



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística
Estágio Supervisionado 2

Valor em Risco Operacional com Cópulas de Valores Extremos

Diego Rodrigues Oliveira - 08/27941

Relatório Final

Orientadora: Cira E. Guevara Otiniano

Brasília
Junho de 2012

Diego Rodrigues Oliveira

Valor em Risco Operacional com Cópulas de Valores Extremos

Relatório Final de conclusão de curso,
apresentado à disciplina Estágio Supervisionado
II do Curso de Bacharelado em Estatística –
Instituto de Ciências Exatas da Universidade de
Brasília.

Orientadora:
Cira Etheowalda Guevara Otiniano

Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Brasília – DF, 18 de junho de 2012

Sumário

Listas de Ilustrações	v
L.1 – Lista de Tabelas.....	v
L.2 – Lista de Figuras	v
Lista de Abreviações	vi
Resumo	vii
1 Introdução	1
2 O Novo Acordo de Basileia	4
2.1 Introdução.....	4
2.2 Risco Operacional	6
2.1.1 O Método de Indicador Básico	7
2.1.2 O Método Padronizado	7
2.1.3 Métodos de Mensuração Avançada	9
3 Valor em Risco Operacional	11
3.1 Introdução.....	11
3.2 Fundamentação Teórica	11
3.3 Medidas de Risco Coerentes	12
3.4 Comonotonicidade	14
4 Distribuições de Cauda Pesada	15
4.1 Introdução.....	15
4.2 Fundamentação Teórica	16
4.2.1 Distribuições α -estáveis	16
4.2.2 Teoria do Valor Extremo	17

4.2.3	Distribuições de Variação Regular	21
5	Cópuas	22
5.1	Introdução.....	22
5.2	Fundamentação Teórica	22
5.3	Limites de Fréchet-Hoeffding	24
5.4	Dependência Caudal.....	25
5.5	Famílias de Cópuas	27
5.5.1	Introdução	27
5.5.2	Cópuas Elípticas	27
5.5.3	Cópuas Arquimedianas	29
6	Cópuas de Valores Extremos	32
6.1	Introdução.....	32
6.2	Fundamentação Teórica	32
6.3	Estimação da Cauda Conjunta.....	35
6.4	Dependência Caudal.....	36
7	Estimadores	37
7.1	Introdução.....	37
7.2	Estimação Paramétrica	38
7.3	Estimação Não Paramétrica	41
8	Aplicação Prática	44
8.1	Introdução.....	44
8.2	Análise Descritiva	44
8.3	Estimação dos Parâmetros.....	49
8.4	Resultados	51
8.5	Observações	53
	Conclusão	54

Referências	55
Apêndice	56
Código SAS	56
Código R.....	57

Listas de Ilustrações

L.1 – Lista de Tabelas

2.1 – Fatores Beta para as Linhas de Negócio de uma Instituição Financeira	pg.7
8.1 – Estatísticas Descritivas para os Conjuntos de Dados FLT e DNN	pg.43
8.2 – Estimativas dos Parâmetros da GEV	pg.48
8.3 - Estimativa do parâmetro da cópula, função de dependência e dependência caudal .	pg.49
8.4 – Valor em Risco para FLT e DNN	pg.50
8.5 – Estrutura de dependência	pg.50
8.6 – Valor em Risco de FLT e DNN considerando a estrutura de dependência	pg.51

L.2 – Lista de Figuras

8.1 - Série temporal dos dados de máximos mensais FLT	pg.44
8.2 - Série temporal dos dados de máximos mensais DNN	pg.45
8.3 - Histograma dos dados de máximos mensais FLT	pg.46
8.4 - Histograma dos dados de máximos mensais DNN	pg.46
8.5 - Gráfico de dispersão dos dados de máximos mensais de FLT e DNN	pg.48
8.6 - QQ-Plots de FLT (x) e DNN (y)	pg.50
8.7 - F.d. empírica (azul) x estimada (vermelho) de FLT e DNN	pg.52

Lista de Abreviações

- **AMA:** *Advanced Method Approach*
- **BMM:** *Block Maxima Method*
- **CBSB:** Comitê da Basileia sobre Supervisão Bancária
- **edf:** *Excess Distribution Function*
- **fd:** Função de Distribuição
- **GEV:** *Generalized Extreme Value distribution*
- **GPD:** *Generalized Pareto Distribution*
- **IF:** Instituição Financeira
- **iid:** Independente e Identicamente Distribuída
- **mef:** *Mean Excess Function*
- **POT:** *Peaks Over Threshold*
- **RC:** Risco de Crédito
- **RM:** Risco de Mercado
- **RO:** Risco Operacional
- **TVE:** Teoria de Valor Extremo
- **va:** Variável Aleatória
- **VaR:** *Value at Risk*
- **VaR_{Op}:** Valor em Risco Operacional

Resumo

Desde Mandelbrot (1963) e Fama (1965), dados financeiros e de riscos, como dados de perdas, taxas de câmbio, índices econômicos e retornos em geral, são, em sua maioria, modelados por distribuições de cauda pesada. A partir da modelagem das distribuições de perdas em Risco Operacional é necessário construir um modelo conjunto para que se estabeleçam pressupostos apropriados sobre a dependência entre as variáveis de perdas e também para que se possa obter o valor em risco operacional (VaR_{Op}) da perda operacional total. As funções cópula são reconhecidas como ferramentas poderosas para tais fins. Portanto, através de cópulas de valores extremos pretende-se obter o modelo conjunto associado às distribuições de perdas, permitindo o cálculo do capital regulatório mínimo, associado ao VaR_{Op} , que a instituição financeira deve alocar para perdas operacionais e o estabelecimento da estrutura de dependência entre os tipos de perdas.

1 Introdução

O Comitê da Basileia sobre a Supervisão Bancária, estabelecido em 1975 e que já havia feito um acordo em 1988, devido a grandes prejuízos, chegando à falência, sofridos por instituições financeiras na década de 90 e a fim de evitar mais prejuízos, realizou o primeiro ciclo de propostas para revisar a estrutura de adequação de capital que deve ser alocado a riscos sofridos pelas instituições financeiras em 1999. Em 2004, surge a estrutura revisada, programada para estabelecer níveis mínimos de capital para bancos internacionalmente ativos. Também surge uma nova classe de risco, o Risco Operacional (RO). Essa nova estrutura é chamada de Novo Acordo da Basileia ou Basileia II.

O Modelo de Distribuições de Perdas, principal método previsto em Basileia II para cálculo do capital regulatório mínimo alocado à perdas operacionais, baseia-se nos processos estocásticos inerentes às perdas oriundas de RO, estando este modelo relacionado às distribuições discretas (frequência) e contínuas (severidade). As perdas são separadas nesses dois processos devido às componentes que regem essas distribuições são por hipótese independentes. A convolução desses dois processos é feita preferencialmente por simulação de Monte Carlo, essa união resulta na distribuição de perdas de uma linha de negócio e/ou evento de perda.

Basileia II define oito linhas de negócios e sete tipos de eventos de perdas. A metodologia mais utilizada atualmente pelas instituições financeiras para o cálculo do requerimento total de capital do banco para perdas operacionais é, basicamente, a soma dos requerimentos obtidos individualmente para cada tipo de perda operacional. O problema é que, ao realizar tal cálculo, o banco estará considerando a hipótese de que as correlações entre os diferentes tipos de perdas operacionais são perfeitas. Isso implica que todos os eventos extremos devem ocorrer ao mesmo tempo em todas as classes de risco, o que é uma hipótese irreal. Portanto, esse cálculo superestima o valor que deve ser alocado.

Portanto, na estimação do Valor em Risco Operacional, que resulta capital regulatório mínimo, precisamos de um modelo conjunto para o vetor aleatório de perdas, ou seja, pressupostos concretos sobre a dependência entre as classes de perdas devem ser feitos. Devido ao leque de modelos de distribuição conjunta específicos para o vetor aleatório de perdas operacionais, Basiléia II apenas sugere o estudo da diversificação da redução de capital alocado por meio de pressupostos de correlação apropriados.

A função cópula tem sido reconhecida como uma ferramenta poderosa para modelar dependência entre variáveis aleatórias. A palavra cópula foi empregada pela primeira vez no sentido probabilístico em um teorema conhecido como Teorema de Sklar (1959), descrevendo as funções que “ligam/conectam” as distribuições univariadas à distribuição multivariada.

A ideia essencial de cópulas é que a distribuição conjunta pode ser fatorada em distribuições marginais e uma função de dependência, a qual é chamada de cópula. A relação de dependência é inteiramente determinada pela função cópula, enquanto que localização, escala e forma são inteiramente determinadas pelas marginais. Como as distribuições de perdas de risco operacional são descritas na literatura existente como de cauda pesada, daremos maior atenção para cópulas de valores extremos.

Assim, tem-se por objetivo que as funções cópulas nos permitam elaborar pressupostos de correlações apropriados sobre a dependência de tipos de perdas operacionais, para que dessa forma consigamos obter um capital regulatório adequado, suficiente, para cobrir eventos de perdas decorrentes do risco operacional sem ser excessivamente conservador.

Nesse sentido, o trabalho de Neslehová, Embrechts e Demoulin (2006) prevê o uso de distribuições de cauda pesada para a modelagem de perdas operacionais, o que nos faz supor que os dados reais com que trabalharemos sejam modelados por tais distribuições. O trabalho de Lu, Tian e Zhang (2008) estabelecem a metodologia para o cálculo do *Value-at-Risk* (*VaR*) de uma cópula de valores extremos, que será replicada neste trabalho.

As próximas seções desse trabalho se desenvolvem da seguinte forma. A seção 2 trata do Novo Acordo de Basiléia, contando um pouco de sua história e definições, além de mostrar as três metodologias usadas para o cálculo do capital regulatório. Na seção 3, apresenta-se o Valor em Risco Operacional, sua definição, propriedade de medidas de risco coerentes e a comonotonicidade. A seção 4 concerne na definição de distribuições de cauda

pesada, explanação da Teoria do Valor Extremo e suas principais distribuições. Na seção 5 apresentamos as principais noções de cópulas e suas famílias. Nas seções 6 e 7 estão contidas as fundamentações teóricas sobre cópulas de valores extremos e métodos de estimação, respectivamente.

