aleatória multiplicativa” é uma variável aleatória que tem a operação de multiplicação bem definida, isto é, o produto de duas variáveis aleatórias multiplicativas resulta em uma variável aleatória com sentido físico, econômico ou financeiro. Esta qualificação é importante porque nem todas as variáveis aleatórias, quando multiplicadas, geram um resultado com sentido físico, como por exemplo, as que possuem valores negativos no domínio. De forma análoga, uma “variável aleatória aditiva” é uma variável aleatória que tem a operação de soma bem definida. Se uma variável aleatória é do tipo retorno, mas é definida sobre retornos negativos, isto é, existe possibilidade de falência, então é possível definir uma regra que, antes de serem multiplicadas, as densidades

devem ser truncadas no ponto 0, e nesse ponto deve ser adicionada uma Delta de Dirac com a massa da densidade de probabilidade dos retornos negativos. Entretanto, isto não será feito neste trabalho, porque foi adotada a hipótese de que há probabilidade zero de haver retornos menores ou iguais a 0.

Este conceito é conhecido pelos matemáticos pelo menos desde 1900. Galton sabia que o produto de duas variáveis aleatórias independentes que seguem distribuições lognormais gera uma variável aleatória que também segue uma lognormal; veja Aitchison e Brown (1963).

6.4. Distribuição multiperiódica dos retornos

No capítulo anterior, foi visto que a distribuição uniperiódica dos retornos de uma carteira é obtida por uma combinação linear dos ativos, na qual os pesos têm soma unitária. Neste trabalho, os pesos são, por hipótese, não-negativos.

Neste capítulo, foi visto que o retorno (em tempo) discreto é uma variável aleatória aditiva em um dado instante de tempo e multiplicativa ao longo do tempo. Isto significa que existe sentido financeiro na soma de dois retornos discretos em um mesmo instante de tempo e na multiplicação de dois retornos discretos em diferentes instantes de tempo. A média aritmética pode ser aplicada somente se a soma estiver bem definida; e a média geométrica, se o produto estiver bem definido.

A distribuição multiperiódica de uma sequência de variáveis aleatórias é a distribuição seguida pelo produto dessas variáveis. É importante lembrar que o retorno discreto é uma variável aleatória multiplicativa ao longo do tempo e, portanto, a convolução aditiva não pode ser aplicada a esses retornos. Neste trabalho, as distribuições são, por hipótese, independentes ao longo do tempo.

Uma forma de se calcular a distribuição multiperiódica de variáveis aleatórias independentes é levar a distribuição dos retornos discretos para o mundo contínuo, realizar convoluções aditivas e trazer o resultado de volta para o mundo discreto. É necessário fazer a transformação $\tilde{R} = \ln(R)$, onde $\tilde{R}$ é o retorno contínuo, que pode ser negativo, e R é o retorno discreto, que deve ser não-negativo. A densidade no mundo discreto é transformada na densidade no mundo contínuo da seguinte forma: $\tilde{f}(\tilde{R}) = e^{\tilde{R}} f(e^{\tilde{R}})$, onde $\tilde{f}(\tilde{R})$ é a densidade contínua definida sobre o domínio dos retornos contínuos, e $f(R)$ é a densidade discreta definida sobre o domínio dos retornos discretos. No mundo dos retornos contínuos, a adição ao longo do tempo está bem definida e a convolução aditiva pode ser aplicada, usando-se, por exemplo, a transformada de Laplace. Esta técnica é uma mudança de variáveis e é conhecida pelos matemáticos; veja Williamson (1989, p.30). O exemplo mais conhecido é o da distribuição lognormal, que pode ser mapeada para a distribuição normal no mundo contínuo. Nesse mundo, a soma de duas variáveis aleatórias independentes que seguem distribuições normais gera uma outra que também segue uma normal. Esta última pode mapeada de volta para uma distribuição lognormal. Isto significa que o produto de duas variáveis independentes que seguem distribuições lognormais gera uma variável que também segue uma distribuição lognormal.

Outra forma de se calcular a distribuição multiperiódica de variáveis aleatórias independentes é a técnica de convolução multiplicativa. Isto será visto a seguir.

6.5. Convolução multiplicativa

Se duas variáveis aleatórias multiplicativas independentes X_1 e X_2 têm suas funções densidade de probabilidade conhecidas, então é possível determinar a densidade de $X_1 \cdot X_2$ aplicando a integral da convolução multiplicativa; ou o teorema da convolução multiplicativa ou de escala, juntamente com a transformada de Mellin; veja o **Corolário 3.7.2**.

O teorema da convolução multiplicativa e a transformada de Mellin foram dados na **Seção 3.7**. É sempre bom lembrar que a convolução só vale para variáveis aleatórias independentes, e esta hipótese foi adotada neste trabalho. Matematicamente,

$$\begin{aligned} X \sim f(x) &\xrightarrow{\mathcal{M}} F(s) = \int_0^{\infty} x^{s-1} f(x) dx \\ &\Downarrow \\ X_1 \cdot X_2 \sim (f_1 \otimes f_2)(x) &= \int_0^{\infty} f_1(t) f_2\left(\frac{x}{t}\right) \frac{1}{t} dt \xrightarrow{\mathcal{M}} F_1(s) F_2(s), \quad X_1 \text{ indep. } X_2. \end{aligned} \quad (6.1)$$

Note que fazendo $s = 2$ na transformada de Mellin, obtém-se o primeiro momento de X , ou seja, $F(2) = E[X]$, e fazendo $s = 3$, obtém-se o segundo momento.

A transformada de Mellin geralmente é mais fácil de se calcular que a transformada de Laplace.

6.6. Esperança e variância de variáveis aleatórias multiplicativas em função do tempo

Um resultado importante relacionado com as variáveis aleatórias multiplicativas (que podem ser multiplicadas) é o comportamento de sua esperança e variância no tempo.

Teorema 6.6.1. Esperança e variância de uma variável aleatória multiplicativa em função do tempo. Se X é uma variável aleatória multiplicativa, é i.i.d., segue uma densidade $f(x)$ e tem uma transformada de Mellin $F(s)$, então sua esperança e variância após $t \geq 0$ períodos são

$$E \left[\underbrace{X \cdot X \cdot \dots \cdot X}_{t \text{ cópias independentes}} \right] = E[X^{\times t}] = (F(2))^t \quad (6.2)$$

e

$$V \left[\underbrace{X \cdot X \cdot \dots \cdot X}_{t \text{ cópias independentes}} \right] = V[X^{\times t}] = (F(3))^t - (F(2))^{2t}. \quad (6.3)$$

Prova. Para t inteiro positivo,

$$Y = X^{\times t} = \underbrace{X \cdot X \cdot \dots \cdot X}_t \sim g(x) = \underbrace{(f \otimes \dots \otimes f)}_t(x) = f^{\otimes t}(x) \xrightarrow{\mathcal{M}} G(s) = (F(s))^t.$$

A esperança de Y é

$$E[Y] = G(2) = (F(2))^t.$$

A variância de Y é

$$V[Y] = G(3) - (G(2))^2 = (F(3))^t - (F(2))^{2t}.$$

Para t positivo e não-inteiro (tempo contínuo), as expressões são as mesmas. A mudança do domínio do expoente dos inteiros positivos para os racionais positivos e, então, para os reais positivos, pode ser justificada por um salto indutivo.

□

Definições relacionadas de integral fracionária podem ser encontradas em Gorenflo e Mainardi (2000).

Para o caso $t = 1$, as expressões (6.2) e (6.3) podem ser encontradas em Park e Sharp-Bette (1990, p.394). Para o caso $t = 1, 2, \dots$, expressões equivalentes podem ser encontradas em Tobin (1965, p.46), Levy (1972), Gressis, Philippatos e Hayya (1976), Marshall (1994) e Gunthorpe e Levy (1994).

A esperança e a variância de uma variável aleatória multiplicativa e i.i.d. são, em geral, não-lineares ao longo do tempo. Por outro lado, para uma variável aleatória aditiva e i.i.d., estes momentos crescem linearmente no tempo. Em grande parte da literatura não ocorre a devida distinção entre estes dois tipos de variáveis. É muito comum encontrar artigos cujos autores dizem que para se obter o desvio-padrão de T períodos, basta multiplicar o desvio-padrão uniperiódico por $\sqrt{T}$. Isto vale somente para variáveis aleatórias aditivas e i.i.d. Note que é possível transformar uma variável aleatória multiplicativa em uma aditiva, levando-a para o mundo aditivo ou contínuo com a transformada $\tilde{R} = \ln(R)$. Nesse mundo, as variáveis aleatórias independentes podem ser somadas e ter o desvio-padrão (volatilidade) multiplicado por $\sqrt{T}$. Após os cálculos, a variável aleatória resultante deve ser trazida de volta para o mundo multiplicativo ou discreto.

Exemplo 6.6.1. Variável aleatória multiplicativa e i.i.d. seguindo uma lognormal. Para uma variável aleatória multiplicativa e i.i.d. que segue uma distribuição lognormal, foi calculada sua transformada de Mellin, e plotados o gráfico da densidade e os gráficos da esperança, do desvio-padrão e da variância em função do tempo.

$$X \sim f(x) = \text{Log}N_{\nu, \sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{(\ln x - \nu)^2}{2\sigma^2}} \xrightarrow{\mathcal{M}} F(s) = e^{\nu(s-1) + \frac{\sigma^2}{2}(s-1)^2}$$

A esperança e variância de $X^{\times t}$ são

$$E[X^{\times t}] = (F(2))^t = e^{t(\nu + \sigma^2/2)} \quad (6.4)$$

e

$$V[X^{\times t}] = (F(3))^t - (F(2))^{2t} = e^{2t(\nu + \sigma^2)} - e^{t(2\nu + \sigma^2)} = e^{t(2\nu + \sigma^2)} (e^{t\sigma^2} - 1). \quad (6.5)$$

Para o caso particular $\nu = 0.01$ e $\sigma = 0.02$, os gráficos a seguir foram plotados.

Figura 6.6.1. Gráfico da densidade, e gráficos da esperança, desvio-padrão e variância em função do tempo para uma variável aleatória lognormal multiplicativa e i.i.d. Para os parâmetros $\nu = 0.01$ e $\sigma = 0.02$, foram traçados os gráficos de $X \sim \text{LogN}_{\nu, \sigma}(x)$ e de $E[X^{xt}]$, $\sqrt{V[X^{xt}]}$ e $V[X^{xt}]$ em função do tempo.

■

Para o caso particular da lognormal, as expressões (6.4) e (6.5) são bem conhecidas; veja, por exemplo, Marshall (1994), Hull (1998, p.250) ou Luenberger (1998, p.310).

6.7. Distribuição composta média

Nesta seção, será apresentado o conceito de distribuição composta média, que foi criado por Hakansson (1971a). Na seção seguinte, este conceito será generalizado para a distribuição composta média para investidores “longevos” (ou com “longevidade”).

Após a obtenção da densidade multiperódica de uma sequência de T carteiras, é necessário calcular o retorno esperado e o risco, que serão usados na construção da fronteira eficiente multiperódica. Ao longo de T períodos, os retornos uniperiódicos se acumularão e, considerando-se que a média uniperiódica é maior ou igual a 1, a distribuição multiperódica terá média e variância maiores ou iguais às da distribuição uniperiódica; veja o **Exemplo 6.6.1**.

Há duas alternativas para o cálculo do risco e do retorno de uma carteira multiperódica.

1. Encontrar a distribuição multiperódica dos T períodos ou do instante final, ajustar a medida de risco para a escala multiperódica e, então, calcular sua esperança e risco, que estarão nessa

escala.

2. Ajustar a distribuição multiperíódica dos T períodos para a escala uniperíódica, de forma que ela se torne uma “distribuição composta média” e, então, calcular sua esperança e seu risco da forma usual, resultando em valores na escala uniperíódica.

Se a primeira alternativa for usada, então a questão se torna como ajustar a medida de risco. Por exemplo, se o valor do retorno padrão do *Downside Risk* for $R_B = 1$, então seu valor elevado à potência T será o mesmo R_B . Como a distribuição multiperíódica se torna bastante achatada, o risco calculado para os retornos menores que $R_B = 1$ se torna bem pequeno. É necessário alterar essa medida de risco de alguma forma não-trivial; caso contrário, à medida que T aumenta, as carteiras se tornam cada vez mais arriscadas.

Na segunda alternativa, a distribuição composta média é obtida com a elevação da distribuição multiperíódica dos T períodos à potência $1/T$ (na seção seguinte será mostrado que a potência pode ter outros valores além desse). A distribuição multiperíódica é levada para a escala uniperíódica. Neste caso, a medida de risco não precisa ser alterada (embora possa ser, caso o investidor queira), e os cálculos do retorno esperado e do risco são feitos da mesma forma que no caso uniperíódico. Note que a elevação de uma variável aleatória à potência $n \in \{1, 2, \dots\}$ não é o mesmo que fazer convoluções multiplicativas com n variáveis, porque a potenciação é feita sobre uma mesma variável aleatória, enquanto a convolução é feita sobre variáveis aleatórias (ou cópias) independentes.

A distribuição da primeira alternativa é a de uma variável aleatória no instante T . A distribuição da segunda alternativa é a de uma variável aleatória tal que se a primeira realização for elevada à potência T , irá gerar uma variável aleatória com a distribuição da primeira alternativa.

A seguir, o conceito de distribuição composta média será definido para o caso particular em que a longevidade do investidor não é considerada.

Definição 6.7.1. Distribuição composta média. Dadas as variáveis aleatórias multiplicativas $X_0, X_1, \dots, X_{T-1}$, para $T \geq 1$, então $(X_0 \cdot X_1 \cdot \dots \cdot X_{T-1})^{1/T}$ segue a densidade de probabilidade composta média $\varphi(x)$.

Se $X_0, \dots, X_{T-1}$ são variáveis aleatórias independentes que seguem, respectivamente, as densidades $f_0(x), \dots, f_{T-1}(x)$, então

$$(X_0 \cdot X_1 \cdot \dots \cdot X_{T-1})^{1/T} \sim \varphi(x) = \left((f_0 \otimes f_1 \otimes \dots \otimes f_{T-1})^{\otimes \frac{1}{T}} \right)(x), \quad (6.6)$$

e se as densidades forem i.i.d., isto é, $f(x) = f_0(x) = f_1(x) = \dots = f_{T-1}(x)$, então

$$\underbrace{(X \cdot X \cdot \dots \cdot X)}_T^{1/T} = (X^{\times T})^{1/T} \sim \varphi(x) = \left(\underbrace{(f \otimes \dots \otimes f)}_T^{\otimes \frac{1}{T}} \right)(x) = (f^{\otimes T})^{\frac{1}{T}}(x). \quad (6.7)$$

A seguir, serão dadas as relações entre as densidades e transformadas de Mellin de X , $X^{\times T}$, $X^{1/T}$ e $(X^{\times T})^{1/T}$.

$$X \sim f(x) \xrightarrow{\mathcal{M}} F(s) \quad (6.8)$$

$$X^{\times T} = \underbrace{X \cdot \dots \cdot X}_T \sim \underbrace{(f \otimes \dots \otimes f)}_T(x) \xrightarrow{\mathcal{M}} (F(s))^T \quad (6.9)$$

$$X^{1/T} \sim f^{\circledast \frac{1}{T}}(x) = Tx^{T-1}f(x^T) \xrightarrow{\mathcal{M}} F\left(1 + \frac{s-1}{T}\right) \quad (6.10)$$

$$\left(X^{\frac{1}{T}}\right)^{\times T} = (X^{\times T})^{\frac{1}{T}} \sim \varphi(x) = \left(\left(f^{\circledast \frac{1}{T}}\right)^{\otimes T}\right)(x) = \left((f^{\otimes T})^{\circledast \frac{1}{T}}\right)(x) \xrightarrow{\mathcal{M}} \left(F\left(1 + \frac{s-1}{T}\right)\right)^T \quad (6.11)$$

A esperança é calculada fazendo $s = 2$ na transformada de Mellin. As duas alternativas têm, em geral, valores diferentes para a esperança. Por exemplo, suponha que as T carteiras sejam iguais, isto é, há T cópias independentes ao longo do tempo. Na primeira alternativa, a esperança será $(F(2))^T$, que, trazida para a escala uniperiódica resulta em $F(2)$. Na segunda alternativa, a esperança na escala uniperiódica será $F(1 + 1/T)$. Para o caso uniperiódico, quando $T = 1$, ambas as alternativas recomendam a maximização de $F(2)$ no instante $t = 0$. Mas se $T > 1$, as esperanças das distribuições multiperiódicas das duas alternativas serão, em geral, diferentes. Esta diferença causa uma dúvida, que será discutida a seguir.

Se um investidor com horizonte de tempo finito T deseja um bom resultado ao final desse período, qual deve ser a alternativa a ser escolhida por ele: maximizar a esperança da distribuição no instante final T ou maximizar a esperança da distribuição composta média? Muitas pessoas diriam ao investidor que maximize a esperança da distribuição no instante final T , mas isto causa um problema que será mostrado no exemplo seguinte.

Exemplo 6.7.1. Comparação entre os critérios de maximização da esperança da distribuição no instante final e de maximização da esperança da distribuição composta média. Foram escolhidos dois ativos com distribuição de retornos i.i.d.; o primeiro com duas Deltas de Dirac e média 5.9 e o segundo com uma Delta de Dirac e média 5.0 ao período. O investidor quer escolher um desses ativos e mantê-lo por $T \geq 1$ períodos (carteira constante). As densidades uniperiódicas dos retornos desses dois ativos e as transformadas de Mellin foram traçadas na **Figura 6.7.1**. Um investidor que segue a maximização da esperança da distribuição no instante final escolherá o ativo com máximo $(F(2))^T$, ou, neste exemplo, máximo $F(2)$, qualquer que seja seu horizonte de tempo. Este investidor escolherá sempre o primeiro ativo. Um investidor que segue a maximização da distribuição composta média escolherá a carteira com máximo $(F(1 + 1/T))^T$, ou, neste exemplo, máximo $F(1 + 1/T)$. Pelo gráfico da transformada de Mellin, para um horizonte de tempo $T = 1$, este investidor escolherá o primeiro ativo e, para $T \geq 2$, ele irá preferir o segundo.

Foram feitos quatro experimentos Monte Carlo, com $n = 100$ simulações cada, para diferentes valores de T . Em cada simulação, para cada ativo, foram sorteados T retornos e, então, foi calculado o seu produto. Os gráficos das simulações Monte Carlo, mostrados na **Figura 6.7.2**, indicam que o critério de

maximização da distribuição composta média produz melhores resultados e, quanto maior o horizonte de tempo, maior a sua vantagem sobre o outro critério. A implementação em Mathematica está no **Apêndice A**.

Figura 6.7.1. Gráficos da densidade uniperiódica dos retornos de dois ativos e de suas transformadas de Mellin. O ativo 1) tem uma densidade dada por $0.9\delta(x-1)+0.1\delta(x-50)$. O ativo 2) tem uma densidade dada por $\delta(x-5)$. As médias dos ativos são, respectivamente, 5.9 e 5.0.

Figura 6.7.2. Gráficos de quatro experimentos Monte Carlo comparando os critérios de escolha de carteira. A esperança uniperiódica do primeiro ativo é maior que a do segundo, mas à medida que T aumenta, a variância do primeiro ativo cresce muito rapidamente e torna a média cada vez menos relevante para uma decisão financeira.

No exemplo anterior, foi mostrado que o ativo com a maior média uniperiódica não é necessariamente o que deve ser escolhido para um horizonte de tempo finito maior que um período. Os retornos dos dois ativos são sempre maiores ou iguais a 1 e, portanto, seus riscos, medidos pelo *Downside Risk* com um

α genérico e $R_B = 1$, serão iguais a zero. À medida que o tempo aumenta, a variância do primeiro ativo aumenta rapidamente; ele poderá ter altos retornos, mas com baixa probabilidade. Isto faz a esperança dos retornos discretos no instante T perder relevância em uma decisão financeira. Uma análise média *versus* variância detectaria um grande risco no ativo 1, mas já foi visto que a variância não é uma boa medida de risco por penalizar os retornos altos de distribuições genéricas.

No capítulo seguinte, será provado que, se T tender a infinito, o critério de maximização da distribuição composta média tende ao critério de Kelly. Esta é uma propriedade importante que não é contemplada por outros modelos multiperiodicos na literatura.

Um exemplo extremo que mostra a diferença entre estes dois critérios ou distribuições é imaginar um investidor que tem um horizonte de tempo finito T e precisa escolher uma carteira constante dentre as alternativas a seguir.

1. Um ativo com 1% de probabilidade de retorno 0 e 99% de probabilidade de retorno 2 por período.
2. Um ativo com 100% de probabilidade de retorno 1.01 por período.

A esperança uniperiódica da primeira alternativa é 1.98 e é bem maior que a da segunda, que é 1.01. Mas à medida que o horizonte de tempo T aumenta, as chances são de que o ativo 1 tenha um retorno igual a zero e se mantenha neste nível (falência) devido à natureza multiplicativa dos retornos.

A distribuição composta média é bastante conveniente, pois é fácil de se calcular, evita ajustes na medida de risco (mas se o investidor desejar, ele pode alterar sua aversão ao risco em função de seu horizonte de tempo), evita problemas tal como o mostrado no **Exemplo 6.7.1** e generaliza as distribuições usadas nos modelos de Markowitz e de Kelly. O conceito de retorno esperado é generalizado e passa a ser “retorno composto médio” esperado.

O critério de maximização da esperança da distribuição composta média não invalida o da maximização da esperança (da utilidade) da distribuição multiperiodica terminal da sequência de T carteiras. O critério de escolha de carteiras depende do investidor e, portanto, não cabe discussão. Mas ele deve estar ciente dos efeitos mostrados por experimentos Monte Carlo quando T é grande ou tende a infinito. Um problema com a adoção da distribuição multiperiodica terminal é que, antes de isso ser feito na abordagem retorno *versus* risco, deve-se encontrar uma forma de ajustar a medida de risco para a escala multiperiodica.

Hakansson (1971a) definiu o conceito de distribuição composta média para seqüências de carteiras nas quais as carteiras podem ter probabilidade de falência, isto é, a densidade dos retornos de uma dada carteira pode ter uma Delta de Dirac no ponto 0. Como isto causa complicações matemáticas, neste trabalho, esta situação está excluída por hipótese.

Hakansson (1971a) separou o caso em que o horizonte de tempo T é grande, mas finito, do caso em que T tende a infinito. O segundo caso será discutido no capítulo seguinte. No primeiro caso, sabendo que a potência $1/T$ de uma variável aleatória que segue uma lognormal gera uma variável aleatória que também segue uma lognormal, Hakansson (1971a, p.869) calculou as expressões para a esperança e variância de uma

variável lognormal, na medida de probabilidade composta média, para um horizonte de tempo finito T . A seguir, ele usou essas expressões para traçar, para um dado T , a fronteira eficiente de variáveis aleatórias que seguem lognormais no gráfico variância *versus* esperança. Ele não mencionou programação dinâmica, que é uma técnica importante para a determinação da sequência ótima de carteiras. Isto será discutido mais tarde, neste capítulo.

6.7.1. Função utilidade potência

A distribuição composta média é uma mudança na medida de probabilidade da distribuição terminal ou multiperíodica. Além disso, este conceito tem uma importante relação com a função utilidade potência. Hakansson (1971a) mencionou a função utilidade potência $U(x) = x^{1/c}$, para $c \geq 1$. Supondo que a riqueza inicial do investidor seja $W_0 = 1$ e que $X_0 \cdot \dots \cdot X_{T-1} \sim g(x)$, então

$$(X_0 \cdot \dots \cdot X_{T-1})^{1/T} \sim \varphi(x) = g^{\otimes 1/T}(x) = Tx^{T-1}g(x^T).$$

Para $c = T$, a esperança da utilidade é

$$E[U(X)] = \int_0^\infty x^{1/T} g(x) dx,$$

e a esperança da distribuição composta média é

$$E(\varphi(x)) = \int_0^\infty x\varphi(x)dx = \int_0^\infty xTx^{T-1}g(x^T)dx = T \int_0^\infty x^T g(x^T)dx = T \int_0^\infty yg(y)\frac{1}{T}y^{\frac{1-T}{T}}dy = \int_0^\infty y^{\frac{1}{T}}g(y)dy.$$

Portanto, a esperança da utilidade $U_1(x) = x^{1/T}$, para $T \geq 1$, na medida de probabilidade (composta) do instante T , é igual à esperança da utilidade $U_2(x) = x$, na medida de probabilidade composta média. Note que é possível a adoção de uma outra função utilidade sobre a distribuição composta média.

A função utilidade potência é um caso particular das funções utilidade isoelástica, CRRA (*constant relative risk aversion*) e HARA (*hyperbolic absolute risk aversion*). O valor do coeficiente de aversão relativa ao risco é constante. Entretanto, para haver equivalência entre essa função utilidade e o conceito de distribuição composta média, o parâmetro c deve variar com o horizonte de tempo T . Pode-se dizer que a distribuição composta média “induz” um caráter normativo à função utilidade potência. Muitos autores discutiram a maximização da função utilidade potência; veja Samuelson e Merton (1974), Elton e Gruber (1974a, 1974b), Thorley (1995), Eaton e Conover (1998, 2004), Gollier (2001) e Stutzer (2004). Hakansson, em vários artigos posteriores, também usou o conceito de utilidade potência ou isoelástica, mas assumiu que c é constante. Entretanto, veja Grauer e Hakansson (1982, nota 14). A função utilidade CRRA tem uma importante relação com o conceito de separação em um fundo; veja as referências em Constantinides e Malliaris (1995, p.17).

Pelo teorema do limite central no caso multiplicativo, sob certas condições, a densidade de probabilidade do produto de variáveis aleatórias irá tender a uma densidade lognormal. Este caso do teorema do

limite central pode ser entendido com o homeomorfismo dos mundos multiplicativo e aditivo.

Hakansson (1971a) aplicou o teorema do limite central a uma seqüência de carteiras quando T é grande, mas finito, e obteve uma distribuição aproximadamente lognormal. A seguir, ele calculou a esperança da utilidade potência sobre a lognormal. Entretanto, Merton e Samuelson (1974, p.77) dizem que isto não pode ser feito por causa de um intercâmbio impróprio do limite e da integral.

Para mais discussões sobre este tema, veja Merton e Samuelson (1974), Hakansson (1974), Samuelson e Merton (1974), Hakansson e Miller (1975) e Hakansson e Ziemba (1995). Devido à existência de uma salutar discordância entre estes distintos professores, o autor quer deixar clara sua opinião. O conceito de distribuição composta média é entendido como uma mudança na medida de probabilidade. A distribuição do caso multiperíodico deve ser uma transição suave da usada no caso uniperíodico para a usada (implicitamente) no caso com infinitos períodos pelo critério de Kelly. O cálculo do risco, do retorno esperado ou da esperança de uma função utilidade qualquer deve ser feito sobre a distribuição composta média. Estes conceitos passam a ter um significado ligeiramente diferente e, por isso, devem necessariamente receber a qualificação “na medida de probabilidade composta média”.

6.8. Distribuição composta média com longevidade

A distribuição composta média pode ser generalizada de forma a possibilitar um ajuste mesmo para $T = 1$. Este tipo de ajuste pode ser desejado pelo investidor se um ativo possuir uma densidade na qual ocorrem altos retornos com baixa probabilidade. Isto será mostrado com o exemplo a seguir.

Exemplo 6.8.1. Necessidade de ajuste da distribuição no caso uniperíodico. Considere que os ativos 1) e 2) tenham densidades uniperíodicas dos retornos dadas por $f_1(R) = (1 - \varepsilon)\delta(R - 1) + \varepsilon\delta(R - 1/\varepsilon)$, onde $\varepsilon = 10^{-6}$, e $f_2(R) = \delta(R - 1.8)$. Os gráficos destas duas densidades e de suas transformadas de Mellin foram traçados na figura a seguir.

Figura 6.8.1. Gráficos da densidade dos retornos de dois ativos e de suas transformadas de Mellin. Embora a esperança dos retornos do primeiro ativo seja maior, muitos investidores iriam escolher o segundo.

As médias dos retornos dos ativos 1) e 2) são, respectivamente, $2 - 10^{-6} \simeq 2$ e 1.8. Embora o primeiro ativo tenha uma média maior, muitas pessoas (incluindo o autor) irão preferir o segundo, pois consideram a probabilidade $\varepsilon = 10^{-6}$ muito pequena. Claramente, é necessário um ajuste.

■

No exemplo anterior, o investidor percebe intuitivamente que retornos muito altos (baixos) com baixa probabilidade forçam um aumento (uma diminuição) na média. Esta é uma situação na qual é desejável algum tipo de ajuste, que está relacionado com o desdém ou subrelevância do investidor aos resultados com baixa probabilidade. Este ajuste pode ser feito com uma função utilidade ou com uma mudança na distribuição de probabilidades. Neste último caso, a mudança deve ser tal que os retornos com probabilidade pequena sejam penalizados; que a taxa de crescimento do ativo seja preservada, de forma que se o horizonte de tempo T tender a infinito, não haja discordância entre o ajuste e o critério de Kelly; e que a esperança da distribuição ajustada seja côncava em relação aos pesos (não-negativos) dos ativos da carteira.

Pelo gráfico da transformada de Mellin do exemplo anterior, um possível critério seria escolher a carteira com maior $F(\hat{s})$, onde $\hat{s}$ é um valor entre 1 e 1.95, para $T = 1$, e que pode depender de parâmetros do investidor tais como o horizonte de tempo T , longevidade e a aversão a risco, incerteza, entropia ou probabilidades pequenas.

O ajuste pode ser feito com uma transformação na distribuição da variável aleatória, com o conceito de distribuição composta média com longevidade, que é a generalização do conceito apresentado na seção anterior.

Definição 6.8.1. Distribuição composta média com longevidade. *Dadas as variáveis aleatórias multiplicativas $X_0, X_1, \dots, X_{T-1}$, para $T \geq 1$, então $(X_0 \cdot X_1 \cdot \dots \cdot X_{T-1})^{1/(T+\tau)}$ segue a densidade composta média com longevidade $\varphi(x)$, onde τ é uma constante que representa a longevidade do investidor ou sua preferência à taxa de crescimento.*

Se $X_0, \dots, X_{T-1}$ são variáveis aleatórias independentes que seguem, respectivamente, as densidades $f_0(x), \dots, f_{T-1}(x)$, então

$$(X_0 \cdot X_1 \cdot \dots \cdot X_{T-1})^{\frac{1}{T+\tau}} \sim \varphi(x) = \left((f_0 \otimes f_1 \otimes \dots \otimes f_{T-1})^{\otimes \frac{1}{T+\tau}} \right)(x), \quad (6.12)$$

e se as densidades forem i.i.d., isto é, $f(x) = f_0(x) = f_1(x) = \dots = f_{T-1}(x)$, então

$$\underbrace{(X \cdot X \cdot \dots \cdot X)}_T^{\frac{1}{T+\tau}} = (X^{\times T})^{\frac{1}{T+\tau}} \sim \varphi(x) = \left(\underbrace{(f \otimes \dots \otimes f)}_T^{\otimes \frac{1}{T+\tau}} \right)(x) = \left((f^{\otimes T})^{\otimes \frac{1}{T+\tau}} \right)(x). \quad (6.13)$$

A densidade $\varphi(x)$ tem sentido financeiro para um investidor racional se $\tau \geq 0$ e $T \geq 1$. Mais tarde, será mostrado que estas são as condições que tornam a esperança na medida composta média côncava sobre os pesos dos ativos. Os termos “longevidade” e “preferência à taxa de crescimento” foram escolhidos porque à medida que τ aumenta, a taxa de crescimento da carteira escolhida pelo investidor tende para a taxa de crescimento da carteira de Kelly, também conhecida por carteira de crescimento ótimo ou log-ótimo. Esta carteira será vista no capítulo seguinte.

Matematicamente, se uma variável aleatória X é multiplicativa, i.i.d. e tal que

$$X \sim f(x) \xrightarrow{\mathcal{M}} F(s),$$

então

$$\left(X^{\times T} \right)^{\frac{1}{T+\tau}} \sim \varphi(x) = \left((f^{\otimes T})^{\otimes \frac{1}{T+\tau}} \right)(x) \xrightarrow{\mathcal{M}} \left(F \left(1 + \frac{s-1}{T+\tau} \right) \right)^T. \quad (6.14)$$

A carteira a ser escolhida é a com maior esperança ($s = 2$) da distribuição ajustada, isto é, com máximo

$$E_{\varphi}[X] = \left(F \left(1 + \frac{1}{T+\tau} \right) \right)^T = \left(\int_0^{\infty} x^{\frac{1}{T+\tau}} f(x) dx \right)^T.$$

Por exemplo, para $\tau = 1$ e $T = 1$, a carteira a ser escolhida seria a com maior $F(1.5)$. Isto resolve o problema do exemplo anterior.

A esperança da distribuição composta média pode ser considerada uma medida de recompensa, ou seja, uma generalização da esperança da distribuição.

6.8.1. Concavidade da esperança

Nesta seção, será provado que a esperança na medida de probabilidade composta média é côncava sobre os pesos não-negativos dos ativos. A concavidade é uma propriedade desejável para medidas de recompensa, de maneira análoga à propriedade da convexidade para medidas de risco.

Teorema 6.8.1. Concavidade da esperança na medida composta média. Sejam X e Y variáveis aleatórias aditivas e multiplicativas; um horizonte de tempo fixo T tal que $1 \leq T < \infty$ e um coeficiente de longevidade $\tau \geq 0$. Se os pesos são não-negativos, então a esperança na medida de probabilidade composta média é côncava sobre os pesos, isto é,

$$E_{\varphi}[uX + vY] \geq uE_{\varphi}[X] + vE_{\varphi}[Y],$$

onde $0 \leq u \leq 1$ e $v = 1 - u$.

Prova. Considere que $[X, Y]^T \sim p(x, y)$ e que

$$X \sim f(x) \xrightarrow{\mathcal{M}} F(s), Y \sim g(y) \xrightarrow{\mathcal{M}} G(s) \text{ e } Z = uX + vY \sim h(z) \xrightarrow{\mathcal{M}} H(s).$$

Então

$$E_{\varphi}[X] = F \left(1 + \frac{1}{T+\tau} \right)^T = \left(\int_0^{\infty} x^{\frac{1}{T+\tau}} f(x) dx \right)^T = \left(\int_0^{\infty} \int_0^{\infty} x^{\frac{1}{T+\tau}} p(x, y) dy dx \right)^T,$$

$$E_{\varphi}[Y] = G \left(1 + \frac{1}{T+\tau} \right)^T = \left(\int_0^{\infty} y^{\frac{1}{T+\tau}} g(y) dy \right)^T = \left(\int_0^{\infty} \int_0^{\infty} y^{\frac{1}{T+\tau}} p(x, y) dy dx \right)^T$$

e

$$E_{\varphi}[uX + vY] = H \left(1 + \frac{1}{T + \tau} \right)^T = \left(\int_0^{\infty} z^{\frac{1}{T+\tau}} h(z) dz \right)^T = \left(\int_0^{\infty} \int_0^{\infty} (ux + vy)^{\frac{1}{T+\tau}} p(x, y) dy dx \right)^T.$$

No gráfico da figura a seguir, a densidade $p(x, y)$ deve ser imaginada na terceira dimensão, para cima do papel. A densidade $p(x, y)$ é aproximada por infinitos paralelepípedos i , onde $1 \leq i \leq \infty$, cujas alturas h_i representam a probabilidade de as variáveis aleatórias assumirem valores dentro de sua base, dividida pela área da base.

Figura 6.8.2. Concavidade da esperança na medida composta média.

Na linha L_i da tabela seguinte, é mostrado como a altura h_i de cada paralelepípedo i contribui para $(E_{\varphi}[uX + vY])^{1/T}$, $(E_{\varphi}[X])^{1/T}$ e $(E_{\varphi}[Y])^{1/T}$. A linha L_S contém as somas dos paralelepípedos. Nesta linha, a coluna C_2 foi multiplicada por $u^{1/(T+\tau)}$, e a coluna C_3 , por $v^{1/(T+\tau)}$. Elevando as somas à potência $T + \tau$ e, então, aplicando a (segunda) desigualdade de Minkowski (3.111), a linha L_M é obtida. Simplificando-se, a linha L_N é obtida. Particionando o domínio com infinitos paralelepípedos de modo que as áreas de suas bases tendam a zero (veja Widder, 1989, p.198) e repetindo o raciocínio, a linha L_C é obtida.

Tabela 6.8.1. Concavidade da esperança na medida composta média.

	C_1		C_2		C_3
	$(E_{\varphi}[uX + vY])_i^{1/T}$		$(E_{\varphi}[X])_i^{1/T}$		$(E_{\varphi}[Y])_i^{1/T}$
L_i	$(ux_i + vy_i)^{1/(T+\tau)} h_i$		$(x_i)^{1/(T+\tau)} h_i$		$(y_i)^{1/(T+\tau)} h_i$
L_S	$\sum_i (ux_i h_i^{T+\tau} + vy_i h_i^{T+\tau})^{1/(T+\tau)}$		$\sum_i (ux_i h_i^{T+\tau})^{1/(T+\tau)}$		$\sum_i (vy_i h_i^{T+\tau})^{1/(T+\tau)}$
L_M	$\left(\sum_i (ux_i h_i^{T+\tau} + vy_i h_i^{T+\tau})^{1/(T+\tau)} \right)^{T+\tau}$	$\geq$	$\left(\sum_i (ux_i h_i^{T+\tau})^{1/(T+\tau)} \right)^{T+\tau}$	$+$	$\left(\sum_i (vy_i h_i^{T+\tau})^{1/(T+\tau)} \right)^{T+\tau}$
L_N	$\left(\sum_i (ux_i + vy_i)^{1/(T+\tau)} h_i \right)^{T+\tau}$	$\geq$	$u \left(\sum_i x_i^{1/(T+\tau)} h_i \right)^{T+\tau}$	$+$	$v \left(\sum_i y_i^{1/(T+\tau)} h_i \right)^{T+\tau}$
L_C	$(E_{\varphi}[uX + vY])^{\frac{T+\tau}{T}}$	$\geq$	$u (E_{\varphi}[X])^{\frac{T+\tau}{T}}$	$+$	$v (E_{\varphi}[Y])^{\frac{T+\tau}{T}}$

Portanto, a função

$$\left(E_{\varphi}[\mathbf{X}]\right)^{\frac{T+\tau}{T}}$$

é côncava sobre pesos não-negativos. Como a função

$$r(x) = x^{\frac{T}{T+\tau}}, \text{ para } 0 \leq \frac{T}{T+\tau} \leq 1,$$

é côncava e não-decrescente, então, pela propriedade (3.108), a função

$$E_{\varphi}[\mathbf{X}] = r\left(\left(E_{\varphi}[\mathbf{X}]\right)^{\frac{T+\tau}{T}}\right)$$

também é côncava.

□

6.8.2. Convexidade do *Downside Risk*

Para determinados parâmetros, a medida de risco *Downside Risk* é (muito provavelmente) convexa em relação aos pesos dos ativos, na medida de probabilidade composta média. Infelizmente, o autor não conseguiu provar formalmente esta proposição. Portanto, ela deve ser considerada uma conjectura, que será enunciada a seguir.

Conjectura 6.8.1. Convexidade do *Downside Risk* na medida de probabilidade composta média. Sejam $\mathbf{X}_0$ e $\mathbf{Y}_0$ variáveis aleatórias que tenham as operações de adição e multiplicação bem definidas; um horizonte de tempo fixo T tal que $1 \leq T < \infty$ e um coeficiente de longevidade $\tau \geq 0$. Se os pesos são não-negativos, então o *Downside Risk* com parâmetros $\alpha \geq 1$ e R_B fixo, na medida de probabilidade composta média, é convexa sobre os pesos, isto é,

$$DR_{\alpha, R_B} \left[\left((u\mathbf{X}_0 + v\mathbf{Y}_0) \cdot \mathbf{Z}_T \right)^{\frac{1}{T+\tau}} \right] \leq u DR_{\alpha, R_B} \left[\left(\mathbf{X}_0 \cdot \mathbf{Z}_T \right)^{\frac{1}{T+\tau}} \right] + v DR_{\alpha, R_B} \left[\left(\mathbf{Y}_0 \cdot \mathbf{Z}_T \right)^{\frac{1}{T+\tau}} \right], \quad (6.15)$$

onde $0 \leq u \leq 1$, $v = 1 - u$ e $\mathbf{Z}_T = \mathbf{Z}_1 \cdot \dots \cdot \mathbf{Z}_{T-1}$ é um produto de variáveis aleatórias.

□

Esta conjectura foi testada numericamente com experimentos Monte Carlo usando-se misturas de Deltas de Dirac e diferentes valores para os parâmetros. A implementação em Mathematica está no **Apêndice A**.

A concavidade da esperança e a convexidade do *Downside Risk* na medida de probabilidade composta média implicam que a fronteira eficiente do gráfico retorno esperado *versus Downside Risk* é côncava; veja Yu (1985, p.26, 181).

6.9. Rebalanceamento calculado por programação dinâmica

Como o horizonte de tempo tem T períodos e é possível rebalancear a carteira no início de cada um deles, então uma análise multiperíodica deve gerar seqüências ordenadas de T carteiras; uma para cada período. As carteiras de uma seqüência não precisam ser iguais.

Isto significa que cada ponto da fronteira eficiente multiperíodica deve representar uma ou mais seqüências ordenadas de carteiras uniperiódicas. Cada carteira de uma dada seqüência possui uma distribuição uniperiódica determinada pelos pesos dos ativos. Aplicando a convolução multiplicativa, obtém-se a distribuição multiperiódica da seqüência de carteiras, que deve ser ajustada para a distribuição composta média, de onde o retorno esperado e o risco serão calculados. Variando-se os pesos dos ativos nas carteiras de cada seqüência, uma fronteira eficiente emergirá como no caso uniperiódico, e o investidor poderá escolher a carteira com seu nível de retorno desejado ou seu nível de risco tolerado.

A solução clássica para este tipo de problema é por programação dinâmica, resolvendo subproblemas de trás para frente, ou do futuro para o presente. Suponha que o investidor fixe um nível de retorno composto médio $R_{E,t}$ desejado, para cada instante $t \in \{0, 1, \dots, T-1\}$. Estes retornos não precisam ser iguais, e é desejável que, no limite quando T tende a infinito, haja convergência de $R_{E,0}$ para o retorno da carteira de Kelly.

No último período, $t = T-1$, e o problema é uniperiódico. O retorno $R_{E,t} = R_{E,T-1}$ se reduz ao retorno discreto, e os pesos dos ativos devem ser calculados de forma que o risco seja minimizado, como no capítulo anterior.

No penúltimo período, o problema torna-se a determinação apenas dos pesos dos ativos da carteira do período $t = T-2$, pois a carteira do período seguinte já foi calculada. Como no caso uniperiódico, os pesos devem ser calculados iterativamente, minimizando-se o risco, com a restrição do retorno composto médio. É necessário fazer a convolução multiplicativa das distribuições dos dois períodos, elevar o resultado à potência $1/2$ e, então, calcular seu retorno composto médio, que deve ser maior ou igual a $R_{E,t} = R_{E,T-2}$.

Para os períodos $t = T-3, T-4, \dots, 0$, o procedimento é análogo. Note que, em geral, o retorno uniperiódico da carteira em um instante $t \neq T-1$ será diferente de $R_{E,t}$.

6.10. Fronteira eficiente

Neste trabalho, será usada a distribuição composta média com $\mathcal{T} = 0$, isto é, sem longevidade. Para a seqüência desejada de retornos compostos médios $R_{E,t}$, para $0 \leq t < T$, o seguinte problema de otimização pode ser formulado.

Problema de otimização 6.10.1. Fronteira eficiente multiperíodica. Existem T períodos e D ativos dependentes $[\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_D]^T$ que seguem uma distribuição multivariada de retornos i.i.d. ao longo do tempo. Os pesos $w_{t,d}$, para $0 \leq t < T$, $1 \leq d \leq D$, dos ativos nas carteiras dos T períodos são as variáveis de otimização e devem ser não-negativos. A medida de risco $\mathcal{R}$ e os retornos compostos médios desejados $R_{E,t}$ são dados, para $0 \leq t < T$. As carteiras são otimizadas do futuro para o presente, e são representadas por

$$\mathbf{R}_{P,t} = \left(\sum_{d=1}^D w_{t,d} \mathbf{R}_d \right) \sim f_t(\mathbf{R}; w_{t,1}, \dots, w_{t,D}) \xrightarrow{\mathcal{M}} F_t(s), \quad 0 \leq t < T.$$

Após serem otimizadas, elas se tornam

$$\mathbf{R}_{P,t}^* = \left(\sum_{d=1}^D w_{t,d}^* \mathbf{R}_d \right) \sim f_t^* \left(R; w_{t,1}^*, \dots, w_{t,D}^* \right) \xrightarrow{\mathcal{M}} F_t^* (s), \quad 0 \leq t < T.$$

O problema de otimização se torna a obtenção dos pesos $w_{t,d}^*$ dados por

$$w_{t,d}^* = \arg \min_{w_{t,d}} \mathcal{R} \left[\left(\mathbf{R}_{P,t} \cdot \mathbf{R}_{P,t+1}^* \cdot \dots \cdot \mathbf{R}_{P,T-1}^* \right)^{\frac{1}{T-t}} \right] = \mathcal{R} \left[\left(f_t \otimes f_{t+1}^* \otimes \dots \otimes f_{T-1}^* \right)^{\otimes \frac{1}{T-t}} (R) \right], \quad t = \{T-1, \dots, 0\}, \quad 1 \leq d \leq D$$

s.t.

$$\begin{aligned} E \left[\left(\mathbf{R}_{P,t} \cdot \mathbf{R}_{P,t+1}^* \cdot \dots \cdot \mathbf{R}_{P,T-1}^* \right)^{\frac{1}{T-t}} \right] &= F_t \left(1 + \frac{1}{T-t} \right) F_{t+1}^* \left(1 + \frac{1}{T-t} \right) \dots F_{T-1}^* \left(1 + \frac{1}{T-t} \right) \geq R_{E,t}, \quad 0 \leq t < T \\ w_{t,d} &\geq 0, \quad 0 \leq t < T, \quad 1 \leq d \leq D \\ \sum_{d=1}^D w_{t,d} &= 1, \quad 0 \leq t < T. \end{aligned}$$

6.10.1. Distribuição multivariada dos retornos dos ativos

A densidade multivariada uniperiódica dos retornos dos ativos é aproximada por um produto tensorial finito de Erlangs; veja a **Seção 5.4.1**.

$$[\mathbf{R}_1, \dots, \mathbf{R}_D]^T \sim m(R_1, \dots, R_D) = \sum_{i=1}^n p_i \gamma_{b_{i,1}}^{a_{i,1}}(R_1) \dots \gamma_{b_{i,D}}^{a_{i,D}}(R_D) \geq 0, \quad \sum_{i=1}^n p_i = 1, \quad 1 \leq n < \infty. \quad (6.16)$$

Por hipótese, esta densidade será considerada i.i.d. nos períodos futuros.

6.10.2. Distribuição dos retornos da carteira

A carteira $\mathbf{R}_{P,t}$ de cada período t , é determinada pelos pesos não-negativos (por hipótese) dos ativos, isto é, $w_{t,d} \geq 0$, para $0 \leq t < T$ e $1 \leq d \leq D$. Como já foi explicado na **Seção 5.4.2**, a densidade da carteira é uma mistura finita de Erlangs e é conveniente expressá-la sem componentes Erlang repetidas, isto é,

$$\begin{aligned} \mathbf{R}_{P,t} = \left(\sum_{d=1}^D w_{t,d} \mathbf{R}_d \right) &\sim f_t(R; w_{t,1}, \dots, w_{t,D}) = \sum_{i=1}^n p_i \left(\gamma_{b_{i,1}}^{a_{i,1}/w_{t,1}} \oplus \dots \oplus \gamma_{b_{i,D}}^{a_{i,D}/w_{t,D}} \right) (R) = \\ &= \sum_{d'_i=1}^{D'_i} \sum_{r_i=1}^{b'_{\max,t,d'_i}} k_{t,r_i,d'_i} \gamma_{r_i}^{a'_{d'_i}/w'_{t,d'_i}} (R), \end{aligned} \quad (6.17)$$

onde

$$D'_i = \text{número de componentes distintos do vetor} \left[\frac{a_1}{w_{t,1}}, \dots, \frac{a_{d_i}}{w_{t,d_i}}, \dots, \frac{a_D}{w_{t,D}} \right]^T, \quad (6.18)$$

$$\left[\frac{a'_{t,1}}{w'_{t,1}}, \dots, \frac{a'_{t,d'_i}}{w'_{t,d'_i}}, \dots, \frac{a'_{t,D'}}{w'_{t,D'}} \right]^T = \text{vetor} \left[\frac{a_1}{w_{t,1}}, \dots, \frac{a_{d_i}}{w_{t,d_i}}, \dots, \frac{a_D}{w_{t,D}} \right]^T \text{ sem componentes repetidos}, \quad (6.19)$$

$$b'_{t,i,d'_i} = \sum_{d=1}^D b_{i,d}, \quad 1 \leq i \leq n, \quad 1 \leq d'_i \leq D'_i, \quad (6.20)$$

$$\frac{a_d}{w_{t,d}} = \frac{a'_{d'_i}}{w'_{t,d'_i}}$$

$$b'_{\max,t,d'_i} = \max_{1 \leq i \leq n} b'_{t,i,d'_i}, \quad 1 \leq d'_i \leq D'_i, \quad (6.21)$$

$$k_{t,r_i,d'_i} = \sum_{i=1}^n p_i \left\{ \prod_{\substack{j=1 \\ j \neq d'_i}}^{D'_i} \left(1 - \frac{a'_{t,d'_i} w'_{t,j}}{w'_{t,d'_i} a'_{t,j}} \right)^{-b'_{t,i,j}} \right\} \left(\frac{a'_{t,d'_i}}{w'_{t,d'_i}} \right)^{b'_{t,i,d'_i} - r_i} M_{t,i,d'_i}^{b'_{t,i,d'_i} - r_i}, \quad 1 \leq r_i \leq b'_{\max,t,d'_i}, \quad 1 \leq d'_i \leq D'_i, \quad (6.22)$$

$$M_{t,i,d'_i}^u = \begin{cases} 1 & , \quad u = 0 \\ \frac{1}{u} \sum_{v=0}^{u-1} \left[\sum_{\substack{j=1 \\ j \neq d'_i}}^{D'_i} b'_{t,i,j} \left(\frac{a'_{t,d'_i}}{w'_{t,d'_i}} - \frac{a'_{t,j}}{w'_{t,j}} \right)^{v-u} \right] M_{t,i,d'_i}^v & , \quad u > 0 \end{cases} \quad (6.23)$$

e

$$0 \leq t < T. \quad (6.24)$$

6.10.3. Distribuição composta média

Neste trabalho, as densidades uniperiódicas das carteiras são representadas por misturas finitas de densidades Erlang. A densidade do retorno acumulado de dois períodos é obtida pela convolução multiplicativa das densidades de cada período. O resultado é uma mistura finita de densidades Meijer G, ou mais precisamente, de um caso particular desta função. Veja o **Teorema 3.12.1**.

Antes do cálculo do risco e do retorno, neste trabalho será feita a mudança da distribuição dos T períodos para a composta média, pela elevação da distribuição à potência $1/T$. Portanto, a densidade que deve ser usada em um dado instante t , onde $0 \leq t < T$, é dada por

$$\begin{aligned} (\mathbf{R}_{P,t} \cdots \mathbf{R}_{P,T-1})^{\frac{1}{T-t}} &\sim (f_t \otimes \dots \otimes f_{T-1})^{\otimes \frac{1}{T-t}}(R) = \\ &= \left(\left[\sum_{d'_i=1}^{D'_i} \sum_{r_i=1}^{b'_{\max,t,d'_i}} k_{t,r_i,d'_i} \gamma_{r_i}^{a'_{t,d'_i}/w'_{t,d'_i}} \right] \otimes \dots \otimes \left[\sum_{d'_{T-1}=1}^{D'_{T-1}} \sum_{r_{T-1}=1}^{b'_{\max,T-1,d'_{T-1}}} k_{T-1,r_{T-1},d'_{T-1}} \gamma_{r_{T-1}}^{a'_{T-1,d'_{T-1}}/w'_{T-1,d'_{T-1}}} \right] \right)^{\otimes \frac{1}{T-t}}(R) = \\ &= \sum_{d'_i=1}^{D'_i} \sum_{r_i=1}^{b'_{\max,t,d'_i}} \dots \sum_{d'_{T-1}=1}^{D'_{T-1}} \sum_{r_{T-1}=1}^{b'_{\max,T-1,d'_{T-1}}} k_{t,r_i,d'_i} \dots k_{T-1,r_{T-1},d'_{T-1}} \cdot \left(\gamma_{r_i}^{a'_{t,d'_i}/w'_{t,d'_i}} \otimes \dots \otimes \gamma_{r_{T-1}}^{a'_{T-1,d'_{T-1}}/w'_{T-1,d'_{T-1}}} \right)^{\otimes \frac{1}{T-t}}(R) = \\ &= \sum_{d'_i=1}^{D'_i} \sum_{r_i=1}^{b'_{\max,t,d'_i}} \dots \sum_{d'_{T-1}=1}^{D'_{T-1}} \sum_{r_{T-1}=1}^{b'_{\max,T-1,d'_{T-1}}} k_{t,r_i,d'_i} \dots k_{T-1,r_{T-1},d'_{T-1}} \cdot \\ &\quad \left(\frac{a'_{t,d'_i} \dots a'_{T-1,d'_{T-1}}}{\Gamma(r_i) \dots \Gamma(r_{T-1}) w'_{t,d'_i} \dots w'_{T-1,d'_{T-1}}} G_{0,T-t}^{T-t,0} \left(\frac{a'_{t,d'_i} \dots a'_{T-1,d'_{T-1}} R}{w'_{t,d'_i} \dots w'_{T-1,d'_{T-1}} \mid r_i - 1, \dots, r_{T-1} - 1} \right) \right)^{\otimes \frac{1}{T-t}} = \\ &= \sum_{d'_i=1}^{D'_i} \sum_{r_i=1}^{b'_{\max,t,d'_i}} \dots \sum_{d'_{T-1}=1}^{D'_{T-1}} \sum_{r_{T-1}=1}^{b'_{\max,T-1,d'_{T-1}}} \frac{k_{t,r_i,d'_i} \dots k_{T-1,r_{T-1},d'_{T-1}} a'_{t,d'_i} \dots a'_{T-1,d'_{T-1}}}{\Gamma(r_i) \dots \Gamma(r_{T-1}) w'_{t,d'_i} \dots w'_{T-1,d'_{T-1}}} \cdot \\ &\quad (T-t) R^{T-t-1} G_{0,T-t}^{T-t,0} \left(\frac{a'_{t,d'_i} \dots a'_{T-1,d'_{T-1}} R^{T-t}}{w'_{t,d'_i} \dots w'_{T-1,d'_{T-1}} \mid r_i - 1, \dots, r_{T-1} - 1} \right), \quad 0 \leq t < T. \end{aligned} \quad (6.25)$$

6.10.4. Valor esperado

O valor esperado no instante t é obtido fazendo $s = 2$ na transformada de Mellin.

$$\begin{aligned} E \left[\left(\mathbf{R}_{P,t} \cdot \dots \cdot \mathbf{R}_{P,T-1} \right)^{\frac{1}{T-t}} \right] &= F_t \left(1 + \frac{1}{T-t} \right) \dots F_{T-1} \left(1 + \frac{1}{T-t} \right) = \\ &= \left[\sum_{d'=1}^{D'_t} \sum_{r=1}^{b'_{\max,t,d'}} \frac{k_{t,r,d'} \Gamma \left(r + \frac{1}{T-t} \right)}{\left(\frac{a'_{t,d'}}{w'_{t,d'}} \right)^{\frac{1}{T-t}} \Gamma(r)} \right] \dots \left[\sum_{d'=1}^{D'_{T-1}} \sum_{r=1}^{b'_{\max,T-1,d'}} \frac{k_{T-1,r,d'} \Gamma \left(r + \frac{1}{T-t} \right)}{\left(\frac{a'_{T-1,d'}}{w'_{T-1,d'}} \right)^{\frac{1}{T-t}} \Gamma(r)} \right], \quad 0 \leq t < T \end{aligned} \quad (6.26)$$

6.10.5. Downside Risk

A potência α do *Downside Risk* foi calculada no **Corolário 3.15.6** e requer a avaliação numérica de um caso particular da função Meijer G.

$$\begin{aligned} DR_{\alpha, R_B}^\alpha \left[\left(\mathbf{R}_{P,t} \cdot \dots \cdot \mathbf{R}_{P,T-1} \right)^{\frac{1}{T-t}} \right] &= DR_{\alpha, R_B}^\alpha \left(\left(f_t \otimes \dots \otimes f_{T-1} \right)^{\otimes \frac{1}{T-t}} (R) \right) = \\ &= DR_{\alpha, R_B}^\alpha \left(\left(\left[\sum_{d'_i=1}^{D'_i} \sum_{r_i=1}^{b'_{\max,t,d'_i}} k_{t,r_i,d'_i} \gamma_{r_i}^{\frac{a'_{t,d'_i}}{w'_{t,d'_i}}} \right] \otimes \dots \otimes \left[\sum_{d'_{T-1}=1}^{D'_{T-1}} \sum_{r_{T-1}=1}^{b'_{\max,T-1,d'_{T-1}}} k_{T-1,r_{T-1},d'_{T-1}} \gamma_{r_{T-1}}^{\frac{a'_{T-1,d'_{T-1}}}{w'_{T-1,d'_{T-1}}}} \right] \right)^{\otimes \frac{1}{T-t}} (R) \right) = \\ &= DR_{\alpha, R_B}^\alpha \left(\sum_{d'_i=1}^{D'_i} \sum_{r_i=1}^{b'_{\max,t,d'_i}} \dots \sum_{d'_{T-1}=1}^{D'_{T-1}} \sum_{r_{T-1}=1}^{b'_{\max,T-1,d'_{T-1}}} k_{t,r_i,d'_i} \dots k_{T-1,r_{T-1},d'_{T-1}} \left(\gamma_{r_i}^{\frac{a'_{t,d'_i}}{w'_{t,d'_i}}} \otimes \dots \otimes \gamma_{r_{T-1}}^{\frac{a'_{T-1,d'_{T-1}}}{w'_{T-1,d'_{T-1}}}} \right)^{\otimes \frac{1}{T-t}} (R) \right) = \\ &= \sum_{d'_i=1}^{D'_i} \sum_{r_i=1}^{b'_{\max,t,d'_i}} \dots \sum_{d'_{T-1}=1}^{D'_{T-1}} \sum_{r_{T-1}=1}^{b'_{\max,T-1,d'_{T-1}}} k_{t,r_i,d'_i} \dots k_{T-1,r_{T-1},d'_{T-1}} DR_{\alpha, R_B}^\alpha \left(\left(\gamma_{r_i}^{\frac{a'_{t,d'_i}}{w'_{t,d'_i}}} \otimes \dots \otimes \gamma_{r_{T-1}}^{\frac{a'_{T-1,d'_{T-1}}}{w'_{T-1,d'_{T-1}}}} \right)^{\otimes \frac{1}{T-t}} (R) \right) = \\ &= \sum_{d'_i=1}^{D'_i} \sum_{r_i=1}^{b'_{\max,t,d'_i}} \dots \sum_{d'_{T-1}=1}^{D'_{T-1}} \sum_{r_{T-1}=1}^{b'_{\max,T-1,d'_{T-1}}} k_{t,r_i,d'_i} \dots k_{T-1,r_{T-1},d'_{T-1}} \frac{\Gamma(1+\alpha) \left(\frac{R_B}{T-t} \right)^\alpha}{\Gamma(r_i) \dots \Gamma(r_{T-1})} \cdot \\ &\quad G_{T-t, 2(T-t)}^{T-t, T-t} \left(\frac{a'_{t,d'_i} \dots a'_{T-1,d'_{T-1}} R_B^{T-t}}{w'_{t,d'_i} \dots w'_{T-1,d'_{T-1}}} \left| \begin{array}{c} \frac{1}{T-t}, \frac{2}{T-t}, \dots, \frac{T-t}{T-t} \\ -\alpha, \frac{1-\alpha}{T-t}, \dots, \frac{T-t-1-\alpha}{T-t} \end{array} \right. \right) = \\ &= \Gamma(1+\alpha) \left(\frac{R_B}{T-t} \right)^\alpha \sum_{d'_i=1}^{D'_i} \sum_{r_i=1}^{b'_{\max,t,d'_i}} \dots \sum_{d'_{T-1}=1}^{D'_{T-1}} \sum_{r_{T-1}=1}^{b'_{\max,T-1,d'_{T-1}}} \frac{k_{t,r_i,d'_i} \dots k_{T-1,r_{T-1},d'_{T-1}}}{\Gamma(r_i) \dots \Gamma(r_{T-1})} \cdot \\ &\quad G_{T-t, 2(T-t)}^{T-t, T-t} \left(\frac{a'_{t,d'_i} \dots a'_{T-1,d'_{T-1}} R_B^{T-t}}{w'_{t,d'_i} \dots w'_{T-1,d'_{T-1}}} \left| \begin{array}{c} \frac{1}{T-t}, \frac{2}{T-t}, \dots, \frac{T-t}{T-t} \\ -\alpha, \frac{1-\alpha}{T-t}, \dots, \frac{T-t-1-\alpha}{T-t} \end{array} \right. \right), \quad 0 \leq t < T. \end{aligned} \quad (6.27)$$

6.10.6. Formulação do problema de otimização

O problema de otimização deve ser formulado de modo que as carteiras sejam calculadas do futuro

para o presente, usando-se as fórmulas anteriores do retorno e do risco.