2 O Novo Acordo de Basiléia

2.1 Introdução

A década de 80 foi marcada por uma crise econômica mundial. Após intensas investigações, autoridades financeiras internacionais chegaram à conclusão de que tal crise se desenvolveu a partir da falta de uma regulamentação sólida para reserva de capital alocado às possíveis perdas que podem ser sofridas por instituições financeiras. Essa possibilidade de perda foi, então, chamada de *risco*, assunto, este, que vem sendo estudado e debatido intensamente desde então.

Na tentativa de evitar maiores prejuízos, o Comitê da Basiléia sobre a Supervisão Bancária (CBSB) em 1988, estabeleceu o que foi chamado de Acordo de Basiléia, ou Basiléia I. Este acordo concerne basicamente na definição de Risco de Crédito (RC) e Risco de Mercado (RM), definição de metodologias e normatizações necessárias para o cálculo do capital regulatório mínimo que deve ser alocado por uma instituição financeira internacionalmente ativa.

O esforço conjunto realizado por instituições financeiras de todo o mundo, não sendo restrito apenas aos países participantes do CBSB, na tentativa de implantar o que era ditado em Basiléia I, se mostrou recompensador. Isso porque aquela nova estrutura, apoiada em um organismo internacional de controle, reduzia de forma significativa grandes prejuízos que instituições financeiras poderiam sofrer e um consequente efeito “dominó”, caracterizando uma crise econômica internacional.

Apesar de todo este esforço, a globalização ocorrida na década de 90, acompanhada de grandes avanços tecnológicos e consequente dinamização das atividades, teve um grande impacto negativo sobre essa estrutura. Isso ocasionou enormes perdas. Assim, se fez necessária a revisão de Basiléia I.

O CBSB resolveu publicar um ciclo de propostas para revisar a estrutura de adequação de capital em junho de 1999. Todos os países-membros participaram de um processo consultivo, onde foi enfatizada a criação de propostas para as autoridades de supervisão em todo o mundo. O CBSB, subsequentemente, liberou propostas adicionais para consulta em janeiro de 2001 e abril de 2003 e, ainda, conduziu três estudos de impacto quantitativo relacionados às suas propostas. Como resultado desse esforço, aprimoramentos valiosos foram feitos às propostas originais.

Surge, então, o Novo Acordo da Basiléia, ou Basiléia II, em 2004. Sendo uma declaração aceita por todos os membros do CBSB. Ele estabelece os detalhes da estrutura acordada para mensurar a estrutura de capital e o padrão mínimo a ser atingido que as autoridades nacionais de supervisão proporão para adoção em seus países. Esta estrutura e o padrão que ela contém foram endossados pelos Presidentes dos Bancos Centrais e pelas Autoridades de Supervisão Bancária dos países do Grupo dos Dez.

O objetivo fundamental do trabalho do comitê para revisar o Acordo da Basiléia de 1988 foi desenvolver uma estrutura que fortaleça ainda mais a solidez e a estabilidade do sistema bancário internacional e, ao mesmo tempo, mantenha consistência suficiente para que o regulamento de adequação do capital não seja uma fonte significativa de desigualdade competitiva entre os bancos internacionalmente ativos.

A estrutura revisada oferece uma série de opções para determinar as exigências de capital para risco de crédito e risco operacional (RO), uma classe de risco até então desconsiderada, a fim de permitir que os bancos e as autoridades de supervisão selecionem abordagens que sejam mais adequadas para as suas operações e para a infraestrutura do seu mercado financeiro. Além disso, ela está baseada em uma estrutura formada por três pilares, onde se enfatiza a complementação do Cálculo do Capital Regulatório Mínimo (Primeiro Pilar) através de um Processo de Supervisão Bancária (Segundo Pilar) e da Disciplina de Mercado, ou Transparência, (Terceiro Pilar). Este trabalho se concentra no primeiro pilar e para desenvolvimento dessa seção [2] foi de essencial importância.

2.2 Risco Operacional

De forma geral, risco é entendido como uma condição onde existe uma possibilidade de um evento adverso ao esperado ocorrer. Basiléia II prevê e define três tipos de riscos intrínsecos a uma instituição financeira: Risco de Crédito, Risco de Mercado e Risco Operacional. Cruz (2002) cita um levantamento onde se estima que 50% das perdas de uma instituição financeira se refiram ao Risco de Crédito, 15% ao Risco de Mercado e 35% ao Risco Operacional. Por ser um tema relativamente recente, mas de grande importância, nos concentraremos neste último tipo de risco, cuja definição descrita em Basiléia II é transcrita abaixo.

Definição 2.1: *O risco operacional é definido como o risco de perda resultante de pessoas, sistemas e processos internos inadequados ou deficientes, ou de eventos externos.*

Um exemplo de grande prejuízo sofrido por uma instituição financeira devido a um evento de risco operacional é o caso do banco Barings, onde um operador levou o banco à falência após ocultar a posição de futuros e derivativos na Ásia. A partir daí, percebeu-se que perdas relacionadas às falhas humanas têm um alto potencial de afetar o resultado de uma instituição financeira e, portanto, deveriam ser estudadas profundamente.

A estrutura para as metodologias de mensuração do risco operacional obtida após este necessário estudo e descrita no Novo Acordo de Basiléia apresenta três métodos para calcular encargos de capital de risco operacional em um grau de sofisticação e sensibilidade crescente:

- Método de Indicador Básico
- Método Padronizado
- Métodos de Mensuração Avançada (AMA)

Basiléia II também sugere que as instituições financeiras utilizem os dois primeiros métodos citados, seguidamente, enquanto desenvolvem sistemas e práticas de mensuração de risco operacional mais sofisticadas. Também é permitida a utilização do Método de Indicador Básico ou do Método Padronizado para algumas partes de suas operações e o AMA para

outras, desde que determinados critérios mínimos sejam cumpridos. Esses métodos e seus critérios de qualificação são apresentados abaixo.

2.1.1 O Método de Indicador Básico

Neste método, os bancos devem manter um capital para risco operacional igual à média de uma porcentagem fixa (designada alfa) da receita bruta anual positiva dos três anos anteriores. Caso em algum desses anos a receita bruta for negativa ou zero este deve ser excluído tanto do numerador quanto do denominador ao calcular a média. O encargo pode ser expresso conforme abaixo:

$$C = \alpha \frac{\sum_{i=1}^n R_i}{n} \quad (2.1)$$

Onde:

C = encargo de capital

R = receita bruta anual, quando positiva, durante os três anos anteriores

n = número dos três anos anteriores pelo qual a receita bruta foi positiva

$\alpha = 15\%$, que é estabelecido pelo comitê, em relação ao nível de todo o setor do capital exigido para o nível de todo o setor indicado.

2.1.2 O Método Padronizado

Na implementação do Método Padronizado, as atividades das instituições financeiras foram divididas em oito linhas de negócio: financiamento para pessoa jurídica, negociações e vendas, banco de varejo, banco comercial, pagamento e liquidação, serviços de agência, administração de ativos e corretagem no varejo. Tal divisão tem como objetivo principal

estabelecer diferentes níveis de risco entre as linhas de negócio e relacioná-los com suas respectivas receitas brutas, gerando um capital de alocação mais coerente com o perfil da instituição financeira.

O capital alocado para perdas operacionais em cada linha de negócios é calculado multiplicando-se a receita bruta por um fator (designado beta) atribuído a essa linha de negócios. Os fatores betas foram estipulados através da experiência de perdas operacionais em cada linha de negócios e o nível agregado da receita bruta dessa linha de negócios. Deve-se observar que, no método padronizado, a receita bruta é mensurada para cada linha de negócios, não para a instituição toda.

Tabela 2.1 – Fatores Beta para as Linhas de Negócio de uma Instituição Financeira

Linhas de Negócios	Fatores Beta
Financiamento para pessoa jurídica	18%
Negociação e vendas	18%
Banco de varejo	12%
Banco Comercial	15%
Pagamento e liquidação	18%
Serviços de agência	15%
Administração de ativos	12%
Corretagem no varejo	12%

Fonte: *Bank for International Settlements*

O capital total que deve ser alocado é calculado como a média de três anos da soma dos encargos de capital regulamentares em cada uma das linhas de negócio em cada ano. Em qualquer ano determinado, os encargos de capital negativos (resultante da receita bruta negativa), em qualquer linha de negócios, podem compensar os encargos de capital positivos em outras linhas de negócios dentro de um determinado ano que for negativo, o dado para esse ano será zero. O total do encargo de capital pode ser expresso como:

$$K = \left\{ \sum_{i=1}^3 \max_i \left[\sum_{j=1}^8 R_j \beta_j, 0 \right] \right\} / 3 \quad (2.2)$$

Onde:

K = o encargo de capital

R_j = a receita bruta de um determinado ano para a linha de negócio j .

β_j = uma porcentagem fixa, estabelecida pelo Comitê.

2.1.3 Métodos de Mensuração Avançada

Apesar do esforço feito no Método Padronizado, ao se estabelecer fatores de risco operacional apropriados para cada linha de negócios de uma instituição financeira, os dois métodos iniciais propostos em Basiléia II utilizam como base a receita bruta da instituição para estimar possíveis perdas. Existe, nisso, uma aparente contradição, quando se utiliza ganho para calcular perda. Em decorrência dessa necessidade em se lançar mão de métodos mais realistas para mensurar possíveis perdas, Basiléia II estabelece critérios para criação de tais métodos.

Um esforço inicial se dá através do estabelecimento de sete eventos de perdas operacionais principais, também chamados de eventos de perdas de nível 1:

- Fraude Interna
- Fraude Externa
- Práticas Trabalhistas e Segurança no Ambiente de Trabalho
- Clientes, Produtos e Práticas Comerciais
- Danos aos Ativos Físicos
- Ruptura nos Negócios e Falhas no Sistema
- Administração de Execução, Entrega e Processo

A intenção é que as instituições registrem suas perdas nesses eventos e utilizem a distribuição resultante para estimar perdas operacionais futuras. Considerando a contínua evolução dos métodos analíticos para risco operacional, o CBSB não especifica o método ou as premissas de distribuição usadas para gerar a mensuração de risco operacional para fins de capital regulamentar. Embora, o banco deva ser capaz de demonstrar que seu método registra

eventos de perda potencialmente críticos. Seja qual for o método usado, um banco deve demonstrar que sua mensuração de risco operacional cumpre um padrão de solidez comparável ao padrão do método com base em classificações internas para risco de crédito, isto é, comparável a um período de manutenção do investimento de um ano e um intervalo de confiança do 99,9º percentil. Além disso, qualquer sistema de mensuração de risco operacional interno deve ser consistente com os tipos de eventos de perdas definidos acima.