Problema de otimização 6.10.2. Fronteira eficiente multiperódica calculada com produto tensorial finito de Erlangs, *Downside Risk* e distribuição composta média. A medida de risco é o *Downside Risk* DR_{α, R_B} . O número de períodos é $T \in \{1, 2, \dots\}$. Os parâmetros α , R_B e os retornos compostos médios $R_{E,t}$ desejados são dados, para $0 \leq t < T$. Existem D ativos dependentes com uma distribuição multivariada de retornos $\mathbf{R}_1, \dots, \mathbf{R}_D$ i.i.d. ao longo do tempo representada por um produto tensorial finito de Erlangs.

$$[\mathbf{R}_1, \dots, \mathbf{R}_D]^\top \sim m(R_1, \dots, R_D) = \sum_{i=1}^n p_i \gamma_{b_{i,1}}^{a_{i,1}}(R_1) \dots \gamma_{b_{i,D}}^{a_{i,D}}(R_D) \geq 0, \sum_{i=1}^n p_i = 1, 1 \leq n < \infty.$$

Os pesos $w_{t,d}$, para $0 \leq t < T$ e $1 \leq d \leq D$, dos ativos nas carteiras dos T períodos são as variáveis de otimização e devem ser não-negativos. As carteiras devem ser otimizadas do futuro para o presente e são representadas por

$$\begin{aligned} \mathbf{R}_{P,t} &= \left(\sum_{d=1}^D w_{t,d} \mathbf{R}_d \right) \sim f_t(R; w_{t,1}, \dots, w_{t,D}) = \sum_{i=1}^n p_i \left(\gamma_{b_{i,1}}^{a_{i,1}/w_{t,1}} \oplus \dots \oplus \gamma_{b_{i,D}}^{a_{i,D}/w_{t,D}} \right)(R) = \\ &= \sum_{d'_i=1}^{D'_i} \sum_{r_i=1}^{b'_{\max,t,d'_i}} k_{t,r_i,d'_i} \gamma_{r_i}^{a'_{t,d'_i}/w'_{t,d'_i}}(R) \xrightarrow{\mathcal{M}} F_t(s) = \sum_{d'_i=1}^{D'_i} \sum_{r_i=1}^{b'_{\max,t,d'_i}} k_{t,r_i,d'_i} \frac{\Gamma(r_i - 1 + s)}{\left(\frac{a'_{t,d'_i}}{w'_{t,d'_i}} \right)^{s-1} \Gamma(r_i)}, \quad 0 \leq t < T. \end{aligned} \quad (6.28)$$

Após serem otimizadas, elas se tornam

$$\begin{aligned} \mathbf{R}_{P,t}^* &= \left(\sum_{d=1}^D w_{t,d}^* \mathbf{R}_d \right) \sim f_t^*(R; w_{t,1}^*, \dots, w_{t,D}^*) = \sum_{i=1}^n p_i \left(\gamma_{b_{i,1}}^{a_{i,1}/w_{t,1}^*} \oplus \dots \oplus \gamma_{b_{i,D}}^{a_{i,D}/w_{t,D}^*} \right)(R) = \\ &= \sum_{d'_i=1}^{D'_i} \sum_{r_i=1}^{b'^*_{\max,t,d'_i}} k_{t,r_i,d'_i}^* \gamma_{r_i}^{a'^*_{t,d'_i}/w'^*_{t,d'_i}}(R) \xrightarrow{\mathcal{M}} F_t^*(s) = \sum_{d'_i=1}^{D'_i} \sum_{r_i=1}^{b'^*_{\max,t,d'_i}} k_{t,r_i,d'_i}^* \frac{\Gamma(r_i - 1 + s)}{\left(\frac{a'^*_{t,d'_i}}{w'^*_{t,d'_i}} \right)^{s-1} \Gamma(r_i)}, \quad 0 \leq t < T. \end{aligned} \quad (6.29)$$

O problema de otimização se torna a obtenção dos pesos $w_{t,d}^*$ dados por

$$\begin{aligned} w_{t,d}^* &= \arg \min_{w_{t,d}} DR_{\alpha, R_B}^\alpha \left[\left(\mathbf{R}_{P,t} \cdot \mathbf{R}_{P,t+1}^* \cdot \dots \cdot \mathbf{R}_{P,T-1}^* \right)^{\frac{1}{T-t}} \right] = DR_{\alpha, R_B}^\alpha \left[\left(f_t \otimes f_{t+1}^* \otimes \dots \otimes f_{T-1}^* \right)^{\otimes \frac{1}{T-t}}(R) \right] = \\ &= \Gamma(1 + \alpha) \left(\frac{R_B}{T-t} \right)^\alpha \sum_{d'_i=1}^{D'_i} \sum_{r_i=1}^{b'_{\max,t,d'_i}} \sum_{d'_{i+1}=1}^{D'_{i+1}} \sum_{r_{i+1}=1}^{b'^*_{\max,t+1,d'_{i+1}}} \dots \sum_{d'_{T-1}=1}^{D'_{T-1}} \sum_{r_{T-1}=1}^{b'^*_{\max,T-1,d'_{T-1}}} \frac{k_{t,r_i,d'_i} k_{t+1,r_{i+1},d'_{i+1}}^* \dots k_{T-1,r_{T-1},d'_{T-1}}^*}{\Gamma(r_i) \Gamma(r_{i+1}) \dots \Gamma(r_{T-1})} \cdot \\ &\cdot G_{T-t, 2(T-t)}^{T-t, T-t} \left(\frac{a'_{t,d'_i} a'^*_{t+1,d'_{i+1}} \dots a'^*_{T-1,d'_{T-1}} R_B^{T-t}}{w'_{t,d'_i} w'^*_{t+1,d'_{i+1}} \dots w'^*_{T-1,d'_{T-1}}} \middle| r_i, \dots, r_{T-1}; \frac{1}{T-t}, \frac{2}{T-t}, \dots, \frac{T-t}{T-t}, \frac{-\alpha}{T-t}, \frac{1-\alpha}{T-t}, \dots, \frac{T-t-1-\alpha}{T-t} \right), \quad t = \{T-1, \dots, 0\}, \\ &\quad 1 \leq d \leq D \end{aligned}$$

s.t.

$$\begin{aligned}
E \left[\left(\mathbf{R}_{P,t} \cdot \mathbf{R}_{P,t+1}^* \cdot \dots \cdot \mathbf{R}_{P,T-1}^* \right)^{\frac{1}{T-t}} \right] &= F_t \left(1 + \frac{1}{T-t} \right) F_{t+1}^* \left(1 + \frac{1}{T-t} \right) \dots F_{T-1}^* \left(1 + \frac{1}{T-t} \right) = \\
&= \left[\sum_{d'=1}^{D'} \sum_{r=1}^{b'_{\max,t,d'}} \frac{k_{t,r,d'} \Gamma \left(r + \frac{1}{T-t} \right)}{\left(\frac{a'_{t,d'}}{w'_{t,d'}} \right)^{\frac{1}{T-t}} \Gamma(r)} \right] \left[\sum_{d'=1}^{D'_{t+1}} \sum_{r=1}^{b'^*_{\max,t+1,d'}} \frac{k_{t+1,r,d'}^* \Gamma \left(r + \frac{1}{T-t} \right)}{\left(\frac{a'^*_{t+1,d'}}{w'^*_{t+1,d'}} \right)^{\frac{1}{T-t}} \Gamma(r)} \right] \dots \left[\sum_{d'=1}^{D'_{T-1}} \sum_{r=1}^{b'^*_{\max,T-1,d'}} \frac{k_{T-1,r,d'}^* \Gamma \left(r + \frac{1}{T-t} \right)}{\left(\frac{a'^*_{T-1,d'}}{w'^*_{T-1,d'}} \right)^{\frac{1}{T-t}} \Gamma(r)} \right] \geq R_{E,t}, \quad 0 \leq t < T \\
w_{t,d} &\geq 0, \quad 0 \leq t < T, \quad 1 \leq d \leq D \\
\sum_{d=1}^D w_{t,d} &= 1, \quad 0 \leq t < T.
\end{aligned}$$

6.11. Uma estratégia dinâmica

No fim de cada período, uma nova amostra com os retornos dos ativos é observada e, por isso, é necessário recalcular uma nova estimativa para a densidade multivariada e uma nova fronteira eficiente para os períodos restantes. Se a nova estimativa for exatamente igual à antiga, então não é necessário recalcular as carteiras.

6.12. Conclusões

A variável retorno discreto é do tipo aditivo a cada instante de tempo e do tipo multiplicativo ao longo do tempo. Esta é uma constatação crucial em uma análise multiperódica. Os livros didáticos em Finanças não chamam a atenção para este fato.

O mundo contínuo ou aditivo é um homeomorfismo do mundo discreto ou multiplicativo. Os retornos discretos são mapeados nos retornos contínuos, e a operação de multiplicação do mundo discreto se torna a operação de adição do mundo contínuo. Os retornos discretos podem ser levados para o mundo contínuo, onde operações matemáticas mais simples podem ser realizadas, e o resultado deve ser trazido de volta para o mundo discreto. Esta técnica pode ser usada com as densidades dos retornos em tempo discreto e em tempo contínuo. Os pesos dos ativos nas carteiras não são mapeáveis diretamente para o mundo contínuo. Isto significa que quem deve ser levado para o mundo contínuo são os retornos discretos da carteira, e não os dos ativos componentes.

Se o intervalo entre dois períodos tender a zero (tempo contínuo), então, em uma análise matemática mais rigorosa, é preferível realizar todos os cálculos no mundo discreto e, então, aplicar o limite.

Se o horizonte de tempo T do investidor é maior que 1, então a carteira com maior média uniperiódica não é necessariamente a que deve ser escolhida.

A esperança e a variância de variáveis aleatórias multiplicativas e i.i.d. crescem, em geral, não-linearmente em função do tempo (discreto e contínuo). Muitos autores não diferenciam apropriadamente as variáveis aleatórias multiplicativas das aditivas.

Foram discutidas duas possíveis escolhas para a densidade de probabilidade a ser usada no caso

multiperiódico: a densidade acumulada no instante final e a densidade composta média. A distribuição composta média é obtida multiplicando-se T variáveis aleatórias e elevando o resultado à potência $1/T$. Experimentos Monte Carlo indicam que a utilização da distribuição composta média é preferível ao da distribuição no instante final. Esta distribuição é uma mudança na medida de probabilidade e possibilita uma transição suave da distribuição usada no caso uniperiódico para a usada (implicitamente) no caso com infinitos períodos. Ela também evita a necessidade de ajuste na medida de risco.

A distribuição composta média foi generalizada para a composta média com longevidade, de forma a contemplar casos nos quais há altos retornos com baixa probabilidade, mesmo no caso uniperiódico. Na medida de probabilidade composta média com longevidade, a esperança é côncava e o *Downside Risk* é (muito provavelmente) convexo sobre os pesos dos ativos nas carteiras, para o caso particular no qual $\alpha \geq 1$, R_B é constante, $\tau \geq 0$, $1 \leq T < \infty$ e $w_{t,d} \geq 0$, para todos os t e d tais que $1 \leq t < T$ e $1 \leq d \leq D$. Isto implica que a fronteira eficiente do gráfico retorno esperado *versus* *Downside Risk* é côncava.

O problema de otimização multiperiódico foi formulado com o conceito de distribuição composta média. As carteiras a serem escolhidas formam uma seqüência ordenada e devem ser calculadas do futuro para o presente, com a técnica de programação dinâmica. O número de variáveis de otimização em cada estágio é igual ao número de ativos, e o problema de otimização é convexo.

Se a densidade multivariada dos retornos discretos uniperiódicos dos ativos dependentes for representada por um produto tensorial finito de Erlangs, então a densidade uniperiódica dos retornos discretos de qualquer carteira não-alavancada será uma mistura finita de Erlangs. É possível usar a convolução multiplicativa por causa da hipótese i.i.d. A convolução multiplicativa de misturas finitas de Erlangs gera uma mistura finita de um caso especial da distribuição Meijer G. Esta última é fechada em relação à potenciação.

Na medida de probabilidade composta média, o cálculo da esperança requer a avaliação numérica da função Gama; e o cálculo do *Downside Risk* requer a avaliação numérica de um caso particular da função Meijer G, conhecido por caso logarítmico, mas com pólos no numerador e no denominador separados. A quantidade de funções Meijer G e sua dificuldade de avaliação crescem muito rapidamente em função do horizonte de tempo e do número de ativos.

No fim de cada período, uma nova observação com os retornos dos ativos é obtida e, com ela, deve-se fazer um novo recálculo completo para os períodos restantes.

Uma recomendação popular de pessoas ligadas a Finanças é que os jovens devem ter carteiras mais arriscadas do que os idosos. Considerando-se que o horizonte de tempo do jovem é maior que o do idoso, então as fronteiras eficientes de ambos serão diferentes, mesmo que os parâmetros de aversão ao risco sejam iguais. Em geral, as carteiras eficientes de um não estarão na fronteira eficiente do outro. O horizonte de tempo da pessoa determina parcialmente a carteira a ser escolhida e explica parcialmente essa recomendação popular.

7. OTIMIZAÇÃO DE CARTEIRAS COM INFINITOS PERÍODOS

7.1. Introdução

Na **Seção 6.6**, foi visto que, para ativos com retorno esperado maior que um e variância maior que zero, a esperança e a variância crescem não-linearmente com o número de períodos considerados. Para evitar a divergência desses valores quando o número de períodos tende a infinito, é necessário usar uma distribuição composta média, que foi definida no capítulo anterior.

Neste capítulo, será mostrado que, se o horizonte de tempo T do investidor tende a infinito, então os conceitos apresentados resultarão no bem conhecido critério de Kelly para a escolha de carteiras.

O critério de Kelly também é conhecido por critério de crescimento ótimo ou log-ótimo. Ele será provado de duas maneiras quando T tende a infinito: primeiro, usando o homeomorfismo dos mundos de retornos discretos e contínuos; e segundo, pela esperança do limite da distribuição composta média.

Neste capítulo, as carteiras serão consideradas, por hipótese, de pesos constantes com distribuição i.i.d. ao longo dos infinitos períodos. As carteiras com possibilidade de falência também estão excluídas por hipótese, pois a falência certamente ocorrerá quando T tende a infinito.

7.2. Limite da distribuição composta média

Se o horizonte de tempo T do investidor tender a infinito, então a distribuição composta média de uma variável aleatória multiplicativa tenderá a uma distribuição Delta de Dirac. Isto será provado a seguir.

Teorema 7.2.1. Limite da distribuição composta média quando o horizonte de tempo tende a infinito.

Seja X uma variável aleatória multiplicativa e i.i.d. ao longo do tempo, tal que

$$X \sim f(x) \xrightarrow{\mathcal{M}} F(s);$$

e um coeficiente de longevidade $\tau \geq 0$ tal que

$$\varphi(x) = \left((f^{\otimes T})^{\otimes \frac{1}{T+\tau}} \right)(x) = \left(f^{\otimes \frac{1}{T+\tau}} \right)^{\otimes T}(x).$$

Então

$$\lim_{T \rightarrow \infty} \varphi(x) = \delta(x - e^{F'(1)}). \quad (7.1)$$

Prova. A densidade $\varphi(x)$ e sua transformada de Mellin estão relacionadas pela expressão

$$\varphi(x) = \left((f^{\otimes T})^{\otimes \frac{1}{T+\tau}} \right)(x) = \left(\left(f^{\otimes \frac{1}{T+\tau}} \right)^{\otimes T} \right)(x) \xrightarrow{\mathcal{M}} \left(F \left(1 + \frac{s-1}{T+\tau} \right) \right)^T.$$

Fazendo T tender a infinito e aplicando o **Lema 3.16.3**, obtém-se

$$\lim_{T \rightarrow \infty} \varphi(x) \xrightarrow{\mathcal{M}} \lim_{T \rightarrow \infty} \left(F \left(1 + \frac{s-1}{T+\tau} \right) \right)^T = \exp[F'(1)(s-1)].$$

Como

$$\delta(x - e^k) \xrightarrow{\mathcal{M}} e^{k(s-1)},$$

então

$$\lim_{T \rightarrow \infty} \varphi(x) = \lim_{T \rightarrow \infty} \left((f^{\otimes T})^{\otimes \frac{1}{T+\tau}} \right)(x) = \lim_{T \rightarrow \infty} \left(\left(f^{\otimes \frac{1}{T+\tau}} \right)^{\otimes T} \right)(x) = \delta(x - e^{F'(1)}).$$

□

Este teorema vale para distribuições que têm uma densidade maior que 0 no ponto 0, como a exponencial. Densidades do tipo $k\delta(x-0) + (1-k)f(x)$, onde $0 < k \leq 1$, ou as que assumem valor positivo para $x < 0$, têm, no limite quando T tende a infinito, uma densidade composta média igual a $\delta(x-0)$, isto é, elas causam a ruína do investidor no longo prazo.

A seguir, serão calculados os limites da média e da variância da distribuição composta média quando o horizonte de tempo tende a infinito.

Teorema 7.2.2. Limite da esperança e da variância da distribuição composta média quando o horizonte de tempo tende a infinito. Seja X uma variável aleatória multiplicativa e i.i.d. ao longo do tempo, tal que

$$X \sim f(x) \xrightarrow{\mathcal{M}} F(s);$$

e um coeficiente de longevidade $\tau \geq 0$ tal que

$$\varphi(x) = \left((f^{\otimes T})^{\otimes \frac{1}{T+\tau}} \right)(x) = \left(f^{\otimes \frac{1}{T+\tau}} \right)^{\otimes T}(x).$$

Então

$$\lim_{T \rightarrow \infty} E(\varphi(x)) = e^{F'(1)} \quad (7.2)$$

e

$$\lim_{T \rightarrow \infty} V(\varphi(x)) = 0. \quad (7.3)$$

Prova. A densidade e a transformada de Mellin do produto de T variáveis aleatórias independentes e de sua potência $1/(T+\tau)$ são

$$\underbrace{X \cdot \dots \cdot X}_T = X^{\otimes T} \sim g(x) = f^{\otimes T}(x) \xrightarrow{\mathcal{M}} G(s) = (F(s))^T$$

e

$$\begin{aligned} \underbrace{(X \cdot \dots \cdot X)}_T^{1/(T+\tau)} &= (X^{\times T})^{1/(T+\tau)} \sim \varphi(x) = \left((f^{\otimes T})^{\otimes \frac{1}{T+\tau}} \right)(x) = \left(\left(f^{\otimes \frac{1}{T+\tau}} \right)^{\otimes T} \right)(x) = \left(g^{\otimes \frac{1}{T+\tau}} \right)(x) = \\ &= (T+\tau) x^{T+\tau-1} g(x^{T+\tau}) \xrightarrow{\mathcal{M}} \left(F \left(1 + \frac{s-1}{T+\tau} \right) \right)^T. \end{aligned}$$

A função $M(n)$ é definida como

$$M(n) = \lim_{T \rightarrow \infty} \int_0^\infty x^n \varphi(x) dx = \lim_{T \rightarrow \infty} \int_0^\infty x^n (T+\tau) x^{T+\tau-1} g(x^{T+\tau}) dx = \lim_{T \rightarrow \infty} (T+\tau) \int_0^\infty x^{T+\tau+n-1} g(x^{T+\tau}) dx.$$

Fazendo $y = x^{T+\tau}$, então $x = y^{1/(T+\tau)}$ e $dx = y^{-1+1/(T+\tau)} dy / (T+\tau)$. A função $M(n)$ se torna

$$\begin{aligned} M(n) &= \lim_{T \rightarrow \infty} (T+\tau) \int_0^\infty y^{1+\frac{n-1}{T+\tau}} g(y) \frac{1}{T+\tau} y^{-1+\frac{1}{T+\tau}} dy = \lim_{T \rightarrow \infty} \int_0^\infty y^{\frac{n}{T+\tau}} g(y) dy = \\ &= \lim_{T \rightarrow \infty} \int_0^\infty y^{s-1} g(y) dy \Bigg|_{s=1+\frac{n}{T+\tau}} = \lim_{T \rightarrow \infty} G(s) \Bigg|_{s=1+\frac{n}{T+\tau}} = \lim_{T \rightarrow \infty} G \left(1 + \frac{n}{T+\tau} \right) = \\ &= \lim_{T \rightarrow \infty} \left(F \left(1 + \frac{n}{T+\tau} \right) \right)^T. \end{aligned}$$

Aplicando o **Lema 3.16.3**, $M(n)$ se torna

$$M(n) = \exp(nF'(1)).$$

Logo,

$$\lim_{T \rightarrow \infty} E(\varphi(x)) = \lim_{T \rightarrow \infty} \int_0^\infty x \varphi(x) dx = M(1) = e^{F'(1)}$$

e

$$\lim_{T \rightarrow \infty} V(\varphi(x)) = M(2) - [M(1)]^2 = e^{2F'(1)} - [e^{F'(1)}]^2 = 0.$$

□

Pelo teorema anterior, a média e variância da distribuição limite $\varphi(x) = \delta(x - e^{F'(1)})$ coincidem, respectivamente, com os limites da média e da variância da distribuição composta média quando T tende a infinito. Note que os demais momentos também coincidem.

Hakansson (1971a, p.865) estudou o retorno composto médio de seqüências de carteiras no limite quando o horizonte de tempo T tende a infinito. Em seu artigo, algumas carteiras da seqüência podem causar falência. Nas seqüências em que isto não ocorre, Hakansson obteve expressões para a esperança e variância da distribuição composta média no limite, que são as dadas no **Teorema 7.2.2**.

Quando o horizonte de tempo T tende a infinito, o coeficiente de longevidade τ não aparece na distribuição limite $\varphi(x)$. Isto significa que a carteira a ser escolhida não depende do fator subjetivo τ . Por

ser uma Delta de Dirac, a carteira a ser escolhida também independe da função utilidade (que deve ser crescente) do investidor. Todos os investidores escolherão a carteira com máximo $\exp[F'(1)]$, ou máximo $F'(1)$.

7.2.1. Função utilidade logarítmica

No capítulo anterior, foi visto que a função utilidade potência está relacionada com o conceito de distribuição composta média. Se a função utilidade for multiplicada por um número positivo ou se a ela for adicionada uma constante, então a carteira de máxima utilidade continuará a mesma. Portanto, as funções utilidade potência, isoelástica ou CRRA podem ser definidas como

$$U(W) = \frac{W^{\frac{1}{T+\tau}} - 1}{1/(T+\tau)} \quad (7.4)$$

e, então, obtém-se

$$\lim_{T \rightarrow \infty} U(W) = \lim_{T \rightarrow \infty} \frac{W^{\frac{1}{T+\tau}} - 1}{1/(T+\tau)} = \ln(W). \quad (7.5)$$

Note que o expoente da função utilidade potência varia com o horizonte de tempo T . No limite quando T tende a infinito, essa função utilidade se torna a logarítmica.

7.3. O critério de Kelly

Em seu artigo clássico, Kelly (1956) demonstrou o critério que leva seu nome usando conceitos relacionados à informação mútua, mas na conclusão, ele já reconhecia que o logaritmos dos retornos são aditivos em repetidas apostas no tempo e que é a eles que a lei dos grandes números se aplica. Ele também concluiu que com probabilidade 1, o investidor que seguisse seu critério iria enriquecer mais que um investidor que seguisse qualquer outro critério. Isto foi formalizado posteriormente por Breiman (1961). Kelly estudou jogos nos quais o perdedor perde o valor apostado; portanto, para evitar a ruína, somente uma fração de sua riqueza deve ser investida no jogo. A carteira pode conter um ativo “sem risco” (o bolso) de forma a impedir a ruína do investidor. É claro que, se a esperança dos retornos da carteira for menor que 1, o investidor não deve escolhê-la.

O critério de Kelly, com as hipóteses necessárias, pode ser enunciado com o teorema a seguir.

Teorema 7.3.1. Critério de Kelly para a escolha de uma carteira. Seja R uma variável aleatória multiplicativa e i.i.d. que representa os retornos no mundo discreto de uma carteira, tal que R tem esperança maior que 1 e probabilidade zero de ter retornos menores ou iguais a zero. Se o horizonte de tempo T de um investidor tender a infinito, então ele deve escolher a carteira que maximiza a esperança dos logaritmos dos retornos.

Prova 1. Suponha que a variável aleatória R segue uma densidade $f(R)$. As condições de que R tem esperança maior que 1 e probabilidade zero de ter retornos menores ou iguais a zero excluem as carteiras com possibilidade de causar falência do investidor. Como a variável R é multiplicativa e está no mundo dos retornos discretos, é possível levá-la para o mundo dos retornos contínuos com uma mudança de variáveis. Fazendo $\tilde{R} = \ln(R)$, então $R = e^{\tilde{R}}$, $dR = e^{\tilde{R}} d\tilde{R}$ e $\tilde{f}(\tilde{R}) = e^{\tilde{R}} f(e^{\tilde{R}})$.

No mundo dos retornos contínuos, a variável $\tilde{R}$ é aditiva e i.i.d. ao longo do tempo. Pela lei dos grandes números, a média aritmética após t possíveis realizações, onde t tende a infinito, tenderá para sua média teórica, calculada na medida de probabilidade dos retornos contínuos, que é

$$\tilde{E}[\tilde{R}] = \int_{-\infty}^{\infty} \tilde{R} \tilde{f}(\tilde{R}) d\tilde{R} = \int_{-\infty}^{\infty} \tilde{R} e^{\tilde{R}} f(e^{\tilde{R}}) d\tilde{R} = \int_0^{\infty} \ln(R) f(R) dR = E[\ln(R)]. \quad (7.6)$$

O valor $E[\ln(R)]$ é um número escalar no mundo dos retornos contínuos, o qual deve ser mapeado para o escalar $e^{E[\ln(R)]}$ no mundo dos retornos discretos. Como o investidor deseja maximizar a esperança da carteira no mundo discreto, ele deve escolher a carteira com máximo $e^{E[\ln(R)]}$, ou, porque a exponencial é uma função crescente, a carteira com máximo $E[\ln(R)]$.

□

Prova 2. Suponha que a variável aleatória R segue uma densidade $f(R)$. As condições de que R tem esperança maior que 1 e probabilidade zero de ter retornos menores ou iguais a zero excluem as carteiras com possibilidade de causar falência do investidor. Como o investidor deseja maximizar a esperança de $\varphi(R)$, então, pelo **Teorema 7.2.2**, ele deve escolher a carteira com o máximo escalar $\exp[F'(1)]$, ou máximo $F'(1)$. Pela propriedade

$$f(R)(\ln R)^n \xrightarrow{\mathcal{M}} F^{(n)}(s),$$

a carteira deve ser a que tiver máximo

$$F'(1) = \int_0^{\infty} R^{1-1} f(R) \ln(R) dR = \int_0^{\infty} f(R) \ln(R) dR = E[\ln(R)]. \quad (7.7)$$

□

A **Prova 1** usa os conceitos de mundos de retornos discretos e contínuos, e é similar à dada por Luenberger (1998, p.419). No mundo dos retornos contínuos ou mundo aditivo, os conceitos usuais da Estatística podem ser aplicados. Nesse mundo, pela lei dos grandes números, o limite da média aritmética de n possíveis realizações multiperiodicas, quando n tende a infinito, é igual à sua média teórica. O resultado é um número escalar que deve ser trazido de volta para o mundo discreto e, então, maximizado.

A **Prova 2** usa o conceito de distribuição composta média no limite e a transformada de Mellin. O conceito de distribuição composta média com longevidade foi projetado, construído ou definido de forma que a hipótese

$$\lim_{T \rightarrow \infty} \frac{T + \tau}{T} = 1$$

do **Lema 3.16.3** pudesse ser usada e que a **Prova 2** pudesse ser feita. O propósito foi obter concordância com resultados conhecidos nos casos com um e com infinitos períodos. Portanto, é discutível se a **Prova 2** é uma “prova” ou uma “construção matemática feita para se obter resultados conhecidos”. Se esta definição puder ser obtida a partir de conceitos mais simples (axiomas), então a **Prova 2** seria uma prova, e o critério de Kelly passaria a ser um teorema. De qualquer forma, este conceito é importante porque traz um novo conhecimento para o caso multiperiódico. Milton Friedman (1953) disse que uma teoria deve ser julgada mais pelo realismo de suas predições e menos pela relevância de suas premissas.

Um investidor que tem a função utilidade potência definida em (7.4) (ela foi usada, com uma transformação linear, na **Seção 6.7.1**), terá, no limite quando o horizonte de tempo T tender a infinito, a função utilidade logarítmica. Veja a **Seção 7.2.1**.

Kelly (1956, p.925) e Bell e Cover (1980) disseram que o fato de o investidor maximizar a esperança dos logaritmos dos retornos não tem relação com a função utilidade que ele atribui à riqueza. Veja também Latané (1979). Mossin (1968) e Hakansson (1971a, 1971b) discutiram o conceito de estratégia míope, na qual a carteira escolhida para o caso uniperiódico é a mesma para o caso multiperiódico (ou com infinitos períodos). Na opinião do autor, um investidor que segue o critério de Kelly não é equivalente a um investidor com utilidade logarítmica porque, embora a função a ser maximizada seja a mesma, as motivações são diferentes. Outro ponto a ser considerado é que a teoria da utilidade não adota sempre a hipótese de horizonte de tempo tendendo a infinito.

Bell e Cover (1980) usaram o critério de Kelly para resolver o Paradoxo de São Petersburgo no caso com infinitos períodos. A solução utilitária para o paradoxo, dada por Bernoulli, tem a desvantagem de que caso a função utilidade não seja limitada superiormente, as condições do paradoxo podem ser alteradas de forma a tornar novamente infinita a esperança da utilidade; veja Samuelson (1977). É possível aplicar o conceito de distribuição composta média para o caso multiperiódico do paradoxo, mas o resultado irá depender do coeficiente de longevidade, e também será possível alterar as condições do paradoxo de forma a tornar infinita a esperança nessa medida de probabilidade.

Markowitz (1991a, p.365) discutiu uma crítica feita por Samuelson (1969, 1971) ao critério de Kelly e deu um exemplo simples relacionando o ponto de vista dos dois autores. Esta crítica se reduz à diferença discutida no capítulo anterior entre os critérios da maximização da esperança da distribuição no instante final e da maximização da distribuição composta média.

MacLean, Ziemba e Blazenko (1992) listaram várias propriedades do critério de Kelly. A esperança dos logaritmos dos retornos é uma função côncava sobre os pesos não-negativos dos ativos.

O critério de Kelly pode ser formulado com o problema de otimização a seguir.

Problema de otimização 7.3.1. Critério de Kelly. Existem D ativos dependentes. A densidade multivariada de seus retornos $\mathbf{R}_d$, para $1 \leq d \leq D$, é conhecida. O domínio dos retornos dos ativos é o dos reais não-

negativos. Os pesos w_d , para $1 \leq d \leq D$, são as variáveis de otimização, devem somar 1 e ser não-negativos. A densidade dos retornos da carteira $\mathbf{R}_P$ é expressa por

$$\mathbf{R}_P = \sum_{d=1}^D w_d \mathbf{R}_d \sim f(R; w_1, \dots, w_D).$$

A carteira de Kelly é obtida com a formulação a seguir.

$$\begin{aligned} & \max_{w_d} E[\ln(\mathbf{R}_P)], \quad 1 \leq d \leq D \\ & s.t. \\ & w_d \geq 0, \quad 1 \leq d \leq D \\ & \sum_{d=1}^D w_d = 1. \end{aligned}$$

Se a distribuição multivariada dos retornos for representada por um produto tensorial finito de Erlangs, i.i.d. ao longo do tempo, então a formulação a seguir pode ser feita.

Problema de otimização 7.3.2. Critério de Kelly para ativos com distribuição multivariada expressa por um produto tensorial finito de Erlangs. Existem D ativos dependentes. Há probabilidade zero de os retornos dos ativos serem menores ou iguais a zero. Os pesos w_d , para $1 \leq d \leq D$, são as variáveis de otimização, devem somar 1 e ser não-negativos.

A densidade multivariada dos retornos dos ativos $\mathbf{R}_d$, para $1 \leq d \leq D$, é representada por um produto tensorial finito de Erlangs e foi dada na **Seção 5.4.1**.

$$[\mathbf{R}_1, \dots, \mathbf{R}_D]^\top \sim m(R_1, \dots, R_D) = \sum_{i=1}^n p_i \gamma_{b_{i,1}}^{a_{i,1}}(R_1) \dots \gamma_{b_{i,D}}^{a_{i,D}}(R_D) \geq 0, \quad \sum_{i=1}^n p_i = 1, \quad 1 \leq n < \infty.$$

A densidade dos retornos da carteira $\mathbf{R}_P$ é expressa por uma mistura finita de Erlangs e foi dada na **Seção 5.4.2**.

$$\mathbf{R}_P = \sum_{d=1}^D w_d \mathbf{R}_d \sim f(R; w_1, \dots, w_D) = \sum_{i=1}^n p_i \left(\gamma_{b_{i,1}}^{a_{i,1}/w_1} \oplus \dots \oplus \gamma_{b_{i,D}}^{a_{i,D}/w_D} \right)(R) = \sum_{d'=1}^{D'} \sum_{r=1}^{b'_{\max,d'}} k_{r,d'} \gamma_r^{a_{d'}/w_{d'}}(R),$$

onde

$$D' = \text{número de componentes distintos do vetor } \left[\frac{a_1}{w_1}, \dots, \frac{a_d}{w_d}, \dots, \frac{a_D}{w_D} \right]^\top, \quad (7.8)$$

$$\left[\frac{a'_1}{w'_1}, \dots, \frac{a'_{d'}}{w'_{d'}}, \dots, \frac{a'_{D'}}{w'_{D'}} \right]^\top = \text{vetor } \left[\frac{a_1}{w_1}, \dots, \frac{a_d}{w_d}, \dots, \frac{a_D}{w_D} \right]^\top \text{ sem componentes repetidos,} \quad (7.9)$$

$$b'_{i,d'} = \sum_{d=1}^D b_{i,d} \frac{a_d}{w_d} \frac{a'_{d'}}{w'_{d'}}, \quad 1 \leq i \leq n, \quad 1 \leq d' \leq D', \quad (7.10)$$

$$b'_{\max,d'} = \max_{1 \leq i \leq n} b'_{i,d'}, \quad 1 \leq d' \leq D', \quad (7.11)$$

$$k_{r,d'} = \sum_{\substack{i=1 \\ r \leq b'_{i,d'}}}^n p_i \left\{ \prod_{\substack{j=1 \\ j \neq d'}}^{D'} \left(1 - \frac{a'_{d'} w'_j}{w'_{d'} a'_j} \right)^{-b'_{i,j}} \left(\frac{a'_{d'}}{w'_{d'}} \right)^{b'_{i,d'} - r} M_{i,d'}^{b'_{i,d'} - r} \right\}, \quad 1 \leq r \leq b'_{\max,d'}, \quad 1 \leq d' \leq D' \quad (7.12)$$

e

$$M_{i,d'}^u = \begin{cases} 1 & , \quad u = 0 \\ \frac{1}{u} \sum_{v=0}^{u-1} \left[\sum_{\substack{j=1 \\ j \neq d'}}^{D'} b'_{i,j} \left(\frac{a'_{d'}}{w'_{d'}} - \frac{a'_j}{w'_j} \right)^{v-u} \right] M_{i,d'}^v & , \quad u > 0 \end{cases} \quad (7.13)$$

O seguinte resultado vale para uma distribuição Erlang (ou Gama).

$$\int_0^\infty \ln(R) \gamma_b^a(R) dR = \psi(b) - \ln(a) \quad (7.14)$$

O problema de otimização se torna

$$\begin{aligned} \max_{w_d} E[\ln(\mathbf{R}_P)] &= \sum_{d'=1}^{D'} \sum_{r=1}^{b'_{\max,d'}} k_{r,d'} \left(\psi(r) - \ln\left(\frac{a'_{d'}}{w'_{d'}}\right) \right), \quad 1 \leq d \leq D \\ \text{s.t.} \\ w_d &\geq 0, \quad 1 \leq d \leq D \\ \sum_{d=1}^D w_d &= 1 \end{aligned} \quad (7.15)$$

7.4. Fronteira eficiente

Neste capítulo, foi assumido que o horizonte de tempo tende a infinito e, por isso, será usada a distribuição composta média no limite quando o tempo tende a infinito. Pela **Prova 2** do **Teorema 7.3.1**, as densidades de todas as carteiras de pesos constantes serão Deltas de Dirac, isto é,

$$\varphi(R) = \delta(R - e^{F'(1)}) = \delta(R - e^{E[\ln(\mathbf{R}_P)]}). \quad (7.16)$$

No gráfico retorno esperado *versus* risco, a região das possíveis carteiras de pesos constantes, na medida de probabilidade composta média, se torna um segmento de uma curva. A carteira ótima $\mathbf{R}_P^*$ obtida no **Problema de otimização 7.3.1** é a carteira de Kelly ou de crescimento ótimo. Considere que $R_{\text{Kelly}} = \exp(E[\ln(\mathbf{R}_P^*)])$. Se $R_{\text{Kelly}} \geq 1$, então a fronteira eficiente se torna um único ponto. Se $R_{\text{Kelly}} < 1$, então a fronteira eficiente não existe, e o investidor não deve colocar seu dinheiro nos ativos. Note que se o investidor puder deixar o dinheiro em seu bolso ou no ativo livre de risco, então este caso não ocorrerá.

Considere a medida de risco *Downside Risk* com um retorno padrão R_B tal que $1 \leq 1 + r_f \leq R_B$. Então o gráfico será similar ao dado na **Figura 7.4.1**. Se $R_{\text{Kelly}} < R_B$, então a carteira de Kelly terá risco, embora ela tenha a maior taxa de crescimento e torne o investidor infinitamente rico no limite quando T tende a infinito. Por isso, pode-se argumentar que a condição $1 \leq 1 + r_f \leq R_B \leq R_{\text{Kelly}}$ deve ocorrer, e que se R_B depender do horizonte de tempo, deve tender a R_{Kelly} .

Figura 7.4.1. Fronteira eficiente com *Downside Risk* para infinitos períodos na medida de probabilidade composta média. Se $R_{Kelly} \geq 1$, então a fronteira eficiente será um único ponto. O restante da fronteira serão (em geral) dois segmentos de reta.

Hakansson (1971a, p.866) disse que a carteira de crescimento ótimo (de Kelly) é a única carteira eficiente no limite “sem variância”. Veja também Luenberger (1998, p.432).

Hakansson (1971c) deu um exemplo mostrando que, para distribuições genéricas, a carteira de Kelly não está na fronteira eficiente do gráfico retorno esperado *versus* desvio-padrão, na medida de probabilidade uniperiódica. Entretanto, Elton e Gruber (1974c) provaram que, se as carteiras possuem distribuições lognormais, então a carteira de Kelly está sobre essa fronteira.

7.5. Derivadas do critério de Kelly

Pelo critério de maximização da distribuição composta média, a carteira que deve ser escolhida é a com máximo $F(1 + 1/(T + \tau))$. Se o horizonte de tempo T tender a infinito, esses valores tenderão a $F(1) = 1$, que é a propriedade da unidade da distribuição. Intuitivamente, a carteira que deve ser escolhida é a com maior $F'(1)$. Pelo mesmo motivo, dadas duas carteiras com o mesmo valor para $F'(1)$, a que deveria ser escolhida é a que tiver maior $F''(1)$, e assim por diante.

Entretanto, basta calcular a primeira derivada, pois as demais são funções crescentes da primeira. Se

$$G(s) = F\left(1 + \frac{s-1}{T + \tau}\right)^T,$$

então

$$\begin{aligned}
G(s) \Big|_{s=2}^{T \rightarrow \infty} &= F \left(1 + \frac{s-1}{T+\tau} \right)^T \Big|_{s=2}^{T \rightarrow \infty} = \exp[F'(1)(s-1)] \Big|_{s=2}^{T \rightarrow \infty} = \exp[F'(1)], \\
G'(s) \Big|_{s=2}^{T \rightarrow \infty} &= \frac{\partial}{\partial s} F \left(1 + \frac{s-1}{T+\tau} \right)^T \Big|_{s=2}^{T \rightarrow \infty} = \frac{T}{T+\tau} F \left(1 + \frac{s-1}{T+\tau} \right)^{T-1} F' \left(1 + \frac{s-1}{T+\tau} \right) \Big|_{s=2}^{T \rightarrow \infty} = F'(1) \exp[F'(1)], \\
G''(s) \Big|_{s=2}^{T \rightarrow \infty} &= \frac{\partial^2}{\partial s^2} F \left(1 + \frac{s-1}{T+\tau} \right)^T \Big|_{s=2}^{T \rightarrow \infty} = \frac{(T-1)T}{(T+\tau)^2} F \left(1 + \frac{s-1}{T+\tau} \right)^{T-2} \left(F' \left(1 + \frac{s-1}{T+\tau} \right) \right)^2 + \\
&\quad \frac{T}{(T+\tau)^2} F \left(1 + \frac{s-1}{T+\tau} \right)^{T-1} F'' \left(1 + \frac{s-1}{T+\tau} \right) \Big|_{s=2}^{T \rightarrow \infty} = (F'(1))^2 \exp[F'(1)],
\end{aligned}$$

e, em geral,

$$G^{(n)}(s) \Big|_{s=2}^{T \rightarrow \infty} = (F'(1))^n \exp[F'(1)], \quad n = 0, 1, 2, \dots \quad (7.17)$$

É possível generalizar a ordem da derivada para o caso fracionário $n \geq 0$; veja Mainardi, Paridisi e Gorenflo (1998, p.21). Leibniz, em 1695, já conhecia a derivada de ordem $n = 1/2$; veja Gorenflo e Mainardi (2000). A expressão (7.17) se torna

$$G^{(n)}(s) \Big|_{s=2}^{T \rightarrow \infty} = (F'(1))^n \exp[F'(1)], \quad n \in \mathbb{R}_+^*. \quad (7.18)$$

É suficiente escolher a carteira com máximo $F'(1)$, pois as suas derivadas de ordem $n \geq 0$ também estarão maximizadas. Veja também Platen (2002, lema 3.7 e teorema 3.8).

7.6. Conclusões

No capítulo anterior, foram apresentados os conceitos de que o retorno discreto é uma variável multiplicativa ao longo do tempo e de que os mundos de retornos discretos e contínuos são homeomorfos. Se a densidade de probabilidade dos retornos discretos uniperiódicos das carteiras for i.i.d. e o horizonte de tempo T do investidor tender a infinito, então estes conceitos resultam no bem conhecido critério de Kelly para a escolha de uma carteira. Por este critério, a carteira que deve ser escolhida é a que tiver a maior esperança dos logaritmos dos retornos discretos. Esta é a carteira de Kelly ou de crescimento ótimo.

A distribuição composta média, usada no caso multiperiódico, permite uma transição suave da distribuição usada no caso uniperiódico para a usada no caso com infinitos períodos. A esperança e o risco podem ser calculados com esta distribuição (ou medida de probabilidade), qualquer que seja o horizonte de tempo.

O limite da distribuição composta média quando T tende a infinito é uma Delta de Dirac. Esta distribuição é a usada implicitamente pelo critério de Kelly. A carteira que deve ser escolhida é a que tiver maior $F'(1)$. Uma formulação mais precisa para o critério de Kelly seria escolher a carteira que maximiza a esperança da distribuição composta média no limite quando T tende a infinito. A carteira ótima será a mesma nas duas formulações, mas a distribuição em questão fica mais clara e pode evitar confusões.

Pelo conceito de distribuição composta média, todos os investidores concordarão na escolha da carteira de Kelly, independentemente de sua função utilidade, de sua medida de risco e de seu coeficiente de

longevidade.

Caso haja empate entre duas carteiras pelo critério da distribuição composta média, não adianta calcular as derivadas de ordem superior, pois elas serão iguais e já estarão maximizadas.

A fronteira eficiente, calculada na medida de probabilidade composta média, caso exista, será um único ponto, o qual representa a carteira de Kelly. Esta é a carteira a ser escolhida por um investidor racional. O restante da fronteira se torna um segmento de uma curva e não um compacto. Se a medida de risco for o *Downside Risk*, então o restante da fronteira se torna, em geral, dois segmentos de reta.

8. SELEÇÃO DE UM MODELO DE OTIMIZAÇÃO DE CARTEIRAS

8.1. Introdução

Um modelo científico é uma simplificação do mundo real. À medida que a ciência avança, os modelos ficam cada vez mais perto do mundo real, no sentido de explicar mais eventos passados e prever eventos futuros com maior precisão. As premissas simplificadoras dos modelos são chamadas hipóteses. Sobre estas hipóteses, o raciocínio lógico é aplicado, e o modelo é construído. A conclusão ou resultado é a tese. Se as hipóteses forem válidas, então a tese será válida. Matematicamente,

$$\text{HIPÓTESES} \Rightarrow \text{TESE.}$$

Assumindo que a construção de cada modelo está logicamente correta, então possíveis diferenças nos resultados devem ser atribuídas às diferenças nas hipóteses. Modelos têm que ser julgados por quão perto suas hipóteses estão do mundo real. No entanto, se a diferença entre os resultados for pequena, então, por consideração prática, o mais fácil de se calcular pode ser escolhido.

Constructo é uma operacionalização matemática de um conceito. Um conceito pode ser expresso de diferentes formas, isto é, ele pode ter diferentes definições matemáticas. Por exemplo, o retorno esperado pode ser calculado pela média aritmética dos últimos n retornos ou pela média teórica de uma densidade cujos parâmetros foram calibrados para os dados amostrais. Risco possui várias definições, tais como desvio-padrão, semivariância, valor-em-risco e *Downside Risk*. Claramente, a escolha dos constructos afetará o resultado de um modelo de otimização de carteiras.

Tendo dois ou mais modelos de otimização de carteiras, um investidor terá que escolher um deles antes de investir seu dinheiro. Isto é semelhante à escolha de um dentre vários fundos mútuos, mas agora os detalhes internos dos modelos são conhecidos. A seleção de um modelo de otimização de carteiras é um teste de “caixa transparente”, em contraste aos testes de “caixa preta” tais como os índices de Sharpe, Jensen e Treynor, usados para selecionar um dentre vários fundos.

Neste capítulo, será proposto um critério para ajudar o investidor a escolher um dentre vários modelos de otimização de carteiras.

8.2. Motivação

Dados dois ou mais modelos de alocação de capital, um investidor tem interesse nos desempenhos relativos entre eles. Ele precisa escolher um dos modelos antes de investir seu dinheiro. É necessário compará-los de alguma forma.

Considere o caso uniperiódico. O modelo de Markowitz é o modelo 1, e o apresentado neste trabalho é o modelo 2. O modelo 1 assume uma normal multivariada para os retornos dos ativos e usa o desvio-padrão como medida de risco. O modelo 2 assume um produto tensorial finito de Erlangs e usa o *Downside Risk*. Se as hipóteses de um modelo representam exatamente o que ocorre no mundo real, então esse é o modelo que deve ser escolhido. Entretanto, os modelos são apenas aproximações do mundo real. No mundo real, a distribuição dos retornos não é uma normal multivariada, nem um produto tensorial finito de Erlangs, e nem independente ao longo do tempo. Para um investidor descrever sua medida de risco ou função utilidade, ele deve fornecer parâmetros cujos valores são obtidos de forma subjetiva. É necessário usar dados históricos nos testes, mas a probabilidade de que as realizações de variáveis aleatórias contínuas se repitam em um futuro finito é (ou tende a) zero. Estes são problemas que sempre afetarão os testes. Qualquer que seja a forma de comparação entre os modelos, é recomendável prudência sempre que os resultados forem extrapolados, pois dependem dos parâmetros, dados usados, e hipóteses assumidas.

O produto tensorial finito de Erlangs é definido sobre os reais não-negativos e pode ter um número arbitrário de parcelas, de forma que pode aproximar, com precisão arbitrária, qualquer densidade contínua sobre um domínio fechado. O desvio-padrão não tem sentido financeiro para distribuições assimétricas porque penaliza os retornos altos. Por isso, o modelo 2 parece ser qualitativamente superior ao modelo 1. Entretanto, é interessante fazer testes numéricos comparando os modelos, pois nem sempre a superioridade de um modelo é aparente. Os testes podem ser usados para comparar dois modelos similares, tais como modelos que diferem apenas no parâmetro α do *Downside Risk*. Também é possível que existam efeitos no mundo real não considerados no conjunto de hipóteses dos modelos, que podem afetar significativamente o resultado final.

Um possível teste é calcular a fronteira eficiente dos modelos e, para um certo nível de retorno esperado, encontrar a carteira eficiente de cada um. Este procedimento é repetido para vários períodos de um *backtest* e, então, se calcula o desvio-padrão e o *Downside Risk*. Entretanto, os resultados indicarão apenas que o desvio-padrão está minimizado para o modelo 1 e que o *Downside Risk* está minimizado para o modelo 2. Isto não implica que um modelo tem melhor desempenho que o outro. Cheng e Wolverton (2001) reconheceram este problema. A seguir, será descrito um procedimento de teste um pouco melhor.

8.3. Metodologia

Para simplificar a exposição, nesta seção será considerado que o investidor quer escolher um dentre dois modelos. Caso haja mais de dois modelos, o raciocínio é análogo.

Os modelos de alocação de capital assumem uma aproximação para a densidade de probabilidade

multivariada dos retornos dos ativos dependentes. Os resultados dos testes comparativos serão afetados se os estimadores forem tendenciosos. As carteiras calculadas poderão ser tendenciosas, isto é, poderão ter sempre valor esperado ou risco maiores (ou menores) do que os previstos. A obtenção de estimadores melhores (menos tendenciosos ou com menor dispersão) para a distribuição multivariada deve sempre ser tentada, qualquer que seja o modelo de otimização, pois esta etapa independe da seguinte, que é o cálculo do retorno esperado e do risco. Possíveis formas de diminuir o problema dos estimadores tendenciosos são: aumentar o número de amostras; usar variações do critério da máxima verossimilhança que considerem um número pequeno de amostras; e obter multiplicadores empíricos com simulações Monte Carlo controladas. Entretanto, isto está fora do escopo deste trabalho.

Os modelos de otimização de carteiras calculam as carteiras sobre uma fronteira eficiente, mas eles usam diferentes definições para retorno esperado e risco. As fronteiras eficientes são traçadas em diferentes gráficos. Para permitir uma comparação mais justa, é necessário colocá-las no mesmo gráfico. Cada modelo tem como resultado os pesos dos ativos das carteiras sobre sua fronteira eficiente. Portanto é possível calcular o risco e o retorno esperado dessas carteiras de acordo com outra definição. Isto é, a fronteira eficiente de um modelo pode ser traçada no gráfico de outro modelo. A nova curva sempre será interior à fronteira original porque esta já foi otimizada para as definições de retorno esperado e risco. As carteiras sobre uma linha horizontal ou vertical estão totalmente ordenadas e é isto que torna a comparação mais justa.

Um nível de retorno esperado deve ser escolhido arbitrariamente. Posteriormente, o teste deve ser repetido para outros valores de retorno esperado. Embora seja possível escolher um nível de risco, é preferível fixar o retorno esperado porque o retorno é uma variável observável no final de cada período, enquanto o risco é apenas estimável. Além disso, fixar o retorno esperado é computacionalmente mais conveniente. Os diferentes modelos podem calcular o retorno esperado de diferentes formas, mas como as fronteiras estão no mesmo gráfico, isto será considerado automaticamente. Uma linha horizontal interceptará as fronteiras dos modelos, e uma carteira para cada modelo será escolhida.

Um *backtest* usa, em geral, uma janela deslizante de dados históricos para simular a carteira que seria escolhida com base nesses dados. Com isso, obtém-se uma série temporal de resultados realizados. Para cada período do *backtest*, podem ser obtidos quatro pontos: o retorno fixado; o retorno realizado do modelo 1; o retorno realizado do modelo 2; e a diferença entre os retornos realizados dos modelos 2 e 1. Esses pontos formarão séries temporais e resultarão em curvas em um gráfico com o tempo na abscissa. Também é possível traçar um histograma para cada uma das três últimas séries. Estes gráficos e histogramas servirão para comparar os desempenhos relativos entre os modelos.

O resultado dos testes irá depender das hipóteses assumidas. Se a hipótese de independência ao longo do tempo for considerada válida, então a média dos retornos realizados de cada modelo deverá ser igual ao retorno fixado. Caso não seja, os estimadores são tendenciosos, ou a aproximação da densidade multivariada foi ruim, ou a densidade multivariada não é idêntica ao longo do tempo, ou existe dependência temporal.

Usualmente, a calibração da densidade multivariada dos modelos é feita com dados históricos recentes. Caso ocorra um evento extremo sem similar nesse passado recente, então as hipóteses assumidas serão violadas. Isto afetará todos os modelos ao mesmo tempo, mas com diferentes intensidades, devido principalmente à medida de risco adotada pelo modelo. O gráfico da série temporal das diferenças entre os retornos dos modelos 2 e 1, e seu histograma serão bons indicadores dos desempenhos relativos entre os modelos.

No caso uniperiódico, possíveis critérios de comparação são: a média aritmética das diferenças entre os retornos, que é igual à diferença entre as médias aritméticas; o número de períodos com maior retorno; e o erro quadrático médio entre o retorno realizado e o esperado. No caso multiperiódico com um horizonte de tempo finito T , pode-se calcular o retorno efetivo (composto) para este horizonte de tempo, de forma que este caso se reduz ao uniperiódico. No caso com infinitos períodos, o critério de comparação deve ser o retorno composto.

Para evitar uma possível tendência causada pela utilização de apenas um retorno esperado, é necessário repetir os testes para outros valores. A visualização dos resultados é facilitada com gráficos. Para cada período do *backtest*, pode-se calcular as carteiras da fronteira eficiente de cada modelo, os retornos realizados dessas carteiras e as diferenças entre esses retornos. Então os retornos realizados das carteiras eficientes e suas diferenças podem ser traçados em gráficos com o retorno esperado na abscissa. Este procedimento deve ser repetido para todos os períodos do *backtest*, resultando em um gráfico com várias curvas, uma para cada período. A partir desses dados, também é possível traçar gráficos tridimensionais dos histogramas dos retornos realizados e de suas diferenças para intervalos de retornos esperados.

Um possível problema nos *backtests* é a utilização de janelas deslizantes, que causa a sobreposição ou reuso dos dados. Isto aumenta a correlação entre os retornos obtidos ao longo do *backtest*. Veja Cheng e Deets (1971) para maiores detalhes e referências. Entretanto, pode-se argumentar que este procedimento é o que será usado posteriormente na alocação de capital. Em um teste amplo, vários *backtests* devem ser feitos, com dados de diferentes ações, setores, bolsas, países, anos e diferentes parâmetros de aversão a risco, horizonte de tempo e longevidade. Se apenas um *backtest* for feito, os resultados, bons ou ruins, não devem ser extrapolados.

O procedimento proposto é um teste das hipóteses dos modelos, isto é, de quão perto as hipóteses de cada modelo estão do mundo real, e não um teste de hipótese da Estatística. O teste de um modelo é também um teste de suas hipóteses. Ainda é um problema de pesquisa qual teste estatístico usar para a comparação dos resultados dos modelos com um dado nível de confiança. Uma possível escolha é a análise de variância para a comparação de médias de diferentes grupos, mas isto deve ser mais bem detalhado.

8.4. Conclusões

Um investidor que deseja aplicar seu dinheiro tem interesse em conhecer os desempenhos relativos entre os modelos de alocação de capital. Assumindo que o desenvolvimento lógico-matemático de cada

modelo está correto, então são as hipóteses ou premissas adotadas que determinam o quão perto os modelos estão do mundo real.

Para uma comparação mais justa dos modelos, é necessário usar um mesmo padrão para todos. Isto pode ser entendido pela transposição da fronteira eficiente de cada modelo para um único gráfico e pela escolha de um determinado retorno esperado, comum a todos. Para este retorno esperado, é possível calcular o retorno realizado da carteira eficiente de cada modelo ao longo dos períodos de um *backtest*. Com os gráficos temporais dos retornos realizados das carteiras eficientes de cada modelo e das diferenças entre eles, os desempenhos dos modelos podem ser comparados.