A instituição pode ser autorizada a usar correlações determinadas internamente em perdas de risco operacional em estimativas de risco operacional individuais, desde que ela possa demonstrar que seus sistemas para determinar as correlações são sólidos, implementados com integridade e consideram a incerteza que rodeia quaisquer estimativas de correlação. A instituição financeira deve validar suas premissas de correlação usando técnicas quantitativas e qualitativas adequadas. E, finalmente, as mensurações de risco operacional geradas internamente e usadas para fins de capital regulamentar devem ter como base um período de observação mínimo de cinco anos dos dados de perdas internas.

Assim, o objetivo desse trabalho é a implementação do *Value-at-Risk* (VaR) para perdas operacionais. Para tanto, utilizaremos distribuições de cauda pesada para ajustar as distribuições de eventos de perdas reais, já que, como afirmado por Moscadelli (2004) e Fontnouvelle (2004), após utilizarem os dados obtidos pelo CBSB no estudo inicial sobre RO, as caudas das funções de distribuição de perdas são mais bem ajustadas por distribuições de variação regular ou por distribuições tipo Pareto. A estrutura de correlação entre essas distribuições será obtida através de um estudo da função cópula, obtendo a distribuição conjunta dessas variáveis (eventos de perda). Por fim, o capital regulatório que a instituição deve alocar em perdas operacionais será obtido empregando a medida de risco VaR, diretamente na função cópula. O próximo capítulo trata essa medida de risco de forma mais detalhada.

3 Valor em Risco Operacional

3.1 Introdução

A medida de risco VaR surgiu no mercado financeiro, mais especificamente no gerenciamento de risco de mercado, no início da década de 1990 e é atualmente considerada a medida padrão para tal risco e usada intensamente na gestão de riscos. O VaR informa a perda máxima esperada ao longo de um horizonte-alvo dentro de um determinado intervalo de confiança. Mais precisamente, a perda máxima é o x_q , quantil $1 - \alpha$, da distribuição de perdas. Como consequência disso, VaR estima x_q para valores suficientemente baixos de α .

A principal referência para construção desse capítulo foi Cruz (2002).

3.2 Fundamentação Teórica

Definição 3.1 (Valor em Risco Operacional): *Dada uma distribuição de perdas L , com função de distribuição acumulada F , o VaR associado a um nível de confiança $\gamma = 1 - \alpha$ é definido por*

$$VaR_\gamma(L) = \inf\{x | P(L \leq x) \geq \gamma\} \quad (3.1)$$

A principal diferença entre os modelos de VaR de mercado e operacional está relacionada ao fato de que a aplicação da Teoria de Valores Extremos (TVE), ou outras distribuições de cauda pesada, nos faz relaxar a hipótese gaussiana em que se baseiam os modelos de VaR de mercado. Ou seja, as distribuições de perdas operacionais não são

ajustadas de forma alguma por uma distribuição normal, como se poderia esperar no risco de mercado, embora esse pressuposto muitas vezes seja contestado, até mesmo para o risco de mercado.

Além disso, o risco operacional é medido por perdas observadas, não por variações de um valor, como acontece no risco de mercado. Por essas razões, Cruz (2002) afirma que o VaR para o risco operacional satisfaz a propriedade da subaditividade, sendo, portanto, uma medida de risco coerente. A seguir, definem-se as propriedades que uma medida de risco deve satisfazer.

3.3 Medidas de Risco Coerentes

A necessidade de medidas que consigam captar o risco de forma consistente faz com que diversos estudos sejam feitos relacionados a essa necessidade. Portanto, medidas de risco vêm sendo propostas há muito tempo. Artzner et al. (1999), ao perceberem a necessidade de se formalizar um conceito sólido de medida de risco, desenvolveram uma lista de propriedades desejáveis para as quais as medidas de risco podem ser consideradas como sendo “coerentes”.

Considerando-se Ω um espaço amostral, o valor líquido final de uma posição para cada elemento de Ω é uma variável aleatória denotada por X . A variável aleatória identicamente igual a 1 é denotada por $\mathbf{1}$. A função indicadora de estado ω é denotada por $\mathbf{1}_{\{\omega\}}$.

Seja \mathcal{G} o conjunto de todos os riscos, que é o conjunto de todas as funções reais em Ω . Como Ω é finito, \mathcal{G} pode ser identificado como \mathbf{R}^n , onde $n = \text{cardinal}(\Omega)$. O cone de elementos não negativos em \mathcal{G} pode ser denotado por L_+ e o seu negativo por L_- .

Definição 3.2: *Uma medida de risco é uma função de \mathcal{G} em \mathbf{R} .*

De acordo com Artzner et al. (1999), para cada par de riscos X e Y (correlacionados ou não), onde $Y \geq X$, bem como para cada número n e para cada número positivo t , todas as relações seguintes deverão ser satisfeitas.

Sub-aditividade

$$\rho(X + Y) \leq \rho(X) + \rho(Y) \quad (3.2)$$

Essa medida tenta garantir que o risco se comporte de maneira razoável ao somar duas posições. A soma de dois ou mais riscos deve ser menor ou igual à soma dos riscos individuais. Isso permite a mensuração individual de riscos dentro da instituição financeira, uma vez que a soma das medidas individuais é conservadora.

Homogeneidade Positiva

$$\rho(tX) = t\rho(X) \quad (3.3)$$

Esta propriedade tenta garantir que, se você aumentar a sua posição, você ainda terá uma medida de risco confiável.

Invariância da Translação

$$\rho(X + n) = \rho(X) + n \quad (3.4)$$

Esta propriedade garante que quando uma quantidade n for adicionada à posição que estiver sob escrutínio, a necessidade de capital adicional será diminuída por n .

Monotonicidade

$$\rho(Y) \geq \rho(X) \quad (3.5)$$

Esta propriedade garante que o risco será maior para perdas maiores.

Definição 3.3 (Medida de Risco Coerente): *Uma medida de risco que satisfaça às propriedades: subaditividade, homogeneidade positiva, invariância da translação e monotonicidade, é uma medida de risco coerente.*

Artzner et al. (1999) mostram que a medida de VaR de mercado, em alguns casos fracassa em satisfazer a propriedade de subaditividade. Sob a condição de comonotonicidade entre as variáveis de perdas a igualdade é válida.

3.4 Comonotonicidade

Definição 3.4 (Comonotonicidade): *Dois riscos, X e Y , são ditos comonotônicos se a função de distribuição conjunta deles satisfaz*

$$F_{X,Y}(x, y) = \min(F_X(x), F_Y(y)) \quad (3.6)$$

para todo $x, y \geq 0$.

Note que, para que tal condição seja satisfeita, é necessário que ambas as funções de distribuição sejam perfeitamente correlacionadas. Assim, a comonotonicidade estabelece o limite superior da função de distribuição conjunta de dois riscos.

Definição 3.5: *Dois riscos, X e Y , são ditos countermonotônicos se a função de distribuição conjunta deles satisfaz*

$$F_{X,Y}(x, y) = \max(F_X(x) + F_Y(y) - 1, 0) \quad (3.6)$$

para todo $x, y \geq 0$.

Teorema 3.1 (Limites de Fréchet): *A função de distribuição conjunta $F_{X,Y}$ dos riscos X e Y é limitada acima e abaixo por*

$$\max(F_X(x) + F_Y(y) - 1, 0) \leq F_{X,Y}(x, y) \leq \min(F_X(x), F_Y(y)) \quad (3.7)$$

Mais tarde veremos que tais limites também se empregam às funções cópula, já que tais funções são funções de distribuição conjunta. Como tais limites estão intimamente ligados às noções de correlação, o próximo teorema calha em estabelecer essas relações e as propriedades da correlação.

Teorema 3.2 (Hoeffding): *Seja (X, Y) um vetor aleatório com marginais F_1 e F_2 e estrutura de dependência não especificada; assuma que $\text{Var}(X), \text{Var}(Y) < \infty$. Então*

- (a) *O conjunto de todas as possíveis correlações é o intervalo fechado $[\rho_{\min}, \rho_{\max}]$ e para as correlações extremas vale $\rho_{\min} < 0 < \rho_{\max}$.*
- (b) *A correlação extrema $\rho = \rho_{\min}$ é atingida se e somente se X e Y são countermonotônicas; e $\rho = \rho_{\max}$ é atingida se e somente se X e Y são comonotônicas.*
- (c) *$\rho_{\min} = -1$ se e somente se X e $-Y$ são do mesmo tipo; $\rho_{\max} = 1$ se e somente se X e Y são do mesmo tipo.*

4 Distribuições de Cauda Pesada

4.1 Introdução

A propagação do uso de distribuições de cauda pesada por Mandelbrot(1963) e Fama(1965) para modelar dados financeiros, devido a evidências de melhores resultados do que as distribuições gaussianas utilizadas até então, fez com que vários estudos fossem feitos sobre modelagem de dados financeiros através de distribuições de cauda pesada.

A aplicação da Teoria de Valores Extremos, cujas distribuições integram as de cauda pesada, ao Risco Operacional surge do interesse no estudo de riscos extremos no mercado financeiro. Apesar de riscos extremos serem incomuns, resultados estatísticos significativos poderão ser obtidos se métodos de inferência apropriados forem aplicados. Dizemos que riscos extremos são incomuns porque se espera que a quantidade de eventos extremos em uma distribuição não seja grande.

Os eventos de RO que compõem o banco de dados podem não ser frequentes, pois provêm de fraudes ou falhas e ambas são evitadas ao máximo por uma instituição financeira, mas podem ser significativos no seu impacto. E são justamente esses eventos de impacto significativo que devem ser previstos. A TVE pode fornecer estimativas confiáveis das chances de eventos que nunca foram vistos antes, mas que podem afetar de forma considerável as finanças de uma instituição.

As distribuições que integram a TVE se referem à classe de distribuições de variação regular, que, por sua vez, constituem parte das distribuições de cauda pesada. Neste capítulo, conceituam-se tais distribuições e estudamos mais profundamente duas classes de distribuição de cauda pesada, sendo estas as distribuições de Variação Regular e as α -Estáveis.

Nesta seção, as principais referências são: Embrechts *et al.* (1997) e Cruz (2002).

4.2 Fundamentação Teórica

Uma função de distribuição acumulada, $F(x) = P(X \leq x)$, possui cauda leve ou pesada, dependendo da velocidade do decaimento a zero de sua cauda quando comparada com a cauda de funções de distribuição da família exponencial.

Definição 4.1: Uma função de distribuição acumulada é dita de cauda pesada, à direita, se para todo $\varepsilon > 0$,

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(x)}{e^{-\varepsilon x}} = \infty \quad (4.1)$$

onde $\bar{F}(x) = 1 - F(x)$.

Duas classes de distribuições de cauda pesada são: as α -estáveis e as distribuições de Variação Regular, conceituadas a seguir.

4.2.1 Distribuições α -estáveis

Definição 4.2: Dados Z_1, \dots, Z_n , variáveis aleatórias independentes e identicamente distribuídas, com distribuição comum Z , a variável aleatória Z é chamada de α -estável se existem constantes $D_n \in \mathbb{R}$ tais que $S_n = Z_1 + \dots + Z_n$ seja igual em distribuição a $n^{1/2}Z + D_n$.

As α -estáveis não tem função de distribuição fechada, porém a função característica é bem conhecida.

Definição 4.3: Uma distribuição Z , α -estável, tem função característica

$$\varphi(t) = \begin{cases} \exp \left\{ i\mu t - \sigma^\alpha |t|^\alpha \left[1 - i\beta \operatorname{sen}(t) \tan \left(\alpha \frac{\pi}{2} \right) \right] \right\}, & \text{se } \alpha \neq 1 \\ \exp \left\{ i\mu t - \sigma |t| \left[1 - i\beta \frac{2}{\pi} \operatorname{sen}(t) \ln(t) \right] \right\}, & \text{se } \alpha = 1 \end{cases} \quad (4.2)$$

onde $0 < \alpha \leq 2$ (índice de estabilidade), $-1 \leq \beta \leq 1$ (parâmetro de simetria), $\sigma > 0$ (parâmetro de escala) e $\mu \in \mathbb{R}$ (parâmetro de locação) caracterizam a distribuição Z e são únicos.

Definição 4.4: A função de distribuição de Z , pertence ao domínio de atração de uma distribuição α -estável Z_α se existirem constantes $a_n \in \mathbb{R}$ e $b_n > 0$ tais que

$$\frac{s_n - a_n}{b_n} \xrightarrow{d} Z_\alpha$$

Escrevemos, $Z \in D(Z_\alpha)$ e dizemos que $\{Z_n\}$ satisfaz o Teorema Central do Limite com limite Z_α .

Exemplos de distribuições α -estáveis são: a Cauchy Padrão, a Gaussiana Padrão e a Lévy com $\alpha = 1/2$. Dentre elas, a que possui cauda mais pesada é a Lévy $1/2$, já que tanto a Cauchy padrão quanto a Gaussiana padrão têm $\alpha = 1$ e quanto menor for o parâmetro α mais pesada é a cauda da distribuição.

4.2.2 Teoria do Valor Extremo

A literatura informa que um banco de dados de perdas operacionais não apresentará uma distribuição gaussiana. Isso porque tal banco terá vários eventos pequenos e poucos eventos muito grandes. A aplicação da Teoria do Valor Extremo, como sendo a teoria que sustenta esse tipo de distribuição, é conhecida, mas ainda está num estágio embrionário na gestão de riscos.

O principal interesse na TVE é estimar o *parâmetro caudal* ξ que caracteriza a forma da cauda da distribuição. Para tanto, existem, essencialmente, duas abordagens paramétricas para se proceder. A primeira, chamada *abordagem extrema*, estuda a distribuição seguida pelo máximo das perdas dado um horizonte de tempo, também conhecida como *Block Maxima*. Tal distribuição é mais bem ajustada pela Distribuição de Valor Extremo Generalizado (GEV, do inglês *generalized extreme value*), que pode ser estimada e que estima o parâmetro caudal.

A segunda, chamada *abordagem caudal*, ou picos acima de um limite (POT, do inglês *peaks-over-threshold*), considera apenas os valores acima de um limite estabelecido e, portanto, foca diretamente na distribuição da cauda. Tal distribuição é melhor ajustada pela Distribuição Generalizada de Pareto (GPD, do inglês *generalized Pareto distribution*), que pode ser estimada para providenciar uma estimativa do parâmetro caudal. Para propósitos de gestão de riscos, a segunda abordagem poderá ser melhor para vender para gerentes de unidades de negócio. Caso contrário, é mais difícil justificar que capital está sendo alocado a uma unidade para um evento que ocorreu há muitos anos. Vejamos de forma mais detalhada ambas as distribuições.

4.2.2.1 Distribuição Generalizada de Valores Extremos

Suponha que a variável aleatória X denote as perdas operacionais no banco de dados fornecido pelo banco. Sejam X_1, \dots, X_n as perdas monetárias observadas num determinado período. Sendo

$$Y_n = \max(X_1, \dots, X_n) \quad (4.3)$$

Para encontrar uma distribuição limitadora que não seja degenerada, a variável aleatória máxima Y_n é padronizada por parâmetros de localização, escala e forma, escolhidos para dar uma distribuição apropriada de extremos padronizados. Portanto, nos concentramos no comportamento assintótico dos extremos. A função de distribuição GEV é obtida do teorema de Fisher-Tippett dado a seguir.

Teorema 4.1: *Seja $\{X_n\}$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas, com distribuição comum F . Se existirem constantes normalizadoras $c_n > 0$ e $d_n \in \mathbb{R}$ e alguma função não-degenerada H tais que*

$$\lim_{n \rightarrow \infty} P\left(\frac{\max\{X_1, \dots, X_n\} - d_n}{c_n} \leq x\right) = H(x) \quad (4.4)$$

então H é um dos três seguintes tipos de distribuição

Fréchet

$$\Phi_{\alpha,\mu,\sigma}(x) = \begin{cases} \exp[-(\frac{x-\mu}{\sigma})^{-\alpha}], & x > \mu \\ 0, & x < \mu \end{cases} \quad (4.5)$$

Weibull

$$\Psi_{\alpha,\mu,\sigma}(x) = \begin{cases} \exp[-(-\frac{x-\mu}{\sigma})^\alpha], & x > \mu \\ 0, & x < \mu \end{cases} \quad (4.6)$$

Gumbel

$$\Lambda_{\mu,\sigma}(x) = \exp\left(-e^{-\frac{x-\mu}{\sigma}}\right), \quad x \in \mathbb{R} \quad (4.7)$$

sendo que $\alpha > 0$ é o índice de estabilidade, μ é o parâmetro de locação e $\sigma > 0$ é o parâmetro de escala. F_X é dita pertencer ao domínio de atração de H . As três distribuições são chamadas de distribuição padrão de valores extremos.

Uma forma fechada para essas três distribuições pode ser muito útil em várias situações. A representação da GEV, unificando esses três tipos de distribuições em uma única família, foi estabelecida apenas em 1955 por Jenkinson – Von Mises.

Definição 4.5: Define-se a função de distribuição $H_{\xi,\mu,\sigma}$ por

$$H_{\xi,\mu,\sigma} = \begin{cases} \exp\left\{-\left(1 + \xi \frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right\}, & \xi \neq 0 \\ \exp\left\{-e^{-\frac{x-\mu}{\sigma}}\right\}, & \xi = 0 \end{cases} \quad (4.8)$$

onde $\xi \in \mathbb{R}$ é chamado de índice caudal, $\mu \in \mathbb{R}$ é o parâmetro de locação e $\sigma > 0$ é o parâmetro de escala, e $1 + \xi(x - \mu)/\sigma > 0$. H_ξ é padronizada quando $\mu = 0$ e $\sigma = 1$. Assim como a GPD, a distribuição GEV tem cauda pesada quando $\xi > 1/2$, ou seja, no caso particular da Fréchet. A relação entre as distribuições Fréchet, Weibull e Gumbel com a GEV é dada da seguinte forma: se $\xi > 0$, Fréchet; se $\xi = 0$, Gumbel; e se $\xi < 0$, Weibull.

4.2.2.2 Distribuição Generalizada de Pareto

A GPD foi apresentada por Pickands (1975) e estudada por Davison (1983) e Hosking & Wallis (1987) entre outros. Ela costuma ser usada na modelagem de grandes sinistros de seguro e em estudos de confiabilidade. A característica notável da distribuição de Pareto é que o parâmetro de escala é função de certo limite que será determinado pelo analista.

Consideremos uma amostra de observações independente e identicamente distribuídas $\{X_n\}$ com distribuição F_X . Estamos interessados na distribuição que seguem as observações que excedem um determinado limite. Seja x_F o ponto final a direita de F .

Definição 4.6: *Seja $u < x_F$, um número real fixo, o limite. A função*

$$F_u(y) = P(X_t - u \leq y \mid X_t > u) = \frac{F_X(y + u) - F_X(u)}{1 - F_X(u)} \quad (4.9)$$

onde $0 \leq y < x_F - u$, é chamada de função de distribuição do excesso (edf, do *inglês excess distribution function*) das v.a.'s X_n acima do limite u . A função

$$e(u) = E[X_t - u \mid X_t > u] \quad (4.10)$$

é chamada de função do excesso médio (mef, do inglês *mean-excess function*).

A edf mensura a probabilidade de uma realização acima do limite u estar próximo de um determinado valor y . A mef capta aquelas realizações que excedem u e considera a distância entre a média e u .

Teorema 4.2: *Se F_X está no domínio de atração da distribuição de valor extremo H_ξ a edf $F_u(y)$ pode ser aproximada, para u "grande", por*

$$F_u(y) \sim G_{\xi, \mu, \sigma}$$

onde $\mu = u \rightarrow \infty$ e

$$G_{\xi, \mu, \sigma} = \begin{cases} 1 - \left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-\frac{1}{\xi}}, & \xi \neq 0 \\ 1 - e^{-\frac{x - \mu}{\sigma}}, & \xi = 0 \end{cases} \quad (4.11)$$

para $0 \leq y < x_F - u$, onde $\xi \in \mathbb{R}$ é chamado de índice caudal, $\mu \in \mathbb{R}$ é o parâmetro de locação e $\sigma > 0$ é o parâmetro de escala. $G_{\xi, \mu, \sigma}$ é chamada de Distribuição Generalizada de Pareto.