Os resultados sempre dependerão das hipóteses assumidas pelos modelos. Caso ocorra um evento que invalide essas hipóteses, os efeitos terão diferentes intensidades em cada modelo. Ainda assim, é interessante comparar os desempenhos relativos dos modelos.

Neste trabalho, estava planejada a realização de alguns testes numéricos comparativos entre os modelos existentes e os propostos. Entretanto, devido a problemas de convergência numérica com os algoritmos de otimização, foi decidido que os testes serão realizados em um trabalho posterior.

9. CONCLUSÕES

9.1. Conclusões

Neste trabalho, o problema de seleção de carteiras foi estudado com uma abordagem em tempo discreto. Foi proposto o produto tensorial finito de Erlangs para aproximar a densidade multivariada dos retornos discretos uniperiódicos de ativos dependentes. Isto faz com que a densidade dos retornos discretos uniperiódicos das carteiras não-alavancadas seja uma mistura finita de Erlangs. Neste caso, o *Downside Risk* de uma carteira requer, em geral, a avaliação numérica da função Meijer G. No caso multiperiódico, foi usado o conceito de distribuição composta média, que permite uma transição suave do caso uniperiódico até o caso com infinitos períodos. A densidade dos retornos das seqüências de carteiras, na medida de probabilidade composta média, se torna uma mistura finita de densidades Meijer G. No limite quando o horizonte de tempo tende a infinito, o critério de Kelly é obtido. Para comparar os desempenhos relativos entre os modelos, é necessário traçar as fronteiras eficientes dos modelos em um mesmo gráfico, fixar um retorno esperado, fazer um *backtest* e traçar os gráficos temporais dos retornos realizados da carteira eficiente de cada modelo e de suas diferenças.

A distribuição uniperiódica dos retornos discretos de ativos dependentes foi aproximada com um produto tensorial finito de Erlangs. Esta distribuição multivariada possui propriedades muito convenientes para a otimização de carteiras. Entretanto, o esforço computacional necessário para calibrar os pesos e os parâmetros cresce muito rapidamente com o número de ativos e de parcelas. Este efeito é conhecido como “maldição da dimensionalidade”. Para diminuir o esforço computacional, os pesos do produto tensorial foram forçados a ser não-negativos; e os parâmetros de escala, a ser constantes para cada ativo. Com essas restrições, a calibração com o critério da máxima verossimilhança passa a ter uma estrutura especial, e os pesos, em sua maioria, serão iguais a zero. Mesmo assim, a maldição da dimensionalidade é forte no caso multiperiódico.

Se as distribuições de retornos dos ativos são definidas os reais não-negativos, e há ativos com peso negativo na carteira, então existe a possibilidade de falência. Neste caso, é necessário realizar uma análise paretiana em três dimensões: recompensa, risco e probabilidade de falência. O investidor deve fornecer valores aceitáveis para duas das três variáveis. No entanto, para simplificar a análise matemática deste trabalho, os pesos dos ativos nas carteiras são, por hipótese, não-negativos.

No caso uniperiódico, a combinação convexa de ativos dependentes cujos retornos discretos seguem

conjuntamente um produto tensorial finito de Erlangs é uma carteira cujos retornos discretos seguem uma mistura finita de Erlangs. Esta é uma propriedade muito importante, que pode ser provada com a transformada de Laplace. O cálculo da esperança dessas carteiras é simples, porém o cálculo do *Downside Risk* requer a avaliação numérica de um caso particular da função especial Meijer G. Este caso é o logarítmico com pólos no numerador e no denominador separados. Por isso, a avaliação numérica é ligeiramente menos complicada. Os lemas necessários para o cálculo numérico foram deduzidos e a representação da Meijer G em forma de série numérica foi dada.

O *Downside Risk* é uma medida de risco que não satisfaz a propriedade de invariância à translação. Entretanto, a falta desta propriedade parece não causar incoerência financeira. O *Downside Risk*, ao contrário de medidas de risco baseadas em quantis, não causa o problema de sub ou superestimação do risco, e isto é desejável para investidores racionais. Além disso, foi provado que ela é uma função convexa sobre os pesos não-negativos dos ativos, para determinados valores de seus parâmetros. Esta propriedade é muito desejável em problemas de minimização. E ela é necessária do ponto de vista financeiro, pois representa o princípio de que a diversificação não deve aumentar o risco.

O retorno discreto é uma variável aleatória do tipo aditivo em cada instante de tempo e do tipo multiplicativo ao longo do tempo. Os mundos de retornos discretos e de retornos contínuos são homeomorfos, e este conceito possibilita cálculos mais simples em alguns casos. Em outros, é mais conveniente trabalhar somente no mundo discreto com a transformada de Mellin.

A análise multiperiódica de carteiras é uma generalização dos casos com um e com infinitos períodos. O conceito de distribuição composta média permite uma transição suave entre esses dois casos. Este conceito foi generalizado para a distribuição composta média com longevidade. Com a mudança da medida de probabilidade para a composta média, o retorno e o risco podem ser calculados da mesma forma, qualquer que seja o horizonte de tempo. Cada ponto da fronteira eficiente multiperiódica representa uma ou mais seqüências ordenadas de carteiras.

Se a densidade dos retornos discretos uniperiódicos de uma carteira é uma mistura finita de Erlangs, então a densidade dos retornos de uma seqüência de carteiras, na medida de probabilidade composta média, será uma mistura finita de densidades Meijer G. Como as densidades são i.i.d. por hipótese, as carteiras de cada seqüência podem ser calculadas do futuro para o presente, o que reduz o esforço computacional.

Se o horizonte de tempo tende a infinito, então o critério de Kelly é obtido. Na medida de probabilidade composta média, a densidade dos retornos de qualquer seqüência de carteiras de pesos constantes tende a uma Delta de Dirac. Como as carteiras deste trabalho seguem misturas finitas de Erlangs, foi possível obter a expressão algébrica para a esperança dos logaritmos dos retornos. Com a maximização dessa expressão, obtém-se a carteira de Kelly. A fronteira eficiente torna-se um único ponto, e o restante da fronteira torna-se um segmento de uma curva. Se a medida de risco for o *Downside Risk*, então a fronteira serão, em geral, dois segmentos de reta. Na medida de probabilidade composta média, a carteira de Kelly independe da função utilidade, da medida de risco e do coeficiente de longevidade do investidor. Todos os

investidores devem escolhê-la.

Dados dois ou mais modelos de otimização de carteiras, um investidor tem interesse em conhecer seus desempenhos relativos. Para comparar de forma mais justa os diferentes modelos, é necessário levar a fronteira eficiente de cada modelo para um mesmo gráfico, escolher um certo nível de retorno esperado e, então, realizar um *backtest*. Com este procedimento, obtém-se a série temporal e o histograma dos retornos realizados das carteiras eficientes de cada modelo e das diferenças entre esses retornos. Estes gráficos são indicadores dos desempenhos relativos entre os modelos. Os testes devem ser repetidos para vários níveis de retorno esperado.

9.2. Contribuições originais

Até onde vai o conhecimento do autor, as contribuições a seguir são originais.

1. Uso de uma mistura finita de distribuições Erlang para aproximar a densidade univariada dos retornos uniperiódicos de um ativo; e de um produto tensorial finito de distribuições Erlang para aproximar a densidade multivariada dos retornos uniperiódicos de ativos dependentes.
2. Procedimento para a geração de pontos aleatórios de uma distribuição multivariada representada por um produto tensorial finito arbitrário.
3. **Lema 3.9.2, Lema 3.9.3, Lema 3.9.4 e Lema 3.9.6**, que resultam em uma avaliação numérica mais simples e precisa de um caso especial da função Meijer G.
4. **Lema 3.9.5**, que é uma simplificação de um resultado obtido por Mathai (1970).
5. **Teorema 3.10.1**, que é um resultado alternativo ao de Mathai (1982).
6. Expressões para o *Downside Risk* de um caso especial de distribuição envolvendo a função Meijer G.
7. **Teorema 3.11.1, Corolário 3.11.1** e reconhecimento que a combinação linear positiva de variáveis aleatórias que seguem conjuntamente um produto tensorial finito de Erlangs é uma variável aleatória que segue uma mistura finita de Erlangs.
8. **Lema 3.16.3**, que é uma generalização do limite (3.113).
9. Reconhecimento dos problemas de sub e superestimação dos retornos indesejáveis das medidas de risco baseadas em quantil.
10. **Teorema 4.4.1 e Corolário 4.4.1**, nos quais é provado que o *Downside Risk* e o *Lower Partial Moment* com $\alpha \geq 1$ e R_B constante são funções convexas sobre os pesos não-negativos dos ativos.
11. Reconhecimento que a escolha de carteiras capazes de causar a falência do investidor deve ser feita com uma abordagem paretiana com três dimensões: recompensa, risco e probabilidade de falência.
12. Reconhecimento que os mundos de retornos discretos e contínuos não irão convergir para um mesmo mundo quando o período de tempo tender a zero.

13. **Teorema 6.6.1**, no qual são dadas as expressões para a esperança e variância de uma variável aleatória multiplicativa em função do tempo contínuo.
14. Reconhecimento que o conceito de distribuição composta média é uma mudança na medida de probabilidade e que, no caso multiperifódico, seu uso é preferível ao da distribuição no instante final.
15. Generalização do conceito de distribuição composta média para permitir, mesmo no caso uniperifódico, a longevidade do investidor.
16. **Teorema 6.8.1**, no qual é provado que a esperança, na medida de probabilidade composta média, é côncava em relação aos pesos dos ativos.
17. **Conjectura 6.8.1**, no qual é conjecturado que o *Downside Risk*, para determinados parâmetros e na medida composta média, é (muito provavelmente) convexo sobre os pesos dos ativos.
18. **Teorema 7.2.1**, no qual é provado, com o auxílio do **Lema 3.16.3** e de transformadas de Mellin, que a distribuição composta média no limite quando o tempo tende a infinito é uma Delta de Dirac. Esta é uma prova alternativa de um resultado obtido por Hakansson (1971a).
19. **Prova 2** do critério de Kelly, usando o **Teorema 7.2.1** e uma propriedade das transformadas de Mellin.
20. Reconhecimento que, para o caso com infinitos períodos, a fronteira das carteiras de pesos constantes, na medida de probabilidade composta média, será um segmento de uma curva, e não um compacto, qualquer que seja a medida de risco.
21. Reconhecimento que, sob a hipótese i.i.d., a carteira de Kelly deve ser escolhida por todos os investidores com horizonte de tempo tendendo a infinito, independentemente de sua função utilidade ou medida de risco.
22. Reconhecimento que, no caso de haver empate entre duas carteiras pelo critério de Kelly, não adianta calcular as derivadas superiores da transformada de Mellin no ponto $s = 1$.
23. Estabelecimento de um critério para a seleção de um dentre vários modelos de otimização de carteiras.

9.3. Limitações e sugestões para pesquisas futuras

O reconhecimento das limitações de um modelo ou pesquisa é o primeiro passo em direção à sua evolução. As hipóteses de um modelo devem ser claramente enunciadas para facilitar o leitor dizer se ele pode ser usado. O modelo também deve ser analisado com base no ponto de vista de teorias existentes de campos relacionados. Tudo isto motivará a criação de novos testes, ferramentas, trabalhos relacionados em outras áreas e, finalmente, novos modelos que irão suceder o proposto.

Como usualmente acontece, esta dissertação gerou muitas questões relacionadas aos tópicos abordados. As dúvidas e limitações são normais em trabalhos científicos e devem servir para motivar pesquisas futuras.

A principal limitação dos modelos propostos é o grande esforço computacional requerido. Para minimizar este problema, é necessário pesquisar novas classes de distribuições com as propriedades adequadas, novos algoritmos de otimização e novas propriedades e generalizações da função Fox H.

Uma lista de sugestões para trabalhos futuros é dada a seguir.

1. Comparação entre os critérios do erro quadrático médio e da máxima verossimilhança para a aproximação de densidades.
2. Inclusão de ativos com pesos negativos na carteira, de uma barreira absorvente no lado esquerdo e de uma árvore com distribuições truncadas e Deltas de Dirac.
3. Inclusão de ativos com probabilidade de retornos negativos.
4. Desenvolvimento de algoritmos de otimização sob medida para resolver os problemas propostos.
5. Estudo de distribuições estáveis e infinitamente divisíveis nos casos aditivo e multiplicativo. Qual a relação entre estas distribuições e a mistura finita de Erlangs e Meijer G?
6. Estudo das distribuições fechadas à combinação linear e das condições que tornam fechado o produto de funções Fox H.
7. Estudo de uma distribuição baseada em Curvas de Bézier (ou mistura de distribuições Beta).
8. Avaliação numérica mais robusta do caso especial da função Meijer G usada neste trabalho.
9. Inclusão de consumo, em dois casos: quantias predeterminadas ou porcentual da riqueza em cada período.
10. Inclusão de custos de transação, em dois casos: valores fixos e variáveis.
11. Estudo das propriedades desejáveis e relações entre medidas de risco, medidas de recompensa, dominância estocástica, função utilidade e distribuição composta média, incluindo o caso multi-periódico. Esses critérios podem ser unificados com axiomas?
12. Inclusão de dependência temporal por um ou mais períodos.
13. Inclusão de um ativo livre de risco. Estudo do teorema da separação e do CAPM uni e multi-periódico.
14. Estudo da relação entre novos modelos financeiros e a hipótese de mercado eficiente.
15. Estudo de carteiras em tempo contínuo usando Cálculo Fracionário.
16. Estudo das possíveis relações entre as teorias da informação e de otimização funcional com a distribuição composta média.
17. Inclusão de derivativos na carteira. Estudo de distribuições livres de arbitragem.
18. Realização de um teste de hipótese de igualdade de retorno médio dos modelos de otimização de carteiras.

APÊNDICE A – CADERNOS MATHEMATICA

Exemplo 3.5.1

```
(* MThesisExErlangMixture.nb - Fernando Hideki Kato - Master Degree Thesis example *)

<< Graphics`Graphics`; << Graphics`Legend`; << Statistics`ContinuousDistributions`

Off[General::"spell1", General::"spell", Integrate::"gener"]

ErlangPDF[a_, b_, x_] :=  $\frac{a^b x^{b-1}}{\text{Gamma}[b] e^{ax}}$  (* a>0, b∈{1,2,...}, x≥0 *)

ErlangCDF[a_, b_, x_] := Evaluate[FullSimplify[ $\int_0^x \text{ErlangPDF}[a, b, t] dt$ , {a>0, b>0, x>0}]]

a := 440; b1 := 440; b2 := 450; b3 := 525; w1 := 0.3; w2 := 0.3; w3 := 0.4

OrigMixturePDF[x_] := Evaluate[w1 ErlangPDF[a, b1, x] + w2 ErlangPDF[a, b2, x] + w3 ErlangPDF[a, b3, x]]

OrigMixtureCDF[x_] := Evaluate[w1 ErlangCDF[a, b1, x] + w2 ErlangCDF[a, b2, x] + w3 ErlangCDF[a, b3, x]]

Plot[OrigMixturePDF[x], {x, 0.8, 1.4}, PlotRange → {0, 6}]

Plot[OrigMixtureCDF[x], {x, 0.8, 1.4}, PlotRange → All]

SeedRandom[163]; S := 40 (*Size of sample or number of random points*);
unifRandomArray := Evaluate[RandomArray[UniformDistribution[0, 1], S]]

points := Evaluate[Table[FindRoot[OrigMixtureCDF[x] == unifRandomArray[[i]], {x, 0.8, 1.4}], {i, 1, S}]]

pointArray := x /. points

Histogram[pointArray,
  HistogramCategories → {0.8, 0.85, 0.9, 0.95, 1, 1.05, 1.1, 1.15, 1.2, 1.25, 1.3, 1.35, 1.4, 1.45, 1.5}]

pointArray

{1.26011, 1.07765, 1.04734, 1.19546, 1.0056, 1.12446, 1.23907, 1.14188, 1.0447, 1.23796,
 1.04293, 1.0348, 1.21076, 1.0283, 1.20057, 0.942018, 1.01565, 0.954768, 1.13454, 0.974394,
 1.01519, 1.10049, 1.03416, 0.980785, 1.02368, 0.93923, 1.25164, 1.02166, 0.959291, 1.03477,
 0.925508, 1.15479, 1.1276, 1.05531, 1.09177, 1.00624, 1.18584, 0.965531, 1.17679, 0.996574}

C525MixturePDF[x_] :=
  Evaluate[0.130028 ErlangPDF[443.367220, 448, x] + 0.510705 ErlangPDF[443.367220, 449, x] +
    0.359267 ErlangPDF[443.367220, 525, x]] (*LogLikelihood=42.306195*)

Plot[C525MixturePDF[x], {x, 0.8, 1.4}, PlotRange → {0, 6}]

C750MixturePDF[x_] :=
  Evaluate[0.191680 ErlangPDF[621.404696, 607, x] + 0.371213 ErlangPDF[621.404696, 636, x] +
    0.055873 ErlangPDF[621.404696, 637, x] + 0.149367 ErlangPDF[621.404696, 706, x] +
    0.231866 ErlangPDF[621.404696, 750, x]] (*LogLikelihood=42.588620*)
```



```

Plot[C750MixturePDF[x], {x, 0.8, 1.4}, PlotRange -> {0, 6}]

tic := Union[{0.8, 0.9, {1.0000000000001, "1.0"}, 1.1, 1.2, 1.3, 1.4},
  Table[{pointArray[i]}, {"", 0.02}, {i, 1, Length[pointArray]}]];
Plot[0, {x, 0.8, 2}, Ticks -> {tic, Automatic}]

fs := 30; Plot[{OrigMixturePDF[x], C525MixturePDF[x], C750MixturePDF[x]}, {x, 0.8, 1.4}, AxesLabel -> {"x", "f(x)"},
  Ticks -> {tic, Automatic}, AxesOrigin -> {1, 0}, AspectRatio -> 0.4, TextStyle -> {FontSize -> fs},
  PlotStyle -> {AbsoluteThickness[4], AbsoluteThickness[1], Dashing[{.01]}], PlotLegend -> {"forig", "f525", "f750"},
  LegendPosition -> {0.3, 0.08}, LegendOrientation -> Horizontal, LegendSize -> {0.4, 0.2},
  LegendShadow -> {.005, -.005}];

```

Exemplo 3.5.2

```

(* MThesisExErlangTensor.nb - Fernando Hideki Kato - Master Degree Thesis example *)

<< Statistics`ContinuousDistributions`; << Graphics`Graphics`; << Graphics`Graphics3D`

Off[General::"spell1", General::"spell", Integrate::"gener"]

ErlangPDF[a_, b_, x_] := 
$$\frac{a^b x^{b-1}}{\Gamma[b] e^{ax}} \quad (* a > 0, b \in \{1, 2, \dots\}, x \geq 0 *)$$


ErlangCDF[a_, b_, x_] := Evaluate[FullSimplify[ $\int_0^x \text{ErlangPDF}[a, b, t] dt$ , {a > 0, b > 0, x > 0}]]

a := {440, 300}; b := {{440, 300}, {500, 375}, {550, 300}}; w := {0.4, 0.3, 0.3}; DD := Dimensions[a][[1]];
n := Dimensions[w][[1]];

$$\left\{ \sum_{i=1}^n w[[i]], \text{Dimensions}[b][[1]], n, DD \right\}$$


OrigMixturePDF[x_, y_] := Evaluate[ $\sum_{i=1}^n w[[i]] \text{ErlangPDF}[a[[1]], b[[i, 1]], x] \text{ErlangPDF}[a[[2]], b[[i, 2]], y]$ ]

Plot3D[OrigMixturePDF[x, y], {x, 0.8, 1.4}, {y, 0.8, 1.4}, PlotPoints -> 30, PlotRange -> All]

RndErlangTensorProduct[D_, a_, b_, r_] :=
  Table[x /. FindRoot[ErlangCDF[a[[d]], b[[d]], x] == r[[d]], {x, 0.5, 2}], {d, 1, D}]

RndErlangTensorProduct[2, {440, 300}, {440, 300}, {0.5, 0.5}]

RndComponent[nWeights_, weights_, r_] :=
  (For[i = 1; t = weights[[1]], (i < nWeights && t <= r), t += weights[++i]]; Return[i])

RndComponent[n, w, 1]

S := 60; SeedRandom[203]; rndUnifArrayDD := Evaluate[RandomArray[UniformDistribution[0, 1], {S, DD}]];
rndComponentArray := Evaluate[Table[RndComponent[n, w, Random[]], {i, 1, S}]]

points := Evaluate[Table[RndErlangTensorProduct[DD, a, b[[rndComponentArray[[s]]]], rndUnifArrayDD[[s]],
  {s, 1, S}]];
points

```

```

{{1.11318, 1.25238}, {1.1983, 0.919568}, {1.10829, 1.05851}, {1.2716, 0.963914}, {1.21433, 1.02264},
{0.951162, 1.00405}, {1.31888, 0.983074}, {1.19662, 1.3484}, {1.09731, 1.269}, {0.968682, 1.07306},
{0.920063, 1.01312}, {1.03844, 1.0907}, {1.27983, 0.918293}, {1.26636, 1.06252}, {1.02859, 0.957023},
{1.22931, 1.26366}, {0.996247, 1.00945}, {1.04189, 0.909297}, {1.25929, 0.947192}, {1.32961, 0.932891},
{0.981206, 1.0075}, {1.15834, 1.20094}, {0.975511, 1.10494}, {1.0563, 0.966789}, {1.14751, 1.32855},
{1.2144, 0.9883}, {1.05234, 1.02222}, {0.973218, 1.05658}, {1.12126, 1.16186}, {1.13301, 1.29986},
{1.04235, 1.07558}, {1.25587, 1.02307}, {1.24418, 0.975813}, {1.19575, 1.23843}, {0.951507, 0.981555},
{1.28553, 1.18282}, {1.00669, 0.998096}, {1.1691, 1.28718}, {1.16226, 1.31378}, {1.24286, 1.06971},
{0.975843, 1.03927}, {1.23602, 1.16895}, {1.13516, 1.31767}, {0.990045, 0.990421}, {1.15478, 1.0002},
{1.25534, 0.861009}, {1.19637, 1.17697}, {1.00498, 1.06186}, {0.976875, 1.01866}, {1.04072, 0.936822},
{1.15112, 1.30271}, {1.11648, 1.20609}, {1.22288, 0.900946}, {0.984739, 0.939897}, {1.04213, 0.948476},
{1.34274, 1.03063}, {1.29051, 1.02579}, {1.23731, 1.0296}, {1.06102, 1.00619}, {1.05234, 1.0582}}

```

```

ListPlot[points, Frame → True, PlotRange → {{0.8, 1.4}, {0.8, 1.4}}]

a550 := {438.260016, 431.159565};
b550 := {{441, 441}, {449, 418}, {450, 417}, {506, 550}, {507, 550}, {509, 521}, {541, 511}, {542, 511},
{550, 399}, {550, 433}};
w550 := {0.343068, 0.040654, 0.034114, 0.118423, 0.084736, 0.039509, 0.039212, 0.002045, 0.087496, 0.210742};
n550 := Dimensions[w550][[1]];

$$\sum_{i=1}^{n550} w550[[i]] (* \text{LogLikelihood}=125.292272 *)$$

M550MixturePDF[x_, y_] :=
  Evaluate[ $\sum_{i=1}^{n550} w550[[i]] \text{ErlangPDF}[a550[[1]], b550[[i, 1]], x] \text{ErlangPDF}[a550[[2]], b550[[i, 2]], y]$ ];

Plot3D[M550MixturePDF[x, y], {x, 0.8, 1.4}, {y, 0.8, 1.4}, PlotPoints → 30, PlotRange → All]

a750 := {595.765349, 582.439109};
b750 := {{590, 598}, {600, 610}, {613, 559}, {629, 613}, {685, 700}, {688, 750}, {695, 587}, {738, 691},
{739, 691}, {748, 537}, {749, 538}, {750, 540}, {750, 589}, {750, 590}};
w750 := {0.236608, 0.018181, 0.110238, 0.052319, 0.058447, 0.178945, 0.009686, 0.006733, 0.039345,
0.076265, 0.019596, 0.007453, 0.087619, 0.098564}; n750 := Dimensions[w750][[1]];

$$\sum_{i=1}^{n750} w750[[i]] (* \text{LogLikelihood}=128.840960 *)$$

M750MixturePDF[x_, y_] :=
  Evaluate[ $\sum_{i=1}^{n750} w750[[i]] \text{ErlangPDF}[a750[[1]], b750[[i, 1]], x] \text{ErlangPDF}[a750[[2]], b750[[i, 2]], y]$ ];

Plot3D[M750MixturePDF[x, y], {x, 0.8, 1.4}, {y, 0.8, 1.4}, PlotPoints → 30, PlotRange → All]

fs := 26; tic3D := {{0.8, {1.00000001, "1.0"}, 1.2, 1.4}, {0.8, {1.00000001, "1.0"}, 1.2, 1.4}, Automatic};
tic2D := {{0.8000000001, "0.8"}, {1.0000000001, "1.0"}, 1.2, 1.4},
{{0.8000000001, "0.8"}, {1.0000000001, "1.0"}, 1.2, 1.4}, {}, {}

g1 := Plot3D[OrigMixturePDF[x, y], {x, 0.8, 1.4}, {y, 0.8, 1.4}, PlotPoints → 40, PlotRange → {0, 25},
  PlotLabel → "morig", AxesLabel → {"x1", "x2", ""}, Ticks → tic3D, TextStyle → {FontSize → fs},
  DisplayFunction → Identity];
Show[GraphicsArray[{g1}]]

g2 := ListPlot[points, Frame → True, FrameLabel → {"x1", "x2", "Random Points", ""}, RotateLabel → False,
  TextStyle → {FontSize → fs}, PlotRange → {{0.8, 1.4}, {0.8, 1.4}}, Axes → True, FrameTicks → tic2D,
  DisplayFunction → Identity];
Show[GraphicsArray[{g2}]]

g3 := Plot3D[M550MixturePDF[x, y], {x, 0.8, 1.4}, {y, 0.8, 1.4}, PlotPoints → 40, PlotRange → {0, 25},
  PlotLabel → "m550", AxesLabel → {"x1", "x2", ""}, Ticks → tic3D, TextStyle → {FontSize → fs},
  DisplayFunction → Identity];
Show[GraphicsArray[{g3}]]

g4 := Plot3D[M750MixturePDF[x, y], {x, 0.8, 1.4}, {y, 0.8, 1.4}, PlotPoints → 40, PlotRange → {0, 25},
  PlotLabel → "m750", AxesLabel → {"x1", "x2", ""}, Ticks → tic3D, TextStyle → {FontSize → fs},
  DisplayFunction → Identity];
Show[GraphicsArray[{g4}]]

Show[GraphicsArray[Table[{g1, g2}, {g3, g4}], GraphicsSpacing → 0.03]]

```

Exemplo 3.9.1

```
(* MThesisExNumEvalMeijerG.nb - Fernando Hideki Kato - Master Degree Thesis example *)
(* The formulas below work for the conditions given in the thesis. *)
(* In this implementation, parameters  $b_i \in \text{Integers}, 1 \leq i \leq m$  *)

<< Utilities 'MemoryConserve';

Off[MemoryConserve::"start", MemoryConserve::"end"]

PsiGen[d_, n_] := PolyGamma[d, 1] + d!  $\sum_{i=1}^n i^{-d-1}$ 

LimitSpVrKProdGamma[k_, v_, a1_, a2_, b1_, b2_] :=

$$\left( \prod_{i=1}^k \frac{(-1)^{v-b1[[i]]}}{(v-b1[[i]])!} \right) \left( \prod_{ib1=k+1}^{\text{Length}[b1]} \text{Gamma}[b1[[ib1]] - v] \right)$$


$$\left( \prod_{ia1=1}^{\text{Length}[a1]} \text{Gamma}[1 - a1[[ia1]] + v] \right) / \left( \prod_{ia2=1}^{\text{Length}[a2]} \text{Gamma}[a2[[ia2]] - v] \right) / \left( \prod_{ib2=1}^{\text{Length}[b2]} \text{Gamma}[1 - b2[[ib2]] + v] \right)$$


LimitDtDstInSpVrKProdGamma[t_, k_, v_, a1_, a2_, b1_, b2_] :=

$$\sum_{i=1}^k \text{PsiGen}[t-1, v-b1[[i]]] + \sum_{ib1=k+1}^{\text{Length}[b1]} \text{PolyGamma}[t-1, b1[[ib1]] - v] + (-1)^{t-1} \sum_{ib2=1}^{\text{Length}[b2]} \text{PolyGamma}[t-1, 1-b2[[ib2]] + v] +$$


$$(-1)^t \sum_{ia1=1}^{\text{Length}[a1]} \text{PolyGamma}[t-1, 1-a1[[ia1]] + v] - \sum_{ia2=1}^{\text{Length}[a2]} \text{PolyGamma}[t-1, a2[[ia2]] - v]$$


MeijerResidueM[t_, k_, v_, a1_, a2_, b1_, b2_] :=

$$\frac{1}{t} \sum_{u=0}^{t-1} \frac{\text{LimitDtDstInSpVrKProdGamma}[t-u, k, v, a1, a2, b1, b2]}{(t-u-1)!} \text{MeijerResidueM}[u, k, v, a1, a2, b1, b2];$$


MeijerResidueM[0, k_, v_, a1_, a2_, b1_, b2_] = 1;

PoleOrderB1[v_, b1_] := (For[i = Length[b1]; k = 0, i > 0, k += If[(-v + b1[[i]]) ≤ 0, 1, 0]; i--]; Return[k])

PowerAux[x_, y_] := If[y == 0, 1, x^y]

MeijerGAux1[a1_, a2_, b1_, b2_, x_, nPoles_, k_, v_] :=

$$x^v \text{LimitSpVrKProdGamma}[k, v, a1, a2, b1, b2] \sum_{t=0}^{k-1} \left( \frac{\text{PowerAux}[-\text{Log}[x], k-t-1]}{(k-t-1)!} \text{MeijerResidueM}[t, k, v, a1, a2, b1, b2] \right)$$


MeijerGAux2[a1_, a2_, b1_, b2_, x_, initialPole_, nPoles_] :=

$$\sum_{v=\text{initialPole}}^{\text{initialPole} + \text{nPoles} - 1} (\text{MeijerGAux1}[a1, a2, b1, b2, x, \text{nPoles}, \text{PoleOrderB1}[v, b1], v])$$


MeijerGKato[a1_, a2_, b1_, b2_, x_, nPoles_] := MeijerGAux2[a1, a2, b1, b2, x, Min[b1], nPoles]

x := 30; nPoles := 50; nDigits := 40; a1 := {-1, -3}; a2 := {5/2}; b1 := {1, 2, 2, 5}; b2 := {-1, -2};
Table[{Block[{$MaxExtraPrecision = 500}, N[MeijerG[{a1, a2}, {b1, b2}, x], nDigits]],
  N[MeijerGKato[a1, a2, b1, b2, x, nPoles], nDigits]}]

{2.900754807997581278090100785981192163674, 2.900754807997581278090100785996539700500}

x := 30; nPoles := 50; nDigits := 40; a1 := {}; a2 := {}; b1 := {1, 2, 2, 2, 3}; b2 := {};
Table[{Block[{$MaxExtraPrecision = 500}, N[MeijerG[{a1, a2}, {b1, b2}, x], nDigits]],
  N[MeijerGKato[a1, a2, b1, b2, x, nPoles], nDigits]}]

{-154345.3237684791915447372667669382750391, 0.3412453756194701072311323888089946919052}

(* The above result from Mathematica 4.2 and 5.0 is incorrect because this Meijer G function represents
a PDF given by the multiplicative or scale convolution of Erlang PDFs and thus cannot assume negative
values. See the thesis. *)
```

Exemplo 3.10.1

```
(* MThesisExErlangAddConv.nb - Fernando Hideki Kato - Master Degree Thesis example *)

Off[General::"spell1", General::"spell"]

ErlangPDF[a_, b_, x_] :=  $\frac{a^b x^{b-1}}{\text{Gamma}[b] e^{ax}}$  (* a>0, b∈{1,2,...}, x≥0 *)

ErlangAdditiveConvolutionM[a_, b_, D_, t_, j_] :=
 $\frac{1}{t} \sum_{m=0}^{t-1} \left( \left( \sum_{i=1}^D \text{If}[i == j, 0, b[[i]] (a[[j]] - a[[i]])^{m-t}] \right) \right) \text{ErlangAdditiveConvolutionM}[a, b, D, m, j];$ 

ErlangAdditiveConvolutionM[a_, b_, D_, 0, j_] = 1;

ErlangAdditiveConvolutionAux1[a_, b_, D_] :=
 $\sum_{j=1}^D \left( \left( \sum_{i=1}^D \text{If}[i == j, 1, \left( 1 - \frac{a[[j]]}{a[[i]]} \right)^{-b[[i]]}] \right) \right.$ 
 $\left. \left( \sum_{r=1}^{b[[j]]} (a[[j]]^{b[[j]]-r} \text{ErlangAdditiveConvolutionM}[a, b, D, b[[j]] - r, j] \text{ErlangPDFAux}[a[[j]], r, x]) \right) \right)$ 

ErlangAdditiveConvolution[a_, b_] := Expand[ErlangAdditiveConvolutionAux1[a, b, Length[a]]];
(* There can't be equal elements in a. *)

ErlangAdditiveConvolution[{4, 5}, {2, 3}]

-1500 ErlangPDFAux[4, 1, x] + 125 ErlangPDFAux[4, 2, x] +
1200 ErlangPDFAux[5, 1, x] + 160 ErlangPDFAux[5, 2, x] + 16 ErlangPDFAux[5, 3, x]

c := 3/2;
N[{ $\int_0^c \text{ErlangPDF}[4, 2, t] \text{ErlangPDF}[5, 3, c-t] dt$ ,
(-1500 ErlangPDF[4, 1, c] + 125 ErlangPDF[4, 2, c] + 1200 ErlangPDF[5, 1, c] + 160 ErlangPDF[5, 2, c] +
16 ErlangPDF[5, 3, c])}, 25]

{0.4451957446075532900679990, 0.4451957446075532900679990}

ErlangAdditiveConvolution[{4, 5, 6}, {3, 3, 4}]

2349000 ErlangPDFAux[4, 1, x] - 202500 ErlangPDFAux[4, 2, x] +
10125 ErlangPDFAux[4, 3, x] - 8294400 ErlangPDFAux[5, 1, x] +
414720 ErlangPDFAux[5, 2, x] - 82944 ErlangPDFAux[5, 3, x] + 5346000 ErlangPDFAux[6, 1, x] +
432000 ErlangPDFAux[6, 2, x] + 27000 ErlangPDFAux[6, 3, x] + 1000 ErlangPDFAux[6, 4, x]

c := 1; N[{ $\int_0^c \int_0^c \int_0^{c-t-u} (\text{ErlangPDF}[4, 3, t] \text{ErlangPDF}[5, 3, u] \text{ErlangPDF}[6, 4, v]) dv du dt$ ,
 $\int_0^c (2349000 \text{ErlangPDF}[4, 1, x] - 202500 \text{ErlangPDF}[4, 2, x] + 10125 \text{ErlangPDF}[4, 3, x] -$ 
8294400 ErlangPDF[5, 1, x] + 414720 ErlangPDF[5, 2, x] - 82944 ErlangPDF[5, 3, x] +
5346000 ErlangPDF[6, 1, x] + 432000 ErlangPDF[6, 2, x] + 27000 ErlangPDF[6, 3, x] +
1000 ErlangPDF[6, 4, x]) dx}, 25]

{0.03175220810273944510082946, 0.03175220810273944510082946}
```

Exemplo 3.11.1

(* MThesisExErlangLinearComb.nb - Fernando Hideki Kato - Master Degree Thesis example *)

Off[General::"spell1", General::"spell"]

ErlangPDF[a_, b_, x_] := $\frac{a^b x^{b-1}}{\text{Gamma}[b] e^{ax}}$ (* a>0, b∈{1,2,...}, x≥0 *)

ErlangAdditiveConvolutionM[a_, b_, D_, t_, j_] :=

$\frac{1}{t} \sum_{m=0}^{t-1} \left(\left(\sum_{i=1}^D \text{If}[i == j, 0, b[[i]] (a[[j]] - a[[i]])^{m-t}] \right) \right) \text{ErlangAdditiveConvolutionM}[a, b, D, m, j];$

ErlangAdditiveConvolutionM[a_, b_, D_, 0, j_] = 1;

ErlangAdditiveConvolutionAux1[a_, b_, D_] :=

$\sum_{j=1}^D \left(\left(\prod_{i=1}^D \text{If}[i == j, 1, \left(1 - \frac{a[[j]]}{a[[i]]} \right)^{-b[[i]]}] \right) \left(\sum_{r=1}^{b[[j]]} (a[[j]])^{b[[j]]-r} \text{ErlangAdditiveConvolutionM}[a, b, D, b[[j]] - r, j] \text{ErlangPDFAux}[a[[j]], r, x] \right) \right)$

ErlangAdditiveConvolution[a_, b_] := Expand[ErlangAdditiveConvolutionAux1[a, b, Length[a]]];

(* There can't be equal elements in a. *)

f[x1_, x2_] := $\frac{4}{10} \text{ErlangPDF}[4, 2, x1] \text{ErlangPDF}[5, 3, x2] + \frac{6}{10} \text{ErlangPDF}[6, 3, x1] \text{ErlangPDF}[3, 2, x2]$

G[t_, w1_, w2_] := Evaluate[$\int_0^{t/w1} \left(\int_0^{\frac{t-w1x1}{w2}} f[x1, x2] dx2 \right) dx1$]; g[t_, w1_, w2_] := Evaluate[$\partial_t G[t, w1, w2]$];

ta := 12/10; w1a := 2/10; w2a := 1/2; N[{g[ta, w1a, w2a], G[ta, w1a, w2a]}, 30]

{0.0388725011703947123665221495889, 0.992872450533433633999047511650}

Expand[$\frac{4}{10} \text{ErlangAdditiveConvolution}[\{4/w1a, 5/w2a\}, \{2, 3\}] + \frac{6}{10} \text{ErlangAdditiveConvolution}[\{6/w1a, 3/w2a\}, \{3, 2\}]$]

$-\frac{225}{256} \text{ErlangPDFAux}[6, 1, x] + \frac{75}{64} \text{ErlangPDFAux}[6, 2, x] + \frac{24}{5} \text{ErlangPDFAux}[10, 1, x] - \frac{16}{5} \text{ErlangPDFAux}[10, 2, x] + \frac{8}{5} \text{ErlangPDFAux}[10, 3, x] - \frac{12}{5} \text{ErlangPDFAux}[20, 1, x] - \frac{2}{5} \text{ErlangPDFAux}[20, 2, x] + \frac{45}{256} \text{ErlangPDFAux}[30, 1, x] + \frac{3}{32} \text{ErlangPDFAux}[30, 2, x] + \frac{3}{80} \text{ErlangPDFAux}[30, 3, x]$

$h[x_] := -\frac{225}{256} \text{ErlangPDF}[6, 1, x] + \frac{75}{64} \text{ErlangPDF}[6, 2, x] + \frac{24}{5} \text{ErlangPDF}[10, 1, x] - \frac{16}{5} \text{ErlangPDF}[10, 2, x] + \frac{8}{5} \text{ErlangPDF}[10, 3, x] - \frac{12}{5} \text{ErlangPDF}[20, 1, x] - \frac{2}{5} \text{ErlangPDF}[20, 2, x] + \frac{45}{256} \text{ErlangPDF}[30, 1, x] + \frac{3}{32} \text{ErlangPDF}[30, 2, x] + \frac{3}{80} \text{ErlangPDF}[30, 3, x];$

H[x_] := Evaluate[$\int_0^x h[t] dt$]

N[{h[ta], H[ta]}, 30]

{0.0388725011703947123665221495889, 0.992872450533433633999047511650}

Simplify[g[x, w1a, w2a] == h[x]]

True

Exemplo 6.7.1

```
(* MThesisExMultiPeriodDirac.nb - Fernando Hideki Kato - Master Degree Thesis example *)

<< Graphics`MultipleListPlot`; << Graphics`Legend`; << Graphics`Arrow`

MixtureRnd[p_] :=
  Block[{r, n, i, t}, {r = Random[]; n = Length[p]; For[i = 1; t = p[[1]], (i < n && t ≤ r), t += p[[++i]]]; Return[i]}]

ListProduct[v_] := Block[{ret, i}, {ret = 1; For[i = Length[v], i > 0, ret *= v[[i--]]]; Return[ret]}]

DiracMellin[m_, s_] := m(s-1)

DiracMixturePlot[pm_] := Table[Arrow[{pm[[2, i]], 0}, {pm[[2, i]], pm[[1, i]]}, HeadScaling → Absolute],
  {i, 1, Length[pm[[1]]]}]

DiracMixtureMellin[pm_, s_] :=  $\sum_{i=1}^{\text{Length}[pm[[1]]]}$  (pm[[1, i]] DiracMellin[pm[[2, i]], s])

DiracMixtureRnd[pm_] := pm[[2, MixtureRnd[pm[[1]]]]]

pm1 := {{0.9, 0.1}, {1, 50}}; pm2 := {{1}, {5}}; n := 100; fs := 20

gd := Plot[{-1, -1}, {x, 0, 50},
  Epilog → {AbsoluteThickness[1], DiracMixturePlot[pm1], AbsoluteThickness[3], DiracMixturePlot[pm2]},
  AxesLabel → {"x", "PDF(x)"}, PlotRange → {0, 1}, PlotStyle → {AbsoluteThickness[1], AbsoluteThickness[3]},
  Ticks → {{{0.000000001, "0"}, {1, ""}, {2, ""}, {3, ""}, {4, ""}, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50},
    {{0.000000001, "0.0"}, {0.1, ""}, 0.2, {0.3, ""}, 0.4, {0.5, ""}, 0.6, {0.7, ""}, 0.8, {0.9, ""}, {1, "1.0"}}},
  TextStyle → {FontSize → fs}, PlotLegend → {"1", "2"}, LegendPosition → {0.1, 0.08},
  LegendOrientation → Vertical, LegendSize → {0.4, 0.3}, LegendShadow → {.005, -.005}, LegendTextSpace → 1,
  DisplayFunction → Identity];
Show[GraphicsArray[{gd}]]

gm := Plot[{DiracMixtureMellin[pm1, s], DiracMixtureMellin[pm2, s]}, {s, 1, 2}, PlotRange → All,
  AxesOrigin → {1, 1}, AxesLabel → {"s", "F(s)"}, PlotStyle → {AbsoluteThickness[1], AbsoluteThickness[3]},
  Ticks → {{{1.0000001, "1.0"}, {1.1, ""}, 1.2, {1.3, ""}, 1.4, {1.5, ""}, 1.6, {1.7, ""}, 1.8, {1.9, ""}, {2, "2.0"}},
    {{1.0000001, "1"}, 2, 3, 4, 5, 6}}, TextStyle → {FontSize → fs}, PlotLegend → {"1", "2"},
  LegendPosition → {-0.1, 0.08}, LegendOrientation → Vertical, LegendSize → {0.4, 0.3},
  LegendShadow → {.005, -.005}, LegendTextSpace → 1, DisplayFunction → Identity];
Show[GraphicsArray[{gm}]]

Show[GraphicsArray[{gd, gm}]]

t1 := 15; SeedRandom[211];
list11 := Evaluate[Table[ListProduct[Table[DiracMixtureRnd[pm1], {t, 1, t1}]], {i, 1, n}]];
list12 := Evaluate[Table[ListProduct[Table[DiracMixtureRnd[pm2], {t, 1, t1}]], {i, 1, n}]];
g1 := MultipleListPlot[list11, list12, PlotRange → {0, 25t1}, PlotJoined → True, AxesOrigin → {0, 0},
  PlotStyle → {AbsoluteThickness[1], AbsoluteThickness[3]}, DisplayFunction → Identity];
Show[GraphicsArray[{g1}]]

t1 := 1; SeedRandom[211]; list11 := Evaluate[Table[ListProduct[Table[DiracMixtureRnd[pm1], {t, 1, t1}]], {i, 1, n}]];
list12 := Evaluate[Table[ListProduct[Table[DiracMixtureRnd[pm2], {t, 1, t1}]], {i, 1, n}]];
g1 := MultipleListPlot[list11, list12, PlotRange → {0, 100}, PlotJoined → True, AxesOrigin → {0, 0},
  PlotLabel → "Monte Carlo, T=" <> ToString[t1], AxesLabel → {"n", "Return"},
  PlotStyle → {AbsoluteThickness[1], AbsoluteThickness[3]}, TextStyle → {FontSize → fs},
  PlotLegend → {"1", "2"}, LegendPosition → {0, 0.15}, LegendOrientation → Vertical,
  LegendSize → {0.4, 0.3}, LegendShadow → {.005, -.005}, LegendTextSpace → 1, DisplayFunction → Identity];
Show[GraphicsArray[{g1}]]
```

```

t2 := 3; SeedRandom[163];
list21 := Evaluate[Table[ListProduct[Table[DiracMixtureRnd[pm1], {t, 1, t2}]], {i, 1, n}]];
list22 := Evaluate[Table[ListProduct[Table[DiracMixtureRnd[pm2], {t, 1, t2}]], {i, 1, n}]];
g2 := MultipleListPlot[list21, list22, PlotRange -> {0, 1000}, PlotJoined -> True, AxesOrigin -> {0, 0},
  PlotLabel -> "Monte Carlo, T=" <> ToString[t2], AxesLabel -> {"n", "Return "},
  PlotStyle -> {AbsoluteThickness[1], AbsoluteThickness[3]}, TextStyle -> {FontSize -> fs},
  PlotLegend -> {"1)", "2) "}, LegendPosition -> {-0.5, 0.15}, LegendOrientation -> Vertical,
  LegendSize -> {0.4, 0.3}, LegendShadow -> {.005, -.005}, LegendTextSpace -> 1, DisplayFunction -> Identity];
Show[GraphicsArray[{g2}]]

t3 := 8; SeedRandom[263]; list31 := Evaluate[Table[ListProduct[Table[DiracMixtureRnd[pm1], {t, 1, t3}]], {i, 1, n}]];
list32 := Evaluate[Table[ListProduct[Table[DiracMixtureRnd[pm2], {t, 1, t3}]], {i, 1, n}]];
g3 := MultipleListPlot[list31, list32, PlotRange -> {0, 7 10^6}, PlotJoined -> True, AxesOrigin -> {0, 0},
  PlotLabel -> "Monte Carlo, T=" <> ToString[t3], AxesLabel -> {"n", "Return "},
  PlotStyle -> {AbsoluteThickness[1], AbsoluteThickness[3]}, TextStyle -> {FontSize -> fs},
  PlotLegend -> {"1)", "2) "}, LegendPosition -> {-0.5, 0.15}, LegendOrientation -> Vertical,
  LegendSize -> {0.4, 0.3}, LegendShadow -> {.005, -.005}, LegendTextSpace -> 1, DisplayFunction -> Identity];
Show[GraphicsArray[{g3}]]

t4 := 10; SeedRandom[333];
list41 := Evaluate[Table[ListProduct[Table[DiracMixtureRnd[pm1], {t, 1, t4}]], {i, 1, n}]];
list42 := Evaluate[Table[ListProduct[Table[DiracMixtureRnd[pm2], {t, 1, t4}]], {i, 1, n}]];
g4 := MultipleListPlot[list41, list42, PlotRange -> {0, 10^7}, PlotJoined -> True, AxesOrigin -> {0, 0},
  PlotLabel -> "Monte Carlo, T=" <> ToString[t4], AxesLabel -> {"n", "Return "},
  PlotStyle -> {AbsoluteThickness[1], AbsoluteThickness[3]}, TextStyle -> {FontSize -> fs},
  PlotLegend -> {"1)", "2) "}, LegendPosition -> {0.5, 0}, LegendOrientation -> Vertical,
  LegendSize -> {0.4, 0.3}, LegendShadow -> {.005, -.005}, LegendTextSpace -> 1, DisplayFunction -> Identity];
Show[GraphicsArray[{g4}]]

Show[GraphicsArray[Table[{g1, g2}, {g3, g4}], GraphicsSpacing -> .01]]

```

Conjectura 6.8.1

```
(* MthesisExtMCIsDRConvexInUnipEquivMeasure.nb - Fernando Hideki Kato - Master Degree Thesis example *)

Off[General::"spell1"]

DDMScalarMul[a_, ddm_] := (ret = ddm; For[i = 0, ++i ≤ Length[ddm], ret[[i, 1]] *= a]; Return[ret])

DDMSort[ddm_] := Sort[ddm, #1[[2]] < #2[[2]] &];

DDMSimplify[ddm_] :=
  (ret = {}; For[i = 0, ++i ≤ Length[ddm], If[ddm[[i, 1]] == 0, Continue[]; j = Length[ret];
    If[(j > 0 && ret[[j, 2]] == ddm[[i, 2]]), ret[[j, 1]] += ddm[[i, 1]], AppendTo[ret, ddm[[i]]]]; Return[ret])

DDMAdd[ddm1_, ddm2_] := DDMSimplify[DDMSort[Join[ddm1, ddm2]]]

DDMMul[ddm1_, ddm2_] :=
  (ret = {}; For[i = 0, ++i ≤ Length[ddm1], For[j = 0, ++j ≤ Length[ddm2], AppendTo[ret, ddm1[[i]] + ddm2[[j]]]];
    Return[DDMSimplify[DDMSort[ret]]])

DDMPow[ddm_, t_] := (ret = ddm; For[i = 0, ++i ≤ Length[ddm], ret[[i, 2]] = ret[[i, 2]]t; Return[ret])

DDMRndUnif[n_, l_, r_] := (ret = {}; For[i = 0, ++i ≤ n, AppendTo[ret, {1/n, Random[Real, {l, r}]}]];
  DDMSimplify[DDMSort[ret]]; Return[ret])

DDMRndUnif2[n_, l_, r_] :=
  (ret = {}; For[i = 0, ++i ≤ n, AppendTo[ret, {Random[Real, {l, r}], Random[Real, {l, r}]}]]; Return[ret])

DDMRndUnif[T_, n_, l_, r_] := (ret = {{1, 1}}; For[t = 1, ++t ≤ T, ret = DDMMul[ret, DDMRndUnif[n, l, r]]];
  Return[ret])

DDMLinComb[w1_, w2_, XY_] :=
  DDMSimplify[DDMSort[Table[{1/Length[XY], w1 XY[[i, 1]] + w2 XY[[i, 2]]}, {i, Length[XY]}]]]

DDMDR[alfa_, RB_, ddm_] :=
  (ret = 0; For[i = 0, ++i ≤ Length[ddm], ret += If[ddm[[i, 2]] ≥ RB, 0, (RB - ddm[[i, 2]])alfa * ddm[[i, 1]]];
  Return[ret(1/alfa)])

alfa = 2; RB = 1.1; T = 3; Tau = 0.1; nRnd = 5; l = 0.5; r = 1.5; u = 4/10; v = 1 - u;

TestConvexity[seed_] := (SeedRandom[seed]; ZT := Evaluate[DDMRndUnif[T, nRnd, l, r]];
  XY := Evaluate[DDMRndUnif2[nRnd, l, r]]; X0 := Table[{1/nRnd, XY[[i, 1]]}, {i, nRnd}];
  Y0 := Table[{1/nRnd, XY[[i, 2]]}, {i, nRnd}]; Z0 := DDMLinComb[u, v, XY];
  X := DDMPow[DDMMul[X0, ZT], 1/(T + Tau)]; Y := DDMPow[DDMMul[Y0, ZT], 1/(T + Tau)];
  Z := DDMPow[DDMMul[Z0, ZT], 1/(T + Tau)]; DRZ := DDMDR[alfa, RB, Z];
  DRuXPvY := u DDMDR[alfa, RB, X] + v DDMDR[alfa, RB, Y]; Return[{seed, DRZ ≤ DRuXPvY, DRZ, DRuXPvY}])

seed := 13000; count := 6; result := Evaluate[Table[TestConvexity[i], {i, seed, seed + count - 1}]];
For[i = 0, ++i ≤ Length[result], If[! result[[i, 2]], Print["Error on seed ", result[[i, 1]]]]]

result

{{13000, True, 0.171851, 0.18252}, {13001, True, 0.204005, 0.212798}, {13002, True, 0.160497, 0.180954},
{13003, True, 0.221907, 0.233736}, {13004, True, 0.128913, 0.137992}, {13005, True, 0.18498, 0.195847}}
```


REFERÊNCIAS

- Acerbi, Carlo. 2001. Spectral Measures of Risk: a Coherent Representation of Subjective Risk Aversion.
- Acerbi, Carlo; Tasche, Dirk. 2002a. On the Coherence of Expected Shortfall. *Journal of Banking and Finance*, V.26 n.7 p.1487-1503 Jul.
- Acerbi, Carlo; Tasche, Dirk. 2002b. Expected Shortfall: a Natural Coherent Alternative to Value at Risk. *Economic Notes*, V.31 n.2 p.379-388 Jul.
- Acerbi, Carlo; Nardio, Claudio; Sirtori, Carlo. 2001. Expected Shortfall as a Tool for Financial Risk Management.
- Adamchik, Victor. 1996. The Evaluation of Integrals of Bessel Functions Via G-Function Identities.
- Aitchison, J.; Brown, J.A.C. 1963. *The Lognormal Distribution*. Cambridge University Press. London. 1st published in 1957.
- Albeverio, S.; Blanchar, Ph.; Streit, L. (editors). 1987. *Stochastic Process – Mathematics and Physics II. Lecture Notes in Mathematics*, V.1250. Springer-Verlag. Berlin.
- Albrecht, Peter. 2004. Risk Measures. *Encyclopedia of Actuarial Science*. John Wiley & Sons. New York, NY.
- Alderfer, Clayton P.; Bierman Jr., Harold. 1970. Choices with Risk: Beyond the Mean and Variance. *The Journal of Business*, V.43 N.3 p.341-353 Jul.
- Alexander, Gordon J. 1976. The Derivation of Efficient Sets. *Journal of Financial and Quantitative Analysis*, V.11 N.5 p.817-830 Dec.
- Algoet, Paul H.; Cover, Thomas M. 1988. Asymptotic Optimality and Asymptotic Equipartition Properties of Log-Optimum Investment. *The Annals of Probability*, V.16 N.2 p.876-898.
- Anderson, W.J.; Mathai, A.M. 1984. Various Representations of a Generalized Hypergeometric Function through Statistical Techniques. *Math. Proc. Camb. Phil. Soc.*, N.96 p.325-330.
- Ang, James. 1975. A Note on the E, SL Portfolio Selection Model. *Journal of Financial and Quantitative Analysis*, V.10 N.5 p.849-857 Dec.
- Arditti, Fred D.; Levy, Haim. 1975. Portfolio Efficiency Analysis in Three Moments: The Multiperiod Case. *The Journal of Finance*, V.30 N.3 p.797-809 Jun. Reprinted in Levy; Sarnat, editors (1977).
- Artzner, Philippe. 1999. Application of Coherent Risk Measures to Capital Requirements in Insurance.
- Artzner, Philippe; Delbaen, Freddy; Eber, Jean-Marc; Heath, David. 1999. Coherent Measures of Risk. *Mathematical Finance*, V.9 N.3.
- Arzac, Enrique R.; Bawa, Vijay S. 1977. Portfolio Choice and Equilibrium in Capital Markets with Safety-First Investors. *Journal of Financial Economics*, V.4 N.3 p.277-288.
- Athayde, Gustavo M. de. 2001. Building a Mean-Downside Risk Portfolio Frontier. in Sortino; Satchell, editors (2001).
- Aucamp, Donald. 1978. A Comment on Geometric Mean Portfolios. *Management Science*, V.24 N.8 p.859 Apr.
- Bajeux-Besnainou, Isabelle; Jordan, James V. 2001. Mean-Variance Asset Allocation for Long Horizons. *Finance*, V.22 p.7-24 Dec.
- Bajeux-Besnainou, Isabelle; Jordan, James V.; Portait, Roland. 2001. An Asset Allocation Puzzle: Comment. *The American Economic Review*, V.91 p.1170-1179 Sep.
- Bajeux-Besnainou, Isabelle; Jordan, James V.; Portait, Roland. 2003. Dynamic Asset Allocation for Stocks, Bonds, and Cash. *The Journal of Business*, V.76 N.2 p.263-287 Apr.
- Barberis, Nicholas. 2000. Investing for the Long Run When Returns are Predictable. *The Journal of Finance*, V.55 N.1 p.225-264 Feb.
- Baumol, W.J. 1963. An Expected Gain-Confidence Limit Criterion for Portfolio Selection. *Management Science*, V.10 N.1 Oct.
- Bawa, Vijay S. 1975. Optimal Rules for Ordering Uncertain Prospects. *Journal of Financial Economics*, V.2 N.1 p.95-121.
- Bawa, Vijay S. 1976. Admissible Portfolios for All Individuals. *The Journal of Finance*, V.31 N.4 p.1169-1183 Sep.
- Bawa, Vijay S.; Lindenberg, Eric B. 1977. Capital Market Equilibrium in a Mean-Lower Partial Moment Framework. *Journal of Financial Economics*, V.5 p.189-200.
- Bawa, Vijay S. 1977. Mean-Lower Partial Moments and Asset Prices. *Journal of Financial Economics*, Jun.
- Bawa, Vijay S. 1978. Safety First, Stochastic Dominance and Optimal Portfolio Choice. *Journal of Financial and Quantitative Analysis*, V.13 p.255-271 Jun.