Pode-se apresentar a família G_ξ como Distribuição Generalizada de Pareto Padronizada quando $\mu = 0$ e $\sigma = 1$. Além disso, quando $\xi > 1/2$, então $G_{\xi,\mu,\sigma}$ é de cauda pesada. O suporte para y é $y \geq 0$ se $\xi \geq 0$, e $0 \leq y \leq -1/\xi$ se $\xi < 0$. Tanto a GEV quanto a GPD pertencem à classe de distribuições de cauda pesada conhecidas com distribuições de variação regular.

4.2.3 Distribuições de Variação Regular

Definição 4.7: A cauda de uma distribuição F , em que $\bar{F}(x) = P(X > x)$, é de variação regular no infinito se

$$\lim_{t \rightarrow \infty} \frac{\bar{F}(tx)}{\bar{F}(t)} = x^{-\alpha} \quad (4.12)$$

onde $\alpha \geq 0$ e $x > 0$. Se $\alpha = 0$, dizemos que F é lentamente variante.

Apresentadas as distribuições a que, provavelmente, melhor se ajustam às distribuições das variáveis aleatórias de perdas sofridas por uma instituição financeira, devemos apresentar a teoria de cópulas de valores extremos. As cópulas são funções de distribuição conjunta, mas que nos possibilitam definir a estrutura de correlação entre suas funções de distribuição marginal, um dos objetivos deste trabalho. Assim, vejamos a teoria geral do estudo de cópulas, bem como suas famílias, e depois nos restringiremos à teoria de cópulas de valores extremos, bem como suas famílias.

5 Cópulas

5.1 Introdução

De maneira intuitiva, podemos dizer que uma cópula é uma função de distribuição multivariada definida no cubo $[0,1]^n$ e que possui marginais uniformemente distribuídas. Deriva-se das cópulas a estrutura de dependência existente entre as marginais, sendo, assim, de grande interesse para o objetivo pretendido neste trabalho. As funções cópulas são definidas como se segue. Uma classe de cópulas, chamada de cópulas de valores extremos, é estudada mais especificamente do que as demais classes devido à natureza dos dados que se pretende modelar.

Embrechts et al. (2002) e Nelsen (2006) são as principais fontes de informação para a construção dessa seção.

5.2 Fundamentação Teórica

Seja H uma função; denotaremos por $\text{Dom}(H)$ e $\text{Ran}(H)$ o domínio e a amplitude, respectivamente, de H . Também estabeleceremos que uma função f será dita crescente quando $x \leq y$ implicar em $f(x) \leq f(y)$.

Definição 5.1: *Seja S_1, \dots, S_n conjuntos não vazios de \mathbb{R} , onde \mathbb{R} denota a linha real estendida $[-\infty, \infty]$. Seja H uma função real com n variáveis tais que $\text{Dom}(H) = S_1 \times \dots \times S_n$ e para $a \leq b$ ($a_k \leq b_k$, para todo k) seja*

$B = [a, b] = ([a_1, b_1] \times \dots \times [a_n, b_n])$ uma n -caixa cujos vértices pertencem ao $\text{Dom}(H)$. Então o H -volume de B é a n -ésima diferença de ordem de H em B

$$V_H(B) = \Delta_a^b H(t) = \Delta_{a_n}^{b_n} \dots \Delta_{a_1}^{b_1} H(t) \quad (5.1)$$

onde as primeiras n diferenças de ordem são definidas como

$$\Delta_{a_k}^{b_k} H(t) = H(t_1, \dots, t_{k-1}, b_k, t_{k+1}, \dots, t_n) - H(t_1, \dots, t_{k-1}, a_k, t_{k+1}, \dots, t_n) \quad (5.2)$$

Definição 5.2: Uma função real H de n variáveis é n -crescente se $V_H(B) \geq 0$ para toda n -caixa B cujos vértices pertencem ao $\text{Dom}(H)$.

Suponha que cada S_k tenha um menor elemento a_k . Dizemos que H é não degenerada se $H(t) = 0$ para todo t em $\text{Dom}(H)$ tal que $t_k = a_k$ para pelo menos um k . Se cada S_k é não vazio e tem maior elemento b_k , então H tem marginais e as marginais de dimensão 1 de H são as funções H_k com $\text{Dom}(H_k) = S_k$ e com $H_k(x) = H(b_1, \dots, b_{k-1}, x, b_{k+1}, \dots, b_n)$ para todo x em S_k .

Definição 5.3: Uma cópula n -dimensional é uma função C com domínio em $[0,1]^n$ tal que

- i) C é não degenerada e n -crescente
- ii) C tem marginais C_k , $k=1, \dots, n$, que satisfazem $C_k(u) = u$ para todo u em $[0,1]$

Observe que para qualquer n -cópula C , $n \geq 3$, cada marginal k -dimensional de C é uma k -cópula. Equivalentemente uma n -cópula é uma função C de $[0,1]^n$ para $[0,1]$ com as seguintes propriedades:

- a. Para todo \mathbf{u} em $[0,1]^n$, $C(\mathbf{u}) = 0$ se ao menos uma coordenada de \mathbf{u} for 0, e $C(\mathbf{u}) = u_k$ se todas as coordenadas de \mathbf{u} forem iguais a 1 exceto u_k .
- b. Para todo a e b em $[0,1]^n$ tal que $a_i \leq b_i$ para todo i , $V_C([a, b]) \geq 0$.

Portanto, cópulas são funções de distribuição conjunta (em $[0,1]^n$), uma cópula C indica uma medida de probabilidade em $[0,1]^n$ através de

$$V_C([0, u_1] \times \dots \times [0, u_n]) = C(u_1, \dots, u_n) \quad (5.3)$$

e uma extensão padrão para subconjuntos Borel arbitrários de $[0,1]^n$ (não necessariamente n -caixas).

Teorema 5.1 (Teorema de Sklar): Seja H uma função de distribuição n -dimensional com marginais F_1, \dots, F_n . Então existe uma n -cópula C tal que, para todo x em \mathbb{R}^n ,

$$H(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)) \quad (5.4)$$

Se F_1, \dots, F_n forem todas contínuas, então C é única; caso contrário, C é unicamente determinada em $\text{Ran}(F_1) \times \dots \times \text{Ran}(F_n)$.

A partir do Teorema de Sklar podemos ver que para funções de distribuição multivariadas contínuas, as marginais univariadas e a estrutura de dependência multivariada podem ser separadas, e a estrutura de dependência pode ser representada por uma cópula. Sendo F uma função de distribuição univariada, define-se sua inversa como $F^{-1}(t) = \inf\{x \in R | F(x) \geq t\}$ para todo t em $[0,1]$.

Corolário 5.1: *Seja H uma função de distribuição n -dimensional com marginais F_1, \dots, F_n e função cópula C . Então, para qualquer \mathbf{u} em $[0,1]^n$,*

$$C(u_1, \dots, u_n) = H(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)) \quad (5.5)$$

Devemos tomar cuidado ao trabalhar sem o pressuposto de continuidade.

5.3 Limites de Fréchet-Hoeffding

Considere as funções M^n, Π^n, W^n definidas em $[0,1]^n$ como segue

$$M^n(\mathbf{u}) = \min(u_1, \dots, u_n) \quad (5.6)$$

$$\Pi^n(\mathbf{u}) = u_1 \dots u_n \quad (5.7)$$

$$W^n(\mathbf{u}) = \max(u_1 + \dots + u_n - n + 1, 0) \quad (5.8)$$

As funções M^n e Π^n são n -cópulas para todo $n \geq 2$, enquanto a função W^n não é uma cópula para qualquer $n \geq 3$. Embora W^n nunca seja uma cópula para $n \geq 3$, é o melhor limite inferior possível.

Teorema 5.2: *Seja C uma n -cópula qualquer, então, para todo \mathbf{u} em $[0,1]^n$*

$$W^n(\mathbf{u}) \leq C(\mathbf{u}) \leq M^n(\mathbf{u}) \quad (5.9)$$

Este teorema é conhecido como Teorema dos Limites de Fréchet-Hoeffding.

Sejam X_1, \dots, X_n variáveis aleatórias com funções de distribuição contínuas F_1, \dots, F_n , respectivamente, e função de distribuição conjunta H . Então $(X_1, \dots, X_n)^T$ tem cópula única C , onde C é dada pelo Teorema de Sklar. A representação da cópula padrão da distribuição do vetor aleatório $(X_1, \dots, X_n)^T$ é dada por

$$H(x_1, \dots, x_n) = P\{X_1 \leq x_1, \dots, X_n \leq x_n\} = C(F_1(x_1), \dots, F_n(x_n)) \quad (5.10)$$

As transformações $X_i \rightarrow F_i(X_i)$, usadas na representação acima, são normalmente referidas como transformações de probabilidade-integral (para uniformidade) e formam uma ferramenta padrão na metodologia de simulação. Portanto, X_1, \dots, X_n são independentes se e somente se $H(x_1, \dots, x_n) = F_1(x_1) \dots F_n(x_n)$ para todo x_1, \dots, x_n em \mathbb{R} , o seguinte resultado segue do Teorema de Sklar.

Teorema 5.3: *Seja $(X_1, \dots, X_n)^T$ um vetor de variáveis aleatórias contínuas com cópula C , então X_1, \dots, X_n são independentes se e somente se $C = \Pi^n$.*

5.4 Dependência Caudal

A dependência caudal é uma medida de dependência importante e que pode ser derivada das cópulas. Ela capta o fenômeno de simultâneos valores extremos em várias variáveis de risco, o que é um dos maiores interesses no gerenciamento de riscos. Quando modelamos a dependência de eventos caudais, temos que levar em consideração a estrutura de dependência caudal. Devemos encontrar características caudais relevantes das perdas com cópulas específicas que mostrem diferentes tipos de dependência caudal.

Definição 5.4: *Seja $(X, Y)^T$ um vetor de variáveis aleatórias contínuas com funções de distribuição marginal $H_X(x)$ e $H_Y(y)$. O coeficiente de dependência caudal superior de $(X, Y)^T$ é*

$$\lambda_u = \lim_{\gamma \rightarrow 1} P(X > H_X^{-1}(\gamma) | Y > H_Y^{-1}(\gamma)) = \frac{\lim_{\gamma \rightarrow 1} [1 - 2\gamma + C_0(\gamma, \gamma)]}{1 - \gamma} \quad (5.11)$$

sendo que o limite λ_u pertencente a $[0,1]$ existe. Se λ_u pertence a $(0,1]$, X e Y são ditos assintoticamente dependentes na cauda superior; se $\lambda_u = 0$, X e Y são ditos assintoticamente independentes na cauda superior.

De forma semelhante podemos estabelecer o coeficiente de dependência caudal inferior do vetor $(X, Y)^T$

$$\lambda_l = \lim_{\gamma \rightarrow 0} P(X < H_X^{-1}(\gamma) | Y < H_Y^{-1}(\gamma)) = \frac{\lim_{\gamma \rightarrow 0} C_0(\gamma, \gamma)}{\gamma} \quad (5.12)$$

Devido à natureza das variáveis de perdas operacionais, o coeficiente de dependência caudal inferior se torna menos importante; por isso, apenas o citamos. Assim, pode-se calcular a probabilidade condicional de uma perda que exceda o seu VaR quando outra perda exceda o VaR correspondente sob diferentes níveis de confiança γ . Isso é dado por

$$P(X > VaR_X(\gamma) | Y > VaR_Y(\gamma)) = \frac{1 - 2\gamma + C_0(\gamma, \gamma)}{1 - \gamma} \quad (5.13)$$

que é o estimador da dependência caudal superior quando γ se aproxima de 1.

Exemplo 5. 1: Considere a família de cópulas de Gumbel bivariada dada por

$$C_\theta(u, v) = \exp(-[(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta}) \quad (5.14)$$

para $\theta \geq 1$. Então

$$\frac{1 - 2u + C(u, u)}{1 - u} = \frac{1 - 2u + \exp(2^{1/\theta} \ln u)}{1 - u} = \frac{1 - 2u + u^{2^{1/\theta}}}{1 - u} \quad (5.15)$$

e segue que

$$\lim_{u \nearrow 1} \frac{1 - 2u + C(u, u)}{1 - u} = 2 - \lim_{u \nearrow 1} 2^{1/\theta} u^{2^{1/\theta}-1} = 2 - 2^{1/\theta} \quad (5.16)$$

Então, para $\theta > 1$, C_θ tem dependência caudal superior.

Devemos voltar nossa atenção para as cópulas de valores extremos, das quais suas marginais são modeladas por distribuições de cauda pesada, tais como as cópulas de Gumbel, exemplificada acima. Mas, antes explanemos as famílias de cópulas mais conhecidas.

5.5 Famílias de Cópulas

5.5.1 Introdução

As duas principais famílias de cópulas são as Cópulas Elípticas e as Cópulas Arquimedianas. Neste trabalho estamos interessados no estudo de conjuntos de cópulas que estão contidos nessas famílias. Tais cópulas são as cópulas de valores extremos estudadas na próxima seção. Apesar desse conjunto ser restrito, pois existe apenas uma cópula arquimediana de valor extremo, tais cópulas são as que melhor se adequam ao tipo de dados estudados segundo a literatura existente. Outro exemplo de cópulas são as de Marshal-Olkin, citadas apenas em termos ilustrativos.

5.5.2 Cópulas Elípticas

As classes de distribuições elípticas proporcionam uma ampla gama de distribuições multivariadas que dividem várias propriedades com a distribuição gaussiana multivariada, não permitindo a plena modelagem de extremos multivariados e outras formas de dependências não normais. Cópulas elípticas nada mais são que cópulas de distribuições elípticas. A correlação por postos e o coeficiente de dependência caudal são facilmente calculados.

5.5.2.1 Distribuições Elípticas

Definição 5.5: Se \mathbf{X} é um vetor aleatório n -dimensional e, para algum $\boldsymbol{\mu} \in \mathbb{R}^n$ e para alguma matriz simétrica $n \times n$ não-negativa $\boldsymbol{\Sigma}$, a função característica $\varphi_{\mathbf{X}-\boldsymbol{\mu}}(t)$ de $\mathbf{X}-\boldsymbol{\mu}$ é uma função da forma quadrática $t^T \boldsymbol{\Sigma} t$, $\varphi_{\mathbf{X}-\boldsymbol{\mu}}(t) = \phi(t^T \boldsymbol{\Sigma} t)$, dizemos que \mathbf{X} tem distribuição elíptica com parâmetros $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ e ϕ , e escrevemos $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$.

Quando $n = 1$, a classe de distribuições elípticas coincide com a classe de distribuições simétricas unidimensionais. A função ϕ é chamada função característica geradora. A cópula das distribuições elípticas de X é conhecida como cópula elíptica. A classe das distribuições elípticas define a classe da cópula elíptica, onde duas são as mais conhecidas: Cópulas Gaussianas e t-Cópulas.

5.5.2.2 Cópulas Gaussianas

Definição 5.6: A cópula da distribuição normal n -variada com matriz de correlação linear R é

$$C_R^{Ga}(\mathbf{u}) = \Phi_R^n(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n)) \quad (5.17)$$

onde Φ_R^n denota a função de distribuição conjunta da função de distribuição normal padrão n -variada com matriz de correlação linear R , e Φ^{-1} denota a inversa da função de distribuição normal padrão univariada. Cópulas da forma acima são chamadas de cópulas gaussianas.

5.5.2.3 t-Cópuas

Definição 5.7: Se \mathbf{X} tem a representação estocástica

$$\mathbf{X} = \boldsymbol{\mu} + \frac{\sqrt{v}}{\sqrt{S}} \mathbf{Z} \quad (5.18)$$

onde $\boldsymbol{\mu} \in \mathbb{R}^n$, $S \sim \chi_v^2$ e $Z \sim N_n(\mathbf{0}, \boldsymbol{\Sigma})$ são independentes, então \mathbf{X} tem uma distribuição *t* de Student *n*-variada com *v* graus de liberdade com média $\boldsymbol{\mu}$ (para $v > 1$) e matriz de covariância $(v/v-2)\boldsymbol{\Sigma}$ (para $v > 2$). Se $v \leq 2$, $\text{Cov}(X)$ não é definida. Nesse caso, apenas interpretamos $\boldsymbol{\Sigma}$ como sendo o parâmetro de forma da distribuição de \mathbf{X} .

A cópula do \mathbf{X} dado acima pode ser escrita como

$$C_{v,R}^d(\mathbf{u}) = t_{v,R}^n(t_v^{-1}(u_1), \dots, t_v^{-1}(u_n)) \quad (5.19)$$

onde $R_{ij} = \boldsymbol{\Sigma}_{ij}/\sqrt{\boldsymbol{\Sigma}_{ii}\boldsymbol{\Sigma}_{jj}}$ para $i, j \in \{1, \dots, n\}$ e $t_{v,R}^n$ denota a função de distribuição de $\mathbf{Y}\sqrt{v}/\sqrt{S}$, onde $S \sim \chi_v^2$ e $\mathbf{Y} \sim N_n(0, \mathbf{R})$ são independentes.

5.5.3 Cópuas Arquimedianas

Várias famílias de cópuas paramétricas interessantes são Arquimedianas e as classes de cópuas arquimedianas nos direcionam para uma grande variedade de diferentes estruturas de dependência. Ao contrário das cópuas elípticas, as arquimedianas têm expressão fechada e não derivam de funções de distribuição multivariadas usando o Teorema de Sklar.

Definição 5.8: Seja φ uma função contínua e estritamente decrescente de $[0,1]$ para $[0, \infty]$ tal que $\varphi(1) = 0$. A pseudo-inversa de φ é a função $\varphi^{[-1]}: [0, \infty] \rightarrow [0,1]$ dada por

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), & 0 \leq t \leq \varphi(0), \\ 0, & \varphi(0) \leq t \leq \infty \end{cases} \quad (5.20)$$

Note que $\varphi^{[-1]}$ é contínua e decrescente em $[0, \infty]$ e estritamente decrescente em $[0, \varphi(0)]$. No mais, $\varphi^{[-1]}(\varphi(u)) = u$ em $[0, 1]$, e

$$\varphi(\varphi^{[-1]}(t)) = \begin{cases} t, & 0 \leq t \leq \varphi(0), \\ \varphi(0), & \varphi(0) \leq t \leq \infty \end{cases} \quad (5.21)$$

Finalmente, se $\varphi(0) = \infty$ então, $\varphi^{[-1]} = \varphi^{-1}$.

Teorema 5.4: *Seja φ uma função contínua e estritamente decrescente de $[0, 1]$ para $[0, \infty]$ tal que $\varphi(1) = 0$, e seja $\varphi^{[-1]}$ a pseudo inversa de φ . Seja C a função de $[0, 1]^2$ para $[0, 1]$ dada por*

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)) \quad (5.22)$$

Então C é uma cópula se e somente se φ é convexa.

Tais cópulas são chamadas cópulas Arquimedianas. A função φ é chamada de função geradora da cópula. Se $\varphi(0) = \infty$, dizemos que φ é uma geradora estrita. Nesse caso, $\varphi^{[-1]} = \varphi^{-1}$ e $C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v))$ é dita ser uma cópula arquimediana estrita.

Exemplo 5.2: Seja $\varphi(t) = 1 - t$ para t em $[0, 1]$. Então $\varphi^{[-1]}(t) = 1 - t$ para t em $[0, 1]$ e o para $t > 1$; isto é, $\varphi^{[-1]}(t) = \max(1 - t, 0)$. Como $C(u, v) = \max(u + v - 1, 0) = W(u, v)$, nós vemos que o limite inferior de Fréchet-Hoeffding bivariado W é arquimediano.

5.5.3.1 Propriedades

Teorema 5.5: *Seja C uma cópula arquimediana com geradora φ . Então*

- 1) C é simétrica, ou seja, $C(u, v) = C(v, u)$ para todo u, v em $[0, 1]$.
- 2) C é associativa, ou seja, $C(u, v) = C(C(v, u), w) = C(u, C(v, w))$ para todo u, v, w em $[0, 1]$.