- Bawa, Vijay S.; Chakrin, Lewis M. 1979. Optimal Portfolio Choice and Equilibrium in a Lognormal Securities Market. *in* Elton; Gruber, editors (1979, p.47-62).
- Bawa, Vijay S.; Elton, Edwin J.; Gruber, Martin J. 1979. Simple Rules for Optimal Portfolio Selection in a Stable Paretian Markets. *The Journal of Finance*, V.34 N.4 p.1041-1047 Sep.
- Baz, Jamil; Briys, Eric; Bronnenberg, Bart J.; Cohen, Michèle; Kast, Robert; Viala, Pascale; Wathieu, Luc; Weber, Martin; Wertenbroch, Klaus. 1999. Risk Perception in the Short Run and in the Long Run. *Marketing Letters*, V.10 N.3 p.267-283 Aug.
- Bell, Robert M.; Cover, Thomas M. 1980. Competitive Optimality of Logarithmic Investment. *Mathematics of Operations Research*, V.5 N.2 p.161-166 May.
- Bell, Robert; Cover, Thomas M. 1988. Game-Theoretic Optimal Portfolios. *Management Science*, V.34 N.6 p.724-733 Jun.
- Bellman, Richard E. 1957. *Dynamic Programming*. Princeton University Press. Princeton, NJ.
- Berkelaar, Arjan; Kouwenberg, Roy. 2000. *Dynamic Asset Allocation and Downside-Risk Aversion*.
- Bernstein, Peter L. 1992. *Capital Ideas – The Improbable Origins of Modern Wall Street*. The Free Press. New York, NY.
- Bernstein, Peter L. 1996. *Against the Gods – The Remarkable Story of Risk*. John Wiley & Sons. New York, NY.
- Best, Michael J.; Grauer, Robert R. 1990. The Efficient Set Mathematics When Mean-Variance Problems are Subject to General Linear Constraints. *Journal of Economics and Business*, V.42 N.2 p.105-120 May.
- Bicksler, J.; Friend, I. (editors). 1977. *Studies in Risk and Return*. Ballinger. Cambridge, MA.
- Bielecki, Tomasz R.; Jin, Hanqing; Pliska, Stanley R.; Zhou, Xun Yu. 2003. Continuous-Time Mean-Variance Portfolio Selection with Bankruptcy Prohibition.
- Bierman, Jr., Harold. 1997. Portfolio Allocation and the Investment Horizon. *The Journal of Portfolio Management*, V.23 N.4 p.51-55 Summer.
- Bierman, Jr., Harold. 1998. A Utility Approach to the Portfolio Allocation Decision and the Investment Horizon. *The Journal of Portfolio Management*, V.25 N.1 p.81-87 Fall.
- Bodie, Zvi. 1995. On the Risk of Stocks in the Long Run. *Financial Analysts Journal*, V.51 N.3 p.18-22 May-Jun.
- Bondesson, Lennart. 1992. *Generalized Gamma Convolutions and Related Classes of Distributions and Densities*. Lecture Notes in Statistics, V.76. Springer-Verlag. Berlin.
- Booth, David; Fama, Eugene. 1992. Diversification Returns and Asset Contributions. *Financial Analysts Journal*, p.26-32 May-Jun.
- Borch, Karl. 1969. A Note on Uncertainty and Indifference Curves. *The Review of Economic Studies*, V.36 p.1-4.
- Boyd, Stephen; Vandenberghe, Lieven. 2004. *Convex Optimization*. Cambridge University Press. Cambridge, UK.
- Braaksma, B.L.G. 1964. Asymptotic Expansions and Analytic Continuation for a Class of Barnes-Integrals. *Compositio Mathematica*, V.15 p.239-341.
- Brealey, Richard A. 1991. Harry M. Markowitz's Contributions to Financial Economics. *Scandinavian Journal of Economics*, V.93 N.1 p.7-17.
- Breiman, L. 1960. Investment Policies for Expanding Businesses Optimal in a Long-Run Sense. *Naval Research Logistics Quarterly*, V.7 p.647-651. Reprinted with *addendum in* Ziemba; Vickson, editors (1975).
- Breiman, L. 1961. Optimal Gambling Systems for Favorable Games. *in* Noyman, Jorzy (editor). *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*. Four Symposia. Regents of the University of California, 1, p.65-78.
- Brennan, Michael J.; Schwartz, Eduardo S. 1985. On the Geometric Mean Index: A Note. *Journal of Financial and Quantitative Analysis*, V.20 N.1 p.119-122 Mar.
- Brogan, Anita Jean. 2003. *Mutual Fund and Security Selection through an Operations Research Lens*. PhD Thesis, University of North Carolina at Chapel Hill.
- Brogan, Anita Jean; Stidham, Jr, Shaler. c2003a. A Note on Separation in Mean-LPM Portfolio Optimization with Fixed and Moving Targets.
- Brogan, Anita Jean; Stidham, Jr, Shaler. c2003b. Non-Separation in the Mean-LPM Portfolio Optimization Problem.
- Brown, David P. 1987. Multiperiod Financial Planning. *Management Science*, V.33 N.7 p.848-875 Jul.
- Bushman, Robert G. 1974. The Asymptotic Expansion of an Integral. *Rendiconti di Matematica*, V.7 N.3 p.481-486.
- Canner, Niko; Mankiw, N.Gregory; Weil, David N. 1997. An Asset Allocation Puzzle. *The American Economic Review*, V.87 N.1 p.181-191 Mar.
- Carter, Bradley D.; Springer, Melvin D. 1977. The Distribution of Products, Quotients and Powers of Independent H-Function Variates. *SIAM Journal of Applied Mathematics*, V.33 N.4 p.542-558 Dec.
- Chamberlain, Gary. 1983. A Characterization of the Distributions That Imply Mean-Variance Utility Functions. *Journal of Economic Theory*, V.29 p.185-201.
- Chen, Andrew H.Y.; Jen, Frank C.; Zions, Stanley. 1971. The Optimal Portfolio Revision Policy. *The Journal of Business*, V.44 N.1 p.51-61 Jan. Reprinted *in* Elton; Gruber, editors (1972).

- Cheng, Pao L.; Deets, M. King. 1971. Test of Portfolio Building Rules: Comment. *The Journal of Finance*, V.26 N.4 p.965-972 Sep.
- Cheng, Ping. 2001. Comparing Downside-Risk and Mean-Variance Analysis Using Bootstrap Simulation. *Journal of Real Estate Portfolio Management*, V.7 N.3 p.225-238.
- Cheng, Ping; Wolverton, Marvin L. 2001. MPT and the Downside Risk Framework: A Comment on Two Recent Studies. *Journal of Real Estate Portfolio Management*, V.7 N.2 p.125-131.
- Chin, Elion; Weigend, Andreas S.; Zimmermann, Heinz. 1999. Computing Portfolio Risk Using Gaussian Mixtures and Independent Component Analysis. *in Proceedings of the 1999 IEEE/IAFE/INFORMS Conference on Computational Intelligence for Financial Engineering*, p.74-117. New York, NY.
- Clarkson, Robert S. c2000. A general theory of Financial Risks.
- Conine Jr., Thomas E.; Tamarkin, Maurry J. 1981. On Diversification Given Asymmetry in Returns. *The Journal of Finance*, V.36 N.5 p.1143-1155 Dec.
- Constantinides, George M. 1976. Optimal Portfolio Revision with Proportional Transaction Costs: Extension to HARA Utility Functions and Exogeneous Deterministic Income. *Management Science*, V.22 N.8 p.921-923 Apr.
- Constantinides, George M. 1979. Multiperiod Consumption and Investment Behavior with Convex Transaction Costs. *Management Science*, V.25 N.11 p.1127-1137.
- Constantinides, George M. 1990. Habit Formation: A Resolution of the Equity Premium Puzzle. *Journal of Political Economy*, V.98 N.3 p.519-543 Jun.
- Constantinides, G.M.; Malliaris, A.G. 1995. Portfolio Theory. *in Jarrow; Maksimovic; Ziemba, editors. (1995, p.1-30).*
- Constantinides, George M.; Donaldson, John B.; Mehra, Rajnish. 2002. Junior Can't Borrow: A New Perspective on the Equity Premium Puzzle. *The Quarterly Journal of Economics*, V.117 p.269-296 Feb.
- Constantinides, George M. 2002. Rational Asset Prices. *The Journal of Finance*, V.57 N.4 p.1567-1591 Aug.
- Cover, Thomas M. 1991. Universal Portfolios. *Mathematical Finance*, V.1 N.1 p.1-29 Jan.
- Cox, J.; Huang, C.F. 1989. Optimal Consumption and Portfolio Policies when Asset Prices Follow a Diffusion Process. *Journal of Economic Theory*, V.49 N.1 p.33-83 Oct.
- Davis, M.H.A.; Norman, A.R. 1990. Portfolio Selection with Transaction Costs. *Mathematics of Operations Research*, V.15 N.4 p.676-713.
- Delbaen, Freddy. 2000. Coherent Risk Measures on General Probability Spaces.
- Dempsey, Mike; Hudson, Robert; Littler, Kevin; Keasey, Kevin. 1996. On the Risk of Stocks in the Long Run: A Resolution to the Debate?. *Financial Analysts Journal*, V.52 N.5 p.57-62 Sep-Oct.
- Denault, Michel. 1999. Coherent allocation of Risk Capital.
- Deng, Xiao-Tie; Wang, Shou-Yang; Xia, Yu-Sen. 2000. Criteria, Models and Strategies in Portfolio Selection. *Advanced Modeling and Optimization*, V.2 N.2 p.79-103.
- Dexter, A.S.; Yu, J.N.W.; Ziemba, W.T. 1980 Portfolio Selection in a Lognormal Market When the Investor Has a Power Utility Function: Computational Results. *in Dempster, M.A.H. (editor). 1980. Stochastic Programming. p.507-523. Academic Press. New York, NY.*
- Ditkin, V.A.; Prudnikov, A.P. 1965. *Integral Transforms and Operational Calculus*. Pergamon Press. Oxford.
- Domar, E.; Musgrave, R.A. 1944. Proportional Income Taxation and Risk Taking. *Quarterly Journal of Economics*, V.57 May.
- Duffie, D. 1992. *Dynamic Asset Pricing Models*. Princeton University Press. Princeton, NJ.
- Dumas, Bernhard; Luciano, Elisa. 1991. An Exact Solution to a Dynamic Portfolio Choice Problem under Transactions Costs. *The Journal of Finance*, V.46 N.2 p.577-595 Jun.
- Dybvig, Philip H.; Roger, L.C.G.; Back, Kerry. 1999. Portfolio Turnpikes. *The Review of Financial Studies*, V.12 N.1 p.165-195 Spring.
- Eaton, R.Douglas Van; Conover, James A. 1998. Misconceptions about Optimal Equity Allocation and Investment Horizon. *Financial Analysts Journal*, V.54 N.2 p.52-59 Mar-Apr.
- Eaton, R.Douglas Van; Conover, James A. 2004. Misconceptions about Optimal Equity Allocation and Investment Horizon: Author's Response. *Financial Analysts Journal*, V.60 N.1 p.14 Jan-Feb.
- Ekern, Steinar. 1981. Time Dominance Efficiency Analysis. *The Journal of Finance*, V.36 N.5 p.1023-1034 Dec.
- Elton, Edwin J.; Gruber, Martin J. (editors). 1972. *Security Evaluation and Portfolio Analysis*. Prentice-Hall. Englewood Cliffs, NJ.
- Elton, Edwin J.; Gruber, Martin J. 1974a. On the Optimality of Some Multiperiod Portfolio Selection Criteria. *The Journal of Business*, V.47 N.2 Apr.
- Elton, Edwin J.; Gruber, Martin J. 1974b. The Multi-Period Consumption Investment Problem and Single Period Analysis. *Oxford Economics Papers*, V.26 N.2 p.289-301 Jul.
- Elton, Edwin J.; Gruber, Martin J. 1974c. On the Maximization of the Geometric Mean with Lognormal Return Distribution. *Management Science*, V.21 N.4 p.483-488 Dec.

- Elton, Edwin J.; Gruber, Martin J.; Padberg, Manfred W. 1976. Simple Criteria for Optimal Portfolio Selection. *The Journal of Finance*, V.31 N.5 p.1341-1384 Dec.
- Elton, Edwin J.; Gruber, Martin J. (editors). 1979. *Portfolio Theory, 25 Years After*. North-Holland. Amsterdam.
- Elton, Edwin J.; Gruber, Martin J. 1995. *Modern Portfolio Theory and Investment Analysis*. 5th ed. John Wiley & Sons. New York, NY. 1st ed. in 1981.
- Emmer, Susanne; Klüppelberg, Claudia; Korn, Ralf. 2000. Optimal Portfolios with Bounded Downside Risks.
- Erdélyi, A.; Magnus, W.; Oberhettinger, F.; Tricomi, F.G. 1953. *Higher Transcendental Functions*, V.1. McGraw-Hill. New York, NY.
- Erdélyi, A.; Magnus, W.; Oberhettinger, F.; Tricomi, F.G. 1954. *Tables of Integral Transforms*, V.1. McGraw-Hill. New York, NY.
- Fair, Ray C. 1986. Evaluating the Predictive Accuracy of Models. *in* Griliches; Intriligator, editors (1986).
- Fama, Eugene F. 1970. Multiperiod Consumption-Investment Decisions. *The American Economic Review*, V.60 N.1 p.163-174. Reprinted *in* Ziemba; Vickson, editors (1975).
- Fama, Eugene F.; Miller, Merton H. 1972. *The Theory of Finance*. Dryden Press. Hinsdale, IL.
- Fama, Eugene F.; Macbeth, James D. 1974. Long-Term Growth in a Short Term Market. *The Journal of Finance*, V.29 N.3 p.857-885 Jun.
- Fama, Eugene F. 1976. Multiperiod Consumption-Investment Decisions: A Correction. *The American Economic Review*, V.66 N.4 p.723-724 Sep.
- Feller, William. 1971. *An Introduction to Probability Theory and Its Applications*. 2nd ed. John Wiley & Sons. New York, NY.
- Ferguson, R.; Simaan, Y. 1996. Portfolio Composition and the Investment Horizon Revisited. *The Journal of Portfolio Management*, V.22 N.4 p.62-67 Summer.
- Fishburn, Peter C. 1977. Mean-Risk Analysis with Risk Associated with Below Target Returns. *American Economic Review*, V.67 N.2 p.116-126 Mar.
- Fishburn, Peter C. 1984. Foundations of Risk Measurement. I. Risk as Probable Loss. *Management Science*, V.30 N.4 p.398-406 Apr.
- Fisher, I. 1906. *The Nature of Capital and Income*. Macmillan. New York, NY.
- Floudas, Christodoulos A. 1999. *Deterministic Global Optimization – Theory, Methods and Applications*. Kluwer Academic. Dordrecht.
- Föllmer, Hans; Schied, Alexander. 2002. Robust Preferences and Convex Measures of Risk.
- Fox, C. 1961. The G and H Functions as Symmetrical Fourier Kernels. *Trans. Amer. Math. Soc.* V.98 p.395-429.
- Frankfurter, George M. 1990. Is Normative Portfolio Dead? *Journal of Economics and Business*, V.42 N.2 p.95-98 May.
- Frankfurter, George M.; Phillips, Herbert E.; Faulk, Greg. 1999. The Ex Post Performance of Four Portfolio Selection Algorithms. *Review of Quantitative Finance and Accounting*, V.13 N.4 p.347-366 Dec.
- Friedman, Milton. 1953. *Essays in Positive Economics*. The University of Chicago Press. Chicago, IL.
- Gollier, Christian. 2001. *The Economics of Risk and Time*. MIT Press. Cambridge.
- Gollier, Christian; Zeckhauser, Richard J. 2002. Horizon Length and Portfolio Risk. *The Journal of Risk and Uncertainty*, V.24 N.3 p.195-212 May.
- Gorenflo, Rudolf; Mainardi, Francesco. 2000. *Essentials of Fractional Calculus*.
- Granito, Michael; Walsh, Patrick. 1978. Portfolio Efficiency Analysis in Three Moments – The Multiperiod Case: Comment. *The Journal of Finance*, V.33 N.1 p.345-348 Mar.
- Grauer, Robert R. 1981. A Comparison of Growth Optimal and Mean Variance Investment Policies. *The Journal of Financial and Quantitative Analysis*, V.16 N.1 p.1-21 Mar.
- Grauer, Robert R.; Hakansson, Nils H. 1982. Higher Return, Lower Risk: Historical Returns on Long-Run, Actively Managed Portfolios of Stocks, Bonds and Bills, 1936-1978. *Financial Analysts Journal*, V.38 N.2 p.39-53.
- Grauer, Robert R.; Hakansson, Nils H. 1986. A Half Century of Returns on Levered and Unlevered Portfolios of Stocks, Bonds, and Bills, with and without Small Stocks. *The Journal of Business*, V.59 N.2 p.287-318.
- Grauer, Robert R.; Hakansson, Nils H. 1993. On the Use of Mean-Variance and Quadratic Approximations in Implementing Dynamic Investment Strategies: A Comparison of Returns and Investment Policies. *Management Science*, V.39 N.7 p.856-871 Jul.
- Gressis, J.; Hayya, J.; Philippatos, G.C. 1974. Multiperiod Portfolio Efficiency Analysis Via the Geometric Mean. *The Financial Review*, V.9 p.46-63.
- Gressis, N.; Philippatos, G.C.; Hayya, J. 1976. Multiperiod Portfolio Analysis and the Inefficiency of the Market Portfolio. *The Journal of Finance*, V.31 N.4 p.1115-1126 Sep.
- Griliches, Z.; Intriligator, M.D. (editors). 1983. *Handbook of Econometrics*, V.1. North-Holland. Amsterdam.
- Griliches, Z.; Intriligator, M.D. (editors). 1986. *Handbook of Econometrics*, V.3. Elsevier. Amsterdam.

- Grootveld, Henk; Hallerbach, Winfried. 1999. Variance vs Downside Risk: Is There Really that Much Difference? *European Journal of Operational Research*, V.114 p.304-319.
- Gujarati, Damodar N. 1995. *Basic Econometrics*. 3rd ed. McGraw-Hill. New York, NY. 1st ed. in 1978.
- Gunthorpe, Deborah; Levy, Haim. 1994. Portfolio Composition and the Investment Horizon. *Financial Analysts Journal*, V.50 N.1 p.51-56 Jan-Feb.
- Gupta, Aparna. 2000. Optimum Asset Allocation with Behavioral Utilities: A Plan for Acquiring and Consuming Retirement Funds. PhD Thesis, Stanford University.
- Haas, Marcus. 2001. New Methods in Backtesting.
- Hakansson, Nils H.; Liu, Tien-Ching. 1970. Optimal Growth Portfolios When Yields Are Serially Correlated. *The Review of Economic and Statistics*, V.52 N.4 p.385-394 Nov.
- Hakansson, Nils H. 1970. Optimal Investment and Consumption Strategies under Risk for a Class of Utility Functions. *Econometrica*, V.38 N.5 p.587-607 Sep. Reprinted in Ziemba; Vickson, editors (1975).
- Hakansson, Nils H. 1971a. Multi-Period Mean-Variance Analysis: Toward a General Theory of Portfolio Choice. *The Journal of Finance*, V.26 N.4 p.857-884 Sep.
- Hakansson, Nils H. 1971b. On Optimal Myopic Portfolio Policies, With and Without Serial Correlation of Yields. *The Journal of Business*, V.44 N.3 p.324-334 Jul. Reprinted in Ziemba; Vickson, editors (1975).
- Hakansson, Nils H. 1971c. Capital Growth and the Mean-Variance Approach to Portfolio Selection. *The Journal of Financial and Quantitative Analysis*, V.6 N.1 p.517-557 Jan.
- Hakansson, Nils H. 1974a. Comment on Merton and Samuelson. *Journal of Financial Economics*, V.1 N.1 p.95 May.
- Hakansson, Nils H. 1974b. Convergence to Isoelastic Utility and Policy in Multiperiod Portfolio Choice. *Journal of Financial Economics*, V.1 N.3 p.201-224 Sep.
- Hakansson, Nils H.; Miller, Bruce L. 1975. Compound-Return Mean-Variance Portfolios Never Risk Ruin. *Management Science*, V.22 N.4 p.391-400 Dec.
- Hakansson, Nils H. 1979. A Characterization of Optimal Multi-Period Portfolio Policies. in Elton; Gruber, editors (1979, p.169-177).
- Hakansson, Nils H.; Ziemba, W.T. 1995. Capital Growth Theory. in Jarrow; Maksimovic; Ziemba, editors. (1995, p.65-86).
- Hansson, Björn; Persson, Mattias. 2000. Time Diversification and Estimation Risk. *Financial Analysts Journal*, V.56 N.5 p.55-62 Sep-Oct.
- Hardy, G; Littlewood, J.E.; Pólya, G. 1952. *Inequalities*. 2nd ed. Cambridge University Press. Cambridge. 1st ed. in 1934.
- Harlow, W.V.; Rao, Ramesh K.S. 1989. Asset Pricing in a Generalized Mean-Lower Partial Moment Framework: Theory and Evidence. *Journal of Financial and Quantitative Analysis*, V.24 N.3 p.285-311.
- Harlow, W.V. 1991. Asset Allocation in a Downside-Risk Framework. *Financial Analysts Journal*, V.47 N.5 p.28-40 Sep-Oct.
- Haubold, H.J.; Mathai, A.M. 2000. The Fractional Kinetic Equation and Thermonuclear Functions. *Astrophysics and Space Science* V.273 N.1-4 p.53-63.
- Hegyi, S. 2001. A Powerful Generalization of the NBD Suggested by Peter Carruthers.
- Hegyi, S. 2002. H-Function Extension of the NBD in the Light of Experimental Data.
- Helmbold, David P.; Schapire, Robert E.; Singer, Yoram; Warmuth, Manfred K. 1995. A Comparison of New and Old Algorithms for a Mixture Estimation Problem. in *Proceedings of the Eighth Annual Workshop on Computational Learning Theory*, p.69-78.
- Helmbold, David P.; Schapire, Robert E.; Singer, Yoram; Warmuth, Manfred K. 1996. On-Line Portfolio Selection Using Multiplicative Updates. in *Machine Learning: Proceedings of the Thirteenth International Conference*.
- Hicks, J.R. 1946. *Value and Capital*. 2nd ed. Oxford University Press. Oxford.
- Hogan, William W.; Warren, James M. 1972. Computation of the Efficient Boundary in the E-S Portfolio Selection Model. *Journal of Financial and Quantitative Analysis*, V.7 N.4 p.1881-1896 Sep.
- Hogan, William W.; Warren, James M. 1974. Toward the Development of an Equilibrium Capital-Market Model Based on Semivariance. *Journal of Financial and Quantitative Analysis*, V.9 N.4 p.1-11 Jan.
- Horvath, Philip A. 1995. Compounding/Discounting in Continuous Time. *Quarterly Review of Economics and Finance*, V.35 N.3 p.315-325 Fall.
- Huang, C.C.; Vertinsky, I.; Ziemba, W.T. 1978. On Multiperiod Stochastic Dominance. *Journal of Financial and Quantitative Analysis*, V.13 N.1 p.1-13 Mar.
- Huang, Chi-fu; Litzenberger, Robert H. 1988. *Foundations for Financial Economics*. North-Holland. New York, NY.
- Huberman, Gur; Ross, Stephen. 1983. Portfolio Turnpike Theorem, Risk Aversion, and Regularly Varying Utility Functions. *Econometrica*, V.51 N.5 p.1345-1361 Sep.
- Hull, John C. 1998. *Introduction to Futures and Options Markets*. 3rd ed. Prentice-Hall. Upper Saddle River, NJ. 1st ed. in 1991.

- Hürlimann, Werner. 2001. Analytical Evaluation of Economic Risk Capital for Portfolios of Gamma Risks. *Astin Bulletin*, V.31 N.1 p.107-122.
- Ingersoll, Jr., Jonathan E. 1987. *Theory of Financial Decision*. Rowman & Littlefield. Savage, MD.
- Jacquier, Eric; Kane, Alex; Marcus, Alan J. 2003. Geometric or Arithmetic Mean: A Reconsideration. *Financial Analysts Journal*, V.59 N.6 p.46-53 Nov-Dec.
- Jagannathan, Ravi; Kocherlakota, Narayana R. 1996. Why Should Older People Invest Less in Stocks Than Younger People? *Federal Reserve Bank of Minneapolis Quarterly Review*, V.20 N.3 p.11-23 Summer.
- Jarrow, R.A. 1985. *Finance Theory*. Prentice-Hall. Englewood Cliffs, NJ.
- Jarrow, R.A.; Maksimovic, V.; Ziemba, W.T. (editors). 1995. *Finance. Handbooks in Operations Research and Management Science*, V.9. Elsevier. Amsterdam.
- Jarrow, Robert; Zhao, Feng. 2003. Downside Loss Aversion, Portfolio Management, and the Equity Premium Puzzle.
- Jean, William H. 1980. The Geometric Mean and Stochastic Dominance. *The Journal of Finance*, V.35 N.1 p.151-158 Mar.
- Jean, William H.; Helms, Billy P. 1983. Geometric Mean Approximations. *Journal of Financial and Quantitative Analysis*, V.18 N.3 p.287-293 Sep.
- Johnson, Norman L.; Kotz, Samuel; Balakrishnan, N. 1994. *Continuous Univariate Distributions*. V.1. 2nd ed. John Wiley & Sons. New York, NY.
- Jones, Irwin E. 1971. Test of Portfolio Building Rules: Comment. *The Journal of Finance*, V.26 N.4 p.973-975 Sep.
- Jørgensen, Bent. 1982. Statistical Properties of the Generalized Inverse Gaussian Distribution. *Lecture Notes in Statistics*, V.9. Springer-Verlag. Berlin.
- Jorion, Philippe. 1997. *Value at Risk: The New Benchmark for Controlling Market Risk*. McGraw-Hill. New York, NY.
- Kallberg, J.G.; Ziemba, W.T. 1983. Comparison of Alternative Utility Functions in Portfolio Selection Problems. *Management Science*, V.29 N.11 p.1257-1276 Nov.
- Kamin, Jules H. 1975. Optimal Portfolio Revision with a Proportional Transaction Cost. *Management Science*, V.21 N.11 p.1263-1271 Jul.
- Kang, Taehoon; Brorsen, B.Wade; Adam, Brian D. 1996. A New Efficiency Criterion: The Mean-Separated Target Deviations Model. *Journal of Economics and Business*, V.48 N.1 p.47-66.
- Kapur, J.N. 1989. *Maximum-Entropy Models in Science and Engineering*. John Wiley & Sons. New York, NY.
- Kapur, J.N. 1994. *Measures of Information and Their Applications*. John Wiley & Sons. New York, NY.
- Kapur, J.N.; Kesavan, H.K. 1992. *Entropy Optimization Principles with Applications*. Academic Press. San Diego, CA.
- Karatzas, I.; Lehoczky, J.; Shreve, S. 1987. Optimal Portfolio and Consumption Decisions for a 'Small Investor' on a Finite Horizon. *SIAM Journal of Control and Optimization*, p.1157-1186.
- Kelly Jr., J.L. 1956. A New Interpretation of Information Rate. *Bell System Technical Journal*, V.35 p.917-926.
- Keynes, J.M. 1937. *The General Theory of Employment Interest and Money*. Macmillan. London.
- Kijima, M.; Ohnishi, M. 1993. Mean-Risk Analysis of Risk Aversion and Wealth Effects on Optimal Portfolios with Multiple Investment Opportunities. *Annals of Operations Research*, V.45 p.147-163.
- Kilbas, Anatoly A.; Saigo, Megumi. 1998. On the H-Function.
- Kira, D.; Ziemba, W.T. 1977. Equivalence among Alternative Portfolio Selection Criteria. *in* Levy; Sarnat, editors (1977).
- Klein, Roger W.; Bawa, Vijay S. 1976. The Effect of Estimation Risk on Optimal Portfolio Choice. *Journal of Financial Economics*, V.3 p.215-231.
- Klein, Roger W.; Bawa, Vijay S. 1977. The Effect of Limited Information and Estimation Risk on Optimal Portfolio Diversification. *Journal of Financial Economics*, V.5 p.89-111.
- Korn, Ralf. 1997. *Optimal Portfolios*. World Scientific. River Edge, NJ.
- Kraus, Alan; Litzenberger, Robert H. 1976. Skewness Preference and the Valuation of Risk Assets. *The Journal of Finance*, V.31 N.4 p.1085-1100 Sep.
- Kritzman, Mark P. 1991. The Nobel Prize. *Financial Analysts Journal*, V.47 N.1 p.10-15 Jan-Feb.
- Kritzman, Mark. 1994. What Practitioners Need to Know... About Time Diversification. *Financial Analysts Journal*, V.50 N.1 p.14-18 Jan-Feb.
- Kritzman, Mark; Rich, Don. 2002. The Mismeasurement of Risk. *Financial Analysts Journal*, V.58 N.3 p.91-99 May-Jun.
- Kroll, Yoram; Levy, Haim; Markowitz, Harry M. 1984. Mean-Variance Versus Direct Utility Maximization. *The Journal of Finance*, V.39 N.1 p.47-61 Mar.
- Kroll, Yoram; Levy, Haim. 1982. Stochastic Dominance: A Note. *The Journal of Finance*, V.37 N.3 p.871-875 Jun.
- Latané, Henry A. 1959. Criteria for Choice among Risky Ventures. *Journal of Political Economy*, V.67 N.2 p.144-155 Apr.
- Latané, Henry A.; Tuttle, Donald L. 1967. Criteria for Portfolio Building. *The Journal of Finance*, V.22 N.3 p.359-373 Sep.

- Latané, Henry A; Young, William E. 1969. Test of Portfolio Building Rules. *The Journal of Finance*, V.24 N.4 p.595-612 Sep.
- Latané, Henry A. Young, William E. 1971. Test of Portfolio Building Rules: A Reply. *The Journal of Finance*, V.26 N.4 p.976-981 Sep.
- Latané, Henry A. 1979. The Geometric Mean Criterion Continued. *Journal of Banking and Finance*, V.3 p.309-311.
- Leamer, Edward E. 1983. Model Choice and Specification Analysis. *in* Griliches; Intriligator, editors (1983).
- Lee, Wayne Y.; Rao, Ramesh K.S. 1988. Mean Lower Partial Moment Valuation and Lognormally Distributed Returns. *Management Science*, V.34 N.4 p.443-453 Apr.
- Lee, Wayne Y. 1990. Diversification and Time: Do Investment Horizons Matter? *The Journal of Portfolio Management*, V.16 N.3 p.21-26 Spring.
- Levy, Haim. 1972. Portfolio Performance and the Investment Horizon. *Management Science*, V.18 N.12 p.B645-B653 Aug.
- Levy, Haim. 1973. Stochastic Dominance, Efficiency Criteria, and Efficient Portfolios: The Multi-Period Case. *The American Economic Review*, V.63 N.5 p.986-994 Dec.
- Levy, Haim; Sarnat, Marshall. (editors). 1977. *Financial Decision Making under Uncertainty*. Academic Press. New York, NY.
- Levy, Haim. 1979. Does Diversification Always Pay? *in* Elton; Gruber, editors (1979, p.63-71).
- Levy, Haim; Markowitz, Harry M. 1979. Approximating Expected Utility by a Function of Mean and Variance. *The American Economic Review*, V.69 p.308-317.
- Levy, Haim. 1984. Measuring Risk and Performance Over Alternative Investment Horizons. *Financial Analysts Journal*, V.40 N.2 p.61-68 Mar-Apr.
- Levy, Haim; Sarnat, Marshall. 1984. *Portfolio and Investment Selection: Theory and Practice*. Prentice Hall. Englewood Cliffs, NJ.
- Levy, Haim; Samuelson, Paul A. 1992. The Capital Asset Pricing Model with diverse Holding Periods. *Management Science*, V.38 N.11 p.1529-1542 Nov.
- Levy, Haim; Spector, Yishay. 1996. Cross-Asset versus Time Diversification. *The Journal of Portfolio Management*, V.22 N.3 p.24-34 Spring.
- Levy, Haim; Cohen, Allon. 1998. On the Risk of Stocks in the Long Run: Revisited. *The Journal of Portfolio Management*, V.24 N.3 p.60-69 Spring.
- Li, Duan; Ng, Wan-Lung. 1998. Optimal Dynamic Portfolio Selection: Multiperiod Mean-Variance Formulation. *Mathematical Finance*, V.10 N.3 p.387-406 Jul.
- Lima, Elon Lages. 1987. *Curso de Análise*. V.1. 5th ed. Projeto Euclides. Instituto de Matemática Pura e Aplicada. Rio de Janeiro, RJ. CNPq. Brasília, DF.
- Liu, Jun. 2001. Dynamic Portfolio Choice and Risk Aversion.
- Litzenberger, Robert; Budd, A.P. 1971. A Note on Geometric Mean Portfolio Selection and the Market Prices of Equities. *Journal of Financial and Quantitative Analysis*, V.6 N.5 p.1277-1282 Dec.
- Lo, Andrew W. 1999. The Three P's of Total Risk Management. *Financial Analysts Journal*, Jan/Feb.
- Luenberger, David G. 1998. *Investment Science*. Oxford University Press. New York, NY.
- MacLean, L.C.; Ziemba, W.T.; Blazenko, G. 1992. Growth versus Security in Dynamic Investment Analysis. *Management Science*, V.38 N.11 p.1562-1585 Nov.
- MacLean, Leonard C.; Ziemba, William T. 1999. Growth versus Security Tradeoffs in Dynamic Investment Analysis. *Annals of Operations Research*, V.85 p.193-225.
- Magill, Michael J.P.; Constantinides, George M. 1976. Portfolio Selection with Transaction Costs. *Journal of Economic Theory*, V.13 N.2 p.245-263 Oct.
- Maier, Steven F.; Peterson, David W.; Vander Weide, James H. 1977. A Monte Carlo Investigation of Characteristics of Optimal Geometric Mean Portfolios. *Journal of Financial and Quantitative Analysis*, V.12 N.2 p.215-233 Jun.
- Mainardi, Francesco; Paradisi, Paolo; Gorenflo, Rudolf. 1998. Probability Distributions Generated by Fractional Diffusion Equations.
- Mainardi, Francesco; Luchko, Yuri; Pagnini, Gianni. 2001. The Fundamental Solution of the Space-Time Fractional Diffusion Equation. *Fractional Calculus and Applied Analysis*, V.2 N.2 p.153-192.
- Markowitz, Harry M. 1952. Portfolio Selection. *The Journal of Finance*, V.7 N.1 Mar. Reprinted *in* Elton; Gruber, editors (1972).
- Markowitz, Harry M. 1976. Investment for the Long Run: New Evidence for an Old Rule. *The Journal of Finance*, V.31 N.5 p.1273-1286 Dec.
- Markowitz, Harry M. 1977. An Algorithm for Finding Undominated Portfolios. *in* Levy; Sarnat, editors (1977).
- Markowitz, Harry M. 1990. Normative Portfolio Analysis: Past, Present, and Future. *Journal of Economics and Business*, V.42 N.2 p.99-103 May.

- Markowitz, Harry M. 1991a. Portfolio Selection – Efficient Diversification of Investments. Blackwell. Cambridge, MA. Reprinted in 2000. 1st published in 1959 by John Wiley & Sons. New York, NY.
- Markowitz, Harry M. 1991b. Foundations of Portfolio Theory. The Journal of Finance, V.46 N.2 p.469-477 Jun.
- Markowitz, Harry M. 1999. The Early History of Portfolio Theory: 1600-1960. Financial Analysts Journal, V.55 N.4 p.5-16 Jul-Aug.
- Markowitz, Harry M. 2000. Mean-Variance Analysis in Portfolio Choice and Capital Markets. Revised Reissue. Frank J.Fabozzi Associates. New Hope, PA. 1st published in 1987.
- Markowitz, Harry M. 2002. Efficient Portfolios, Sparse Matrices, and Entities: A Retrospective. Operations Research, V.50 N.1 p.154-160 Jan-Feb.
- Markowitz, Harry M.; Dijk, Erik L. van. 2003. Single-Period Mean-Variance Analysis in a Changing World. Financial Analysts Journal, V.59 N.2 p.30-44 Mar-Apr.
- Marshall, J. 1938. Money and the Theory of Assets. Econometrica, V.6 Oct.
- Marshall, John F. 1994. The Role of the Investment Horizon in Optimal Portfolio Sequencing (An Intuitive Demonstration in Discrete Time). The Financial Review, V.29 N.4 p.557-576 Nov.
- Mathai, A.M.; Saxena, R.K. 1969. Distribution of a Product and the Structural Set Up of Densities. The Annals of Mathematical Statistics, V.40 N.4 p.1439-1448.
- Mathai, A.M. 1970. Statistical Theory of Distributions and Meijer's G-Function. Metron, N.28.
- Mathai, A.M. 1971. An Expansion of Meijer's G-Function in the Logarithmic Case with Applications. Mathematische Nachrichten, V.48 N.1-6 p.129-139.
- Mathai, A.M. 1972. Products and Ratios of Generalized Gamma Variates. Scandinavian Actuarial Journal, V.55, p.193-198.
- Mathai, A.M. 1973. A Few Remarks on the Exact Distributions of Certain Multivariate Statistics - II. in Kabe, D.G.; Gupta, R.P. (editors). Multivariate Statistical Inference. North-Holland. Amsterdam.
- Mathai, A.M.; Saxena, R.K. 1973. Generalized Hypergeometric Functions with Applications in Statistics and Physical Sciences. Lecture Notes in Mathematics, V.348. Springer-Verlag. Berlin.
- Mathai, A.M.; Saxena, R.K. 1978. The H-Function with Applications in Statistics and other Disciplines. Wiley. New York, NY.
- Mathai, A.M. 1982. Storage Capacity of a Dam with Gamma Type Inputs. Annals of the Institute of Statistical Mathematics, V.34 p.591-597.
- Mathai, A.M. 1983. On Linear Combinations of Independent Exponential Variables. Communications on Statistics Theory and Methods, V.12 N.6 p.625-632.
- McEnally, Richard W. 1985. Time Diversification: Surest Route to Lower Risk? The Journal of Portfolio Management, Fall.
- Mehra, Rajnish; Prescott, Edward C. 1985. The Equity Premium: A Puzzle. Journal of Monetary Economics, V.15 N.2 p.145-161 Mar.
- Mehra, Rajnish; Prescott, Edward C. 1988. The Equity Premium: A Solution? Journal of Monetary Economics, V.22 N.1 p.133-136 Jul.
- Mehra, Rajnish. 2003. The Equity Premium: Why Is It a Puzzle? Financial Analysts Journal, V.59 N.1 p.54-69 Jan-Feb.
- Mellin, HJ. 1910. Abriß einer einheitlichen Theorie der Gamma- und der Hypergeometrischen Funktionen. Math. Ann., V.68 p.305-337.
- Merton, Robert C. 1969. Lifetime Portfolio Selection under Uncertainty: the Continuous-Time Case. The Review of Economic and Statistics, V.51 N.3 p.247-257 Aug.
- Merton, Robert C. 1971. Optimum Consumption and Portfolio Rules in a Continuous-Time Model. Journal of Economic Theory, V.3 p.373-413 Dec. Reprinted with *errata* in Ziemba; Vickson, editors (1975). Reprinted in Merton (1992).
- Merton, Robert C. 1972. An Analytic Derivation of the Efficient Portfolio Frontier. Journal of Financial and Quantitative Analysis, V.7 N.4 p.1851-1872 Sep.
- Merton, Robert C. 1973. An Intertemporal Capital Asset Pricing Model. Econometrica, V.41 N.5 p.867-887 Sep. Reprinted in Bicksler; Friend, editors (1977).
- Merton, Robert C.; Samuelson, Paul A. 1974. Fallacy of Log-Normal Approximation to Optimal Portfolio Decision-Making Over Many Periods. Journal of Financial Economics, V.1 N.1 p.67-94 May. Reprinted in Bicksler; Friend, editors (1977).
- Merton, Robert C. 1975. Theory of Finance from the Perspective of Continuous Time. Journal of Financial and Quantitative Analysis, V.10 N.4 p.659-674 Nov.
- Merton, Robert C. 1982. On the Mathematics and Economics Assumptions of Continuous-Time Models. in Sharpe; Cootner, editors (1982, p.19-51).
- Merton, Robert C. 1992. Continuous Time Finance. Basil Blackwell. Oxford, UK. 1st ed. in 1990.
- Michaud, Richard O. 1998. Efficient Asset Management. Harvard Business School Press. Boston, MA.

- Michaud, Richard O. 2003. A Practical Framework for Portfolio Choice. *Journal of Investment Management*, V.1 N.2 p.1-16.
- Miller, Merton H. 1999. The History of Finance. *The Journal of Portfolio Management*, V.25 N.4 p.95-101.
- Monoyios, Michael. 2001. Finite Horizon Portfolio Selection with Transaction Costs.
- Morgan, J.P.; Reuters. (organizers). 1996. RiskMetrics – Technical Document. 4th ed.
- Mossin, Jan. 1968. Optimal Multiperiod Portfolio Policies. *The Journal of Business*, V.41 N.2 p.215-229 Apr.
- Moschopoulos, P.G. 1985. The Distribution of the Sum of Independent Gamma Random Variables. *Ann. Inst. Statist. Math. Part A*, N.37 p.541-544.
- Musumeci, Jim; Musumeci, Joe. 1999. A Dynamic-Programming Approach to Multiperiod Asset Allocation. *Journal of Financial Services Research*, V.15 N.1 p.5-21. Feb.
- Nair, V.C. 1971. On the Laplace Transform - I. *Portugaliae Mathematica*, V.30 N.1 p.57-69.
- Nantell, Timothy J.; Price, Barbara. 1979. An Analytical Comparison of Variance and Semivariance Capital Market Theories. *Journal of Financial and Quantitative Analysis*, V.14 N.2 p.221-242 Jun.
- Nawrocki, David N. 1999. A Brief History of Downside Risk Measures. *Journal of Investing*, V.8 N.3 p.9-26 Fall.
- Nawrocki, David N. 2000. Portfolio Optimization, Heuristics and the 'Butterfly Effect'. *Journal of Financial Planning*, V.13 N.2 p.68-78 Feb.
- Oberhettinger, Fritz. 1974. *Tables of Mellin Transforms*. Springer-Verlag. Berlin.
- Ogryczak, Włodzimierz; Ruszczyński, Andrzej. 1997a. From Stochastic Dominance to Mean-Risk Models: Semideviations as Risk Measures. *International Institute for Applied Systems Analysis*, IR-97-027 Jun.
- Ogryczak, Włodzimierz; Ruszczyński, Andrzej. 1997b. On Stochastic Dominance and Mean-Semideviation Models. *International Institute for Applied Systems Analysis*, IR-97-043 Jul.
- Ohlson, J.A.; Ziemba, W.T. 1976. Portfolio Selection in a Lognormal Market When the Investor has a Power Utility Function. *Journal of Financial and Quantitative Analysis*, V.11 N.1 p.57-71.
- Park, Chan S.; Sharp-Bette, Gunter P. 1990. *Advanced Engineering Economics*. John Wiley & Sons. New York, NY.
- Pedersen, Christian S.; Satchell, Stephen E. 1998. An Extended Family of Financial-Risk Measures. *The Geneva Papers on Risk and Insurance Theory*, V.23 p.89-117.
- Pedersen, Christian S.; Satchell, Stephen E. 1999. Choosing the Right Measure of Risk: A Survey. *in* Dahiya, Shri B. (editor). 1999. *The Current State of Economic Science*. Spellbound. Rohtak.
- Pedersen, Christian S.; Satchell, Stephen E. 2002. On the Foundation of Performance Measures Under Asymmetric Returns. *Quantitative Finance*, V.2 p.217-223.
- Pelessoni, Renato; Vieig, Paolo. 2001. Coherent Risk Measures and Upper Previsions.
- Petty, W.J.; Scott, D.F.; Bird, M.M. 1975. The Capital Expenditure Decision-Making Process of Large Corporations. *The Engineering Economist*, V.20 N.3 p.159-172 Spring.
- Pflug, Georg Ch. 2002. How to Measure Risk?.
- Philippatos, George C.; Gressis, Nicolas. 1975. Conditions of Equivalence Among E-V, SSD, and E-H Portfolio Selection Criteria: The Case for Uniform, Normal and Lognormal Distributions. *Management Science*, V.21 N.6 p.617-625 Feb.
- Platen, Eckhard. 2002. *Arbitrage in Continuous Complete Markets*.
- Pliska, Stanley R. 1997. *Introduction to Mathematical Finance: Discrete Time Models*. Blackwell. Malden, MA.
- Porter, R. Burr. 1973. An Empirical Comparison of Stochastic Dominance and Mean-Variance Portfolio Choice Criteria. *Journal of Financial and Quantitative Analysis*, V.8 N.4 p.587-608 Sep.
- Porter, R. Burr. 1974. Semivariance and Stochastic Dominance: A Comparison. *The American Economic Review*, V.64 N.1 p.200-204 Mar.
- Powell, M.J.D. 1981. *Approximation Theory and Methods*. Cambridge University Press. Cambridge. Reprinted in 1996.
- Price, Kelly; Price, Barbara; Nantell, Timothy J. 1982. Variance and Lower Partial Moment Measures of Systematic Risk: Some Analytical and Empirical Results. *The Journal of Finance*, V.37 N.3 p.843-855 Jun.
- Provost, S.B. 1989. On Sums of Independent Gamma Random Variables. *Statistics*, V.20 p.583-591.
- Prudnikov, A.P.; Brychkov, Yu.A.; Marichev, O.I. 1990. *Integrals and Series – More Special Functions*, V.3. Gordon and Breach. New York, NY.
- Pulley, Lawrence B. 1983. Mean-Variance Approximations to Expected Logarithmic Utility. *Operations Research*, V.31 N.4 p.685-696 Jul-Aug.
- Pulley, Lawrence B. 1985. Mean-Variance Versus Direct Utility Maximization: A Comment. *The Journal of Finance*, V.40 N.2 p.601-602 Jun.
- Pulley, Lawrence B.; Epps, T.W. 1990. Implementing the Growth-Optimal Policy for Choosing Portfolios. *Journal of Economics and Business*, V.42 N.2 p.153-164 May.
- Ratner, Mitchell; Arbeláez, Harvey; Leal, Ricardo. 1997. A Time Diversification Approach to the Emerging Stock Markets of Latin America: Argentina, Brazil, Chile, and Mexico. *The International Executive*, V.39 N.5 p.707-723 Sep-Oct.

- Rockafellar, R.Tyrrel; Uryasev, Stanislav. 1999. Optimization of Conditional Value-at-Risk.
- Rockafellar, R.Tyrrel; Uryasev, Stanislav. 2002. Conditional Value-at-Risk for General Loss Distributions. *Journal of Banking & Finance*, V.26 p.1443-1471.
- Rockafellar, R.Tyrrel; Uryasev, Stanislav; Zabarankin, Michael. 2003. Deviation Measures in Risk Analysis and Optimization. Research Report, Risk Management and Financial Engineering Lab, University of Florida.
- Rodriguez, Gerardo José Lemus. 1999. Portfolio Optimization with Quantile-based Risk Measures. PhD Thesis, Massachusetts Institute of Technology.
- Roll, Richard. 1973. Evidence on the 'Growth-Optimum' Model. *The Journal of Finance*, V.28 N.3 p.551-556 Jun.
- Ross, Stephen A. 1978. Mutual Fund Separation in Financial Theory – The Separating Distributions. *Journal of Economic Theory*, V.17 N.2 p.254-286 Apr.
- Ross, Stephen A. 1982. On the General Validity of the Mean-Variance Approach in Large Markets. *in Sharpe; Cootner, editors (1982, p.52-84).*
- Ross, S. 1999. Adding Risks: Samuelson's Fallacy of Large Numbers Revisited. *Journal of Financial and Quantitative Analysis*, V.34 N.3 p.323-339 Sep.
- Rotar, Vladimir. 2003. On Optimal Investment in the Long Run: Rank Dependent Expected Utility as a "Bridge" between the Maximum-Expected-Log and Maximum-Expected-Utility Criteria.
- Roy, A.D. 1952. Safety First and the Holding of Assets. *Econometrica*, V.20 N.3 p.431-449 Jul.
- Rubinstein, Mark. 1974. A Discrete-Time Synthesis of Financial Theory: Part 1. Optimal Decision and Sharing Rules. Research Program in Finance. University of California, Berkeley.
- Rubinstein, Mark. 1974. A Discrete-Time Synthesis of Financial Theory: Part 2. Valuation and Efficiency. Research Program in Finance. University of California, Berkeley.
- Rubinstein, Mark. 1974. A Discrete-Time Synthesis of Financial Theory: Part 3. Extensions and Prospective. Research Program in Finance. University of California, Berkeley.
- Rubinstein, Mark. 1976. The Strong Case for the Generalized Logarithmic Utility Model as the Premier Model of Financial Markets. *The Journal of Finance*, V.31 N.2 p.551-571 May. Reprinted *in Levy; Sarnat, editors (1977).*
- Rubinstein, Mark. 2002. Markowitz's 'Portfolio Selection': a Fifty-Year Retrospective. *The Journal of Finance*, V.57 N.3 p.1041-1045 Jun.
- Samuelson, Paul A. 1969. Lifetime Portfolio Selection by Dynamic Stochastic Programming. *The Review of Economic and Statistics*, V.51 N.3 p.139-246 Aug. Reprinted *in Ziemba; Vickson, editors (1975).* Reprinted *in Elton; Gruber, editors (1972).*
- Samuelson, Paul A. 1971. The "Fallacy" of Maximizing the Geometric Mean in Long Sequences of Investing or Gambling. *Proceedings of the National Academy of Sciences*, V.68 p.2493-2496.
- Samuelson, Paul A. 1974. Challenge to Judgement. *The Journal of Portfolio Management*, V.1 p.17-19.
- Samuelson, Paul A.; Merton, Robert C. 1974. Generalized Mean-Variance Tradeoffs for Best Perturbation Corrections to Approximate Portfolio Decisions. *The Journal of Finance*, V.29 N.1 p.27-40 Mar.
- Samuelson, Paul A. 1977. St. Petersburg Paradoxes: Defanged, Dissected, and Historically Described. *Journal of Economic Literature*, V.15 N.1 p.24-55 Mar.
- Samuelson, Paul A. 1979. Why We Should Not Make Mean Log of Wealth Big Though Years to Act Are Long. *Journal of Banking and Finance*, V.3 p.305-307.
- Samuelson, Paul A. 1989. The Judgement of Economic Science on Rational Portfolio Management: Indexing, Timing, and Long-Horizon Effects. *The Journal of Portfolio Management*, V.16 N.1 p.4-12 Fall.
- Samuelson, Paul A. 1990. Asset Allocation Could be Dangerous to Your Health. *The Journal of Portfolio Management*, V.16 N.3 p.5-8 Spring.
- Samuelson, Paul A. 1994. The Long-Term Case for Equities. *The Journal of Portfolio Management*, V.21 N.1 p.15-24 Fall.
- Schneider, W.R. 1986. Stable Distributions: Fox Function Representation and Generalization. *Lecture Notes in Physics* V.262. Springer-Verlag. Berlin.
- Schneider, W.R. 1987. Generalized One-Sided Stable Distributions. *in Albeverio; Blanchar; Streit, editors (1987).*
- Schneider, W.R. 1988. Generalized Levy Distributions as Limit Laws. *in Albeverio; Blanchar; Streit, editors (1987).*
- Schneller, Meir I. 1975. Mean-Variance Portfolio Composition when Investors' Revision Horizon is Very Long. *The Journal of Finance*, V.30 N.5 p.1293-1300 Dec.
- Schulmerich, Marco; Trautmann, Siegfried. 2003. Local Expected Shortfall-Hedging in Discrete Time. *European Finance Review*, V.7, p.75-102.
- Scott, David W. 1992. Multivariate Density Estimation - Theory, Practice, and Visualization. John Wiley & Sons. New York, NY.
- Securato, José Roberto. 2001. *Decisões Financeiras em Condições de Risco*. Atlas. São Paulo, SP.
- Shafer, Glenn; Vovk, Vladimir. 2001. *Probability and Finance – It's Only a Game!* John Wiley & Sons. New York, NY.

- Sharpe, William F. 1963. A Simplified Model for Portfolio Analysis. *Management Science*, V.9 N.2 p.277-293 Jan. Reprinted in Elton; Gruber, editors (1972).
- Sharpe, William F. 1964. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, V.19 N.3 p.425-442 Sep.
- Sharpe, William F.; Cootner, Cathryn M. (editors). 1982. *Financial Economics – Essays in Honor of Paul Cootner*. Prentice-Hall. Englewood Cliffs, NJ.
- Sharpe, William F. 2000. *Portfolio Theory and Capital Markets*. McGraw-Hill. New York, NY. 1st ed. in 1970.
- Silverman, B.W. 1986. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall Ltd. London.
- Sim, C.H. 1992. Point Processes with Correlated Gamma Interarrival Times. *Statistics and Probability Letters*, V.15 p.135-141.
- Simonsen, M.H.; Werlang, Sérgio R.C. 1991. Subadditive Probabilities and Portfolio Inertia. *Revista de Econometria*, V.11 N.1 p.1-19 Apr.
- Sing, Tien Foo; Ong, Seow Eng. 2000. Asset Allocation in a Downside Risk Framework. *Journal of Real Estate Portfolio Management*, V.6 N.3 p.213-223.
- Sneddon, Ian N. 1972. *The Use of Integral Transforms*. McGraw-Hill. New York, NY.
- Sørensen, Carsten. 1999. Dynamic Asset Allocation and Fixed Income Management. *The Journal of Financial and Quantitative Analysis*, V.34 N.4 p.513-531 Dec.
- Sortino, Frank A.; Meer, Robert van der. 1991. Downside Risk. *The Journal of Portfolio Management*, V.17 N.4 p.27-32 Summer.
- Sortino, Frank A.; Price, Lee N. 1994. Performance Measurement in a Downside Risk Framework. *Journal of Investing*, p.59-64 Fall.
- Sortino, Frank A.; Forsey, Hal J. 1996. On the Use and Misuse of Downside Risk. *The Journal of Portfolio Management*, V.22 N.2 p.35-42 Winter.
- Sortino, Frank A. 2001. From Alpha to Omega. in Sortino; Satchell, editors (2001).
- Sortino, Frank A.; Satchell, Stephen E. (editors). 2001. *Managing Downside Risk in Financial Markets: Theory, Practice and Implementation*. Butterworth Heinemann. Oxford.
- Srivastava, H.M.; Buschman, R.G. 1977. *Convolution Integral Equations with Special Function Kernels*. John Wiley & Sons. New York, NY.
- Srivastava, H.M.; Gupta, K.C.; Goyal, S.P. 1982. *The H-Functions of One and Two Variables with Applications*. South Asian. New Delhi.
- Stacy, E.W. 1962. A Generalization of the Gamma Distribution. *Annals of Mathematical Statistics*, V.33 p.1187-1192.
- Stevens, Guy V.G. 1972. On Tobin's Multiperiod Portfolio Theorem. *The Review of Economic Studies*, V.39 N.4 p.461-468 Sep.
- Stone, Bernell K. 1973. A General Class of Three-Parameter Risk Measures. *The Journal of Finance*, V.28 N.3 p.675-685 Jun.
- Stutzer, Michael. 2004. Misconceptions about Optimal Equity Allocation and Investment Horizon: A Comment. *Financial Analysts Journal*, V.60 N.1 p.12-14 Jan-Feb.
- Szegö, Giorgio. 2002. Measures of Risk. *Journal of Banking & Finance*, V.26 p.1253-1272.
- Verity, Dominic; Carmody, Sean. 2000. Risk Management beyond VaR – Understanding Low Probability High Impact Events.
- Tenney, Mark S. 1995. Dynamic Dynamic-Programming Solutions for the Portfolio of Risky Assets.
- Tetens, J.N. 1789. *Einleitung zur Berechnung der Leibrenten und Anwartschaften*. Leipzig.
- Thorley, Steven R. 1995. The Time-Diversification Controversy. *Financial Analysts Journal*, V.51 N.3 p.68-76 May-Jun.
- Thorp, Edward O. 1971. Portfolio Choice and the Kelly Criterion. *Business and Economics Statistics Section. Proceedings of the American Statistical Association*, p.215-224. Reprinted in Ziemba; Vickson, editors (1975).
- Thorp, Edward O. 1997. The Kelly Criterion in Blackjack, Sports Betting, and the Stock Market. *The 10th International Conference on Gambling and Risk Taking*.
- Tobin, James. 1958. Liquidity Preference as Behavior Towards Risk. *The Review of Economics Studies*, V.25 N.67 p.65-87 Feb. Reprinted in Elton; Gruber, editors (1972).
- Tobin, James. 1965. The Theory of Portfolio Selection. in Hahn, F.H.; Brechling, F.P.R. (editors). 1965. *The Theory of Interest Rates*. Macmillan. London.
- Vander Weide, James H.; Peterson, David W.; Maier, Steven F. 1977. A Strategy Which Maximizes the Geometric Mean Return on Portfolio Investments. *Management Science*, V.23 N.10 p.1117-1123 Jun.
- Vander Weide, James H.; Peterson, David W.; Maier, Steven F. 1978. Reply to Aucamp. *Management Science*, V.24 N.8 p.860 Apr.
- Widder, David V. 1989. *Advanced Calculus*. 2nd ed. Dover. Mineola, NY. 1st ed. published in 1947 by Prentice-Hall. Englewood Cliffs, NJ.

- Wilcox, Jarrod W. 1998. Investing at the Edge. *The Journal of Portfolio Management*, V.24 N.3 p.9-21 Spring.
- Wilcox, Jarrod W. 2000. Better Risk Management. *The Journal of Portfolio Management*, V.26 N.4 p.53-64 Summer.
- Wilcox, Jarrod W. 2003. Harry Markowitz and the Discretionary Wealth Hypothesis. *The Journal of Portfolio Management*, V.29 N.3 Spring.
- Williamson, Robert Charles. 1989. Probabilistic Arithmetic. PhD thesis, University of Queensland.
- Wolfe, Philip. 1957. A Simplex Method for Quadratic Programming. Unpublished technical report. RAND P-1025.
- Wolfe, Philip. 1959. The Simplex Method for Quadratic Programming. *Econometrica*, V.27 p.382-398 Jul.
- Yamai, Yasuhiro; Yoshida, Toshinao. 2001. On the Validity of Value-at-Risk: Comparative Analysis with Expected Shortfall. IMES Discussion Paper Series 2001-E-4, March.
- Yamai, Yasuhiro; Yoshida, Toshinao. 2002. Comparative Analysis of Expected Shortfall and Value-at-Risk (2): Expected Utility Maximization and Tail Risk. *Monetary and Economic Studies*, Apr.
- Young, Williams E.; Trent, Robert H. 1969. Geometric Mean Approximations of Individual Security and Portfolio Performance. *Journal of Financial and Quantitative Analysis*, V.4 N.2 p.179-199 Jun.
- Yu, Lian; Zhang, Shuzhong; Zhou, Xun Yu. 2003. A Downside Risk Analysis based on Financial Index Tracking Models.
- Yu, Po-Lung. 1985. Multiple-Criteria Decision Making – Concepts, Techniques, and Extensions. Plenum Press. New York, NY.
- Zeleny, Milan. 1981. Multidimensional Measure of Risk: Prospect Rating Vector (PRV). *Lecture Notes in Economics and Mathematical Systems*, V.190. Springer-Verlag. Berlin.
- Ziemba, William T. 1972. Note on ‘Optimal Growth Portfolios When Yields Are Serially Correlated’. *Journal of Financial and Quantitative Analysis*, V.7 N.4 p.1995-2000 Sep.
- Ziemba, W.T. 1974. Choosing Investment Portfolios When the Returns Have Stable Distributions. *in* Hammer, P.L.; Zoutendijk, G. (editors). 1974. *Mathematical Programming in Theory and Practice*. North-Holland. Amsterdam. Reprinted *in* Ziemba; Vickson, editors (1975).
- Ziemba, W.T.; Parkan, C.; Brooks-Hill, R. 1974. Calculation of Investment Portfolios with Risk Free Borrowing and Lending. *Management Science*, V.21 N.2 p.209-222 Oct.
- Ziemba, W.T.; Vickson, R.G. (editors). 1975. *Stochastic Optimization Models in Finance*. Academic Press. San Diego, CA.
- Zolotarev, V.M. 1986. One-Dimensional Stable Distributions. *Translations of Mathematical Monographs*, V.65. American Mathematical Society. Providence. Translated from the Russian by McFaden, H.H.