A propriedade associativa das cópulas arquimedianas, em geral, não é compartilhada pelas demais cópulas. Uma extensão das cópulas arquimedianas multivariadas são as cópulas de valores extremos, que serão estudadas profundamente, já que a literatura nos informa que

os dados de risco operacional se ajustam melhor à distribuições de cauda pesada, cujas distribuições multivariadas são cópulas de valores extremos.

6 Cópulas de Valores Extremos

6.1 Introdução

Uma vantagem das Cópulas de Valores Extremos em relação à classe de cópulas Arquimedianas muito mais populares é que elas não são simétricas. Assim, a característica assimétrica verificada em dados financeiros ganha relevância, o que produz estimativas mais fiéis. Uma classe híbrida contendo ambas as cópulas Arquimedianas e cópulas extremas como um caso especial são as Cópulas Archimax.

Esta seção foi elaborada a partir de informações obtidas em Cebrián *et al.*(2003), Gudendorf *et al.*(2009a), Lu *et al.*(2008) e Kotz (2000).

6.2 Fundamentação Teórica

Seja $X_i = (X_{i1}, \dots, X_{id})$, onde i pertence a $\{1, \dots, n\}$, uma amostra de vetores aleatórios independentes e identicamente distribuídos com mesma função de distribuição F , marginais F_1, \dots, F_d e cópula C_F . Por conveniência assuma F contínua. Considere o vetor de máximas componentes:

$$M_n = (M_{n,1}, \dots, M_{n,d}) \quad (6.1)$$

onde $M_{n,j} = \max(X_{i,j}, \dots, X_{n,j})$ e $j = 1, \dots, d$. A distribuição conjunta e marginais de M_n são dadas por F^n e F_1^n, \dots, F_d^n respectivamente. Isso significa que a cópula C_n de M_n é dada por

$$C_n(u_1, \dots, u_d) = C_F(u_1^{1/n}, \dots, u_d^{1/n})^n \quad (6.2)$$

em que $(u_1, \dots, u_d) \in [0,1]^d$.

Definição 6.1: Uma cópula C é chamada de cópula de valor extremo se existe uma cópula C_F tal que

$$C_F(u_1^{1/n}, \dots, u_d^{1/n}) \xrightarrow{n \rightarrow \infty} C(u_1, \dots, u_d) \quad (6.3)$$

para todo $(u_1, \dots, u_d) \in [0,1]^d$. A cópula C_F é dita ser do domínio de atração de uma Cópula de Valores Extremos e denotamos por $C_F \in D(C)$.

A seguir, daremos dois resultados, ambos necessários e suficientes para que $C_F \in D(C)$.

O primeiro resultado, dado por Marshall e Olkin (1983), expressa C_F como o limite da distribuição condicional de $\mathbf{u} = (u_1, \dots, u_d)$ dado que a última componente de \mathbf{u} tenha excedido t . Então, $C_F \in D(C)$ se e somente se

$$\frac{-\log C_F(tu_1, \dots, tu_d)}{-\log C_F(t, \dots, t)} \rightarrow \frac{-\log C(u_1, \dots, u_d)}{-\log C(1, \dots, 1)} \quad (6.4)$$

quando $t \rightarrow \infty$ para cada $u_j > 0$, $j = 1, \dots, d$.

O segundo resultado, apresentado por Takahashi(1994a), está em termos de

$$D_{C_F}(u_1, \dots, u_d) = C_F(\Phi_1^-(u_1), \dots, \Phi_1^-(u_d)) \quad (6.5)$$

e

$$D_C(u_1, \dots, u_d) = C(\Phi_1^-(u_1), \dots, \Phi_1^-(u_d)) \quad (6.6)$$

que são as cópulas de C_F e C , respectivamente. Então, $C_F \in D(C)$ se e somente se

$$\lim_{t \rightarrow \infty} t[1 - D_{C_F}(\mathbf{u}^{1/t})] = -\log D_C(\mathbf{u}) \quad (6.7)$$

para todo $\mathbf{u} \in (0,1)^d$.

Em particular, estamos interessados na possível distribuição multivariada limitante para M_n sob normalizações apropriadas. Consideremos uma aplicação bivariada, por motivo de simplificação. Sejam X e Y variáveis aleatórias e G a distribuição de (X,Y) com distribuições marginais H_X e H_Y . Assim, podemos encontrar a fd não-degenerada tal que a distribuição bivariada extrema G possa ser conectada a uma cópula de valores extremos C_0

$$G(x, y) = C_0(H_X(x; \mu_1, \sigma_1, \xi_1), H_Y(y; \mu_2, \sigma_2, \xi_2)) \quad (6.8)$$

onde μ_j, σ_j, ξ_j são parâmetros da GEV e H é a marginal da GEV.

De acordo com Joe (1997), a cópula única C_0 existe e satisfaz

$$C_0(u^t, v^t) = C_0^t(u, v) \quad (6.9)$$

sendo $t > 0$.

Teorema 6.1: *Uma cópula bivariada C é uma cópula de valor extremo se e somente se*

$$C_0(u, v) = \exp \left\{ \ln(uv) A \left[\frac{\ln(v)}{\ln(uv)} \right] \right\} \quad (6.10)$$

para $(u, v) \in (0, 1]^2$, onde A é chamada de função de dependência

$$A(t) = \int_0^1 \max [t(1-x), x(1-t)] dH(x) \quad (6.11)$$

onde $A: [0, 1] \rightarrow [1/2, 1]$ é convexa e satisfaz

$$\max[t, (1-t)] \leq A(t) \leq 1$$

onde $0 \leq t \leq 1$ e para todo $t \in [0, 1]$. Qualquer função que satisfaça essas duas restrições do teorema acima é uma Cópula de Valor Extremo.

O limite superior de A , $A(t) = 1$, corresponde à independência, ou seja, $C(u, v) = uv$, enquanto o limite inferior, $A(t) = \max[t, (1-t)]$ corresponde à dependência perfeita (comonotonicidade), ou seja, $C(u, v) = \min(u, v)$.

A seguir, três cópulas de valores extremas comumente utilizadas na prática são definidas, onde θ é um parâmetro de controle da dependência.

Cópula de Gumbel

$$C(u, v) = \exp \left\{ -[(-\ln(u))^\theta + (-\ln(v))^\theta]^{1/\theta} \right\} \quad (6.12)$$

onde $\theta \in [1, +\infty)$. Sua função de dependência é dada por

$$A(t) = [t^\theta + (1-t)^\theta]^{1/\theta} \quad (6.13)$$

Cópula de Galambos

$$C(u, v) = uv \exp \{ [(-\ln u)^{-\theta} + (-\ln v)^{-\theta}]^{-1/\theta} \} \quad (6.14)$$

e possui função de dependência

$$A(t) = 1 - [t^{-\theta} + (1-t)^{-\theta}]^{-1/\theta} \quad (6.15)$$

Cópula de Husler-Reiss

$$C(u, v) = \exp \left\{ -\tilde{u} \Phi \left[\frac{1}{\theta} + \frac{1}{2} \theta \ln \left(\frac{\tilde{u}}{\tilde{v}} \right) \right] - \tilde{v} \Phi \left[\frac{1}{\theta} + \frac{1}{2} \theta \ln \left(\frac{\tilde{v}}{\tilde{u}} \right) \right] \right\} \quad (6.16)$$

onde $\tilde{u} = -\ln u$, $\tilde{v} = -\ln v$ e Φ é a distribuição normal padrão. Sua função de dependência é dada por

$$A(t) = t \Phi \left[\theta^{-1} + \frac{1}{2} \theta \ln \left(\frac{t}{1-t} \right) \right] + (1-t) \Phi \left[\theta^{-1} - \frac{1}{2} \theta \ln \left(\frac{t}{1-t} \right) \right] \quad (6.17)$$

Para conseguirmos a estimativa de $A(t)$, precisamos estimar o parâmetro θ . Existem alguns métodos alternativos para obter a estimação de $A(t)$ e serão mostrados no próximo capítulo. Agora, devemos voltar nossa atenção para a estimação da cauda conjunta.

6.3 Estimação da Cauda Conjunta

Podemos conseguir a estimação da cauda conjunta correspondendo a um valor em risco sob um nível de confiança γ a partir de

$$G_\gamma = G(VaR_X(\gamma), VaR_Y(\gamma)) = C_0 \left(H_X(VaR_X(\gamma)), H_Y(VaR_Y(\gamma)) \right) = C_0(\gamma, \gamma) \quad (6.18)$$

O VaR , para cada marginal, pode ser calculado segundo a expressão

$$VaR(\gamma) = \mu + [\sigma(-\ln \gamma)^{-\xi} / \xi - 1] \quad (6.19)$$

quando as marginais são modeladas pela GEV. Da identidade

$$P(X > x, Y > y) = 1 - H_X(x) - H_Y(y) + G(x, y) \quad (6.20)$$

definida como função conjunta de sobrevivência, podemos estimar a cauda conjunta excedente para as duas perdas.

6.4 Dependência Caudal

Toda a teoria empregada à dependência caudal para a teoria de cópulas se aplica às cópulas de valores extremos. Ao mesmo tempo podemos obter a relação entre λ_u e $A(t)$:

$$\lambda_u = 2 - 2A(1/2) \quad (6.21)$$

Assim, devemos estimar primeiramente $A(1/2)$. Então fazemos a comparação entre as saídas baseadas nas diferentes cópulas de valores extremos para julgar sua validade e quantificar a dependência da cauda conjunta dos dados empíricos.

A partir das três cópulas de valor extremos estudadas neste capítulo, podemos calcular a probabilidade condicional de uma variável exceder seu VaR quando a outra excedeu seu respectivo VaR a um nível de confiança γ . Tal cálculo é dado por

$$P\left(X > VaR_\gamma(X) \mid Y > VaR_\gamma(Y)\right) = \frac{[1 - 2\gamma + C_0(\gamma, \gamma)]}{1 - \gamma} \quad (6.22)$$

que pode ser considerado o estimador da dependência caudal superior quando γ se aproxima de 1.

7 Estimadores

7.1 Introdução

Para a estimação dos parâmetros das cópulas, existem vários métodos comumente utilizados:

- 1) Método da Máxima Verossimilhança em Dois Passos: os parâmetros para as cópulas são estimados utilizando os parâmetros estimados para as marginais.
- 2) Funções de Inferência para as Marginais: os parâmetros para as marginais e para a cópula são estimados separadamente, no qual é um método seqüencial de dois passos onde se estima primeiramente os parâmetros das marginais e depois o parâmetro de dependência é estimado no segundo passo substituindo os parâmetros estimados na cópula.
- 3) Método da Máxima Verossimilhança Canônica: diferentemente dos dois métodos citados anteriormente, não é necessário estabelecer pressupostos em relação às marginais, apenas utiliza-se as distribuições empíricas de distribuições de perda e as coloca dentro de um procedimento de estimação dos parâmetros da cópula.
- 4) Abordagem Não-Paramétrica: os parâmetros podem ser estimados utilizando coeficiente de correlação por postos de Kendal; ou os estimadores: Estimador Original de Pickands, Estimador de Deheuvels, Estimador de Hall-Tajvidi ou, por último, o Estimador de Capéraà-Fougères-Genest, que segundo a literatura, obtém os estimadores mais consistentes.

Neste trabalho foi utilizado o Método da Máxima Verossimilhança em Dois Passos para estimar os parâmetros da cópula. Tais metodologias foram retiradas de Embrechts *et al.*(2002), Gudendorf *et al.*(2009b) e Kotz (2000).

7.2 Estimação Paramétrica

Nesta seção será descrita uma abordagem paramétrica para estimação dos parâmetros de uma cópula de valores extremos baseada na função de verossimilhança. A partir de (6.4), Ledford e Tawn (1996) e Smith *et al.* (1997) desenvolveram a seguinte abordagem:

$$\lim_{t \rightarrow \infty} \frac{-\log C_F(tu_1, \dots, tu_d)}{-\log C_F(t, \dots, t)} = \frac{-\log C(u_1, \dots, u_d)}{-\log C(1, \dots, 1)} = \frac{V(u_1, \dots, u_d)}{V(1, \dots, 1)} \quad (7.1)$$

Usando este resultado, eles assumem que (7.1) acontece como uma identidade para algum valor fixo e grande $t = t_c$. Como $t_c u_j$ precisa ser grande para cada j , $u'_j = t_c u_j$ está acima de um limite alto. Isto nos dá

$$\log C_F(u'_1, \dots, u'_d) = \log C_F(t_c, \dots, t_c) \frac{V(u'_1/t_c, \dots, u'_d/t_c)}{V(1, \dots, 1)} \quad (7.2)$$

Devido ao fato de que V é uma função homogênea de ordem -1 , e os termos u' serem variáveis *dummy*,

$$C_F(u_1, \dots, u_d) = \exp \left\{ V(u_1, \dots, u_d) \frac{t_c \log C_F(t_c, \dots, t_c)}{V(1, \dots, 1)} \right\} = \exp \{ V(u_1, \dots, u_d) K \} \quad (7.3)$$

para algum K , quando cada u_j está acima de um limite. Ledford e Tawn (1996) tomam esse limite como o quantil $1 - \lambda_j$ da distribuição Fréchet unitária, onde λ_j é alguma probabilidade pequena. Então, $u_j \geq -1/\log(1 - \lambda_j)$ para $j = 1, \dots, d$.

Para darmos valor à constante K , seja $u_1 = -1/\log(1 - \lambda_1)$, $u_2 = \dots = u_d = \infty$ dentro de (7.3) e note que

$$C_F(-1/\log(1 - \lambda_1), \infty, \dots, \infty) = \exp \{ V[-1/\log(1 - \lambda_1), \infty, \dots, \infty] K \}. \quad (7.4)$$

Isso implica que $1 - \lambda_1 = \exp[-K \log(1 - \lambda_1)]$, e, então, $K = -1$. Portanto, combinando (7.3) com a definição de C_F , temos

$$F(x_1, \dots, x_d) = \exp[-V\{-1/\log F_1(x_1), \dots, -1/\log F_d(x_d)\}], \quad (7.5)$$

válido para cada $x_j \geq t_j$, onde os t_j são escolhidos de tal forma que, para cada j , $F_j(t_j) = 1 - \lambda_j$. Assumindo que F_j para $X_j \geq t_j$ tem a distribuição Generalizada de Pareto com parâmetros $(\sigma_j, \xi_j, \lambda_j)$, temos

$$F(x_1, \dots, x_d) = \exp[-V\{Y_1(x_1), \dots, Y_d(x_d)\}], \quad (7.6)$$

válido para cada $x_j \geq t_j, j = 1, \dots, d$. Essa derivação simplesmente equivale a assumir que para excedentes conjuntos de um conjunto de limiares adequadamente altos, a estrutura de dependência da função de distribuição F é exatamente a de uma distribuição de valores extremos multivariada.

Para desenvolver uma verossimilhança baseada em (7.6), Ledford e Tawn (1996) consideram observações marginais abaixo de seus respectivos limiares. Desta maneira, a contribuição para a verossimilhança de um ponto típico (x_1, \dots, x_d) para os quais as marginais j_1, \dots, j_m alcançam ou excedem seus limites é dada por

$$\frac{\partial^m F(x_1, \dots, x_d)}{\partial x_{j_1} \dots \partial x_{j_m}} \quad (7.7)$$

com F dada por (7.6), avaliados em $x_1 = \max(t_1, X_1), \dots, x_d = \max(t_d, X_d)$.

A forma explícita da verossimilhança para o caso em que $d = 2$ é mostrada a seguir. Para os altos limites marginais t_1 e t_2 que definem o espaço $B_{i_1 i_2}$, onde $i_j = 0$ se $X_j < t_j$ e $i_j = 1$ se $X_j \geq t_j$, sendo $j = 1, 2$. Seja $L_{i_1 i_2}(x_1, x_2)$ a expressão da verossimilhança do ponto (x_1, x_2) que cai na região $B_{i_1 i_2}$, $r_j = -1/(1 - \lambda_j)$ e defina

$$\begin{aligned} V_1(u_1, u_2) &= -\frac{\partial}{\partial u_1} V(u_1, u_2), \\ V_2(u_1, u_2) &= -\frac{\partial}{\partial u_2} V(u_1, u_2), \\ V_{12}(u_1, u_2) &= -\frac{\partial^2}{\partial u_1 \partial u_2} V(u_1, u_2). \end{aligned}$$

A partir de (7.7) pode-se obter

$$\begin{aligned} L_{00}(x_1, x_2) &= \exp[-V(r_1, r_2)], \\ L_{01}(x_1, x_2) &= \exp[-V(r_1, Y_2(x_2))] V_2(r_1, Y_2(x_2)) K_2, \\ L_{10}(x_1, x_2) &= \exp[-V(Y_1(x_1), r_2)] V_1(Y_1(x_1), r_2) K_1, \\ L_{11}(x_1, x_2) &= \exp[-V(Y_1(x_1), Y_2(x_2))] \{ V_1(Y_1(x_1), Y_2(x_2)) V_2(Y_1(x_1), Y_2(x_2)) + \\ &\quad + V_{12}(Y_1(x_1), Y_2(x_2)) \} K_1 K_2 \end{aligned}$$

onde

$$K_j = \lambda_j \sigma_j^{-1} \{1 + \xi_j(x_j - t_j)/\sigma_j\}^{-(1+\xi_j^{-1})} + \{Y_j(x_j)\}^2 \exp\{1/Y_j(x_j)\} \quad (7.8)$$

para $j = 1, 2$. Assim, a contribuição de um ponto típico $(x_{i,1}, x_{i,2})$ na verossimilhança para os parâmetros de dependência θ e parâmetros marginais $\sigma = (\sigma_1, \sigma_2)$, $\xi = (\xi_1, \xi_2)$, $\mu = (\mu_1, \mu_2)$ é dada por

$$L_i(\theta, \sigma, \xi, \mu) = \sum_{r,s \in (0,1)} L_{rs}(x_{i,1}, x_{i,2}) I_{rs}(x_{i,1}, x_{i,2}), \quad (7.9)$$

onde $I_{rs}(x_{i,1}, x_{i,2}) = I\{(x_{i,1}, x_{i,2}) \in B_{rs}\}$ com I sendo uma função indicadora. Finalmente, a verossimilhança para um conjunto de n pontos independentes é dada por

$$L_{(n)}(\theta, \sigma, \xi, \mu) = \prod_{i=1}^n L_i(\theta, \sigma, \xi, \mu). \quad (7.10)$$

Assim, no caso em que $d = 2$, dado uma amostra de tamanho n , $(X_1, Y_1), \dots, (X_n, Y_n)$, a log-verossimilhança é dada por

$$\log L = \sum_{i=1}^n [\ln(f_1(x_i; \mu_1, \sigma_1, \xi_1)) + \ln(f_2(y_i; \mu_2, \sigma_2, \xi_2)) + \ln(C''(u, v), \theta)] \quad (7.11)$$

onde

$$C''(u, v) = \frac{\partial^2}{\partial u \partial v} C(u, v). \quad (7.12)$$

Quando as marginais são conhecidas ou estimadas, (7.11) precisa ser estimado somente com relação a

$$\log L = \sum_{i=1}^n [\ln(C''(u, v), \theta)] \quad (7.13)$$

Este método é conhecido como Estimação por Máxima Verossimilhança (MLE) em dois estágios e será adotado para estimar os parâmetros das cópulas neste trabalho.

7.3 Estimação Não Paramétrica

A estimação não paramétrica de distribuições multivariadas de valores extremos concerne na estimação da função de dependência A .

Seja $\{(U_{i,1}, U_{i,2})\}$, Pickands (1981) observou que $1/\max\{(U_{i,1}(1-w), U_{i,2}w)\}$ tem distribuição exponencial com média $1/A(w)$, para cada $w \in [0,1]$. Assim, a partir da amostra $\{(U_{i,1}, U_{i,2})\}, 1 \leq i \leq n$, um estimador consistente trivial para $A(w)$ é

$$A_n(w) = n \left[\sum_{i=1}^n 1/\max\{(U_{i,1}(1-w), U_{i,2}w)\} \right]^{-1}. \quad (7.14)$$

Sendo não convergente e não diferenciável, este estimador não satisfaz $A_n(0) = A_n(1) = 1$.

A simples modificação

$$\bar{A}_n(w) = n \left[\sum_{i=1}^n 1/\max\{(U_{i,1}(1-w), U_{i,2}w)\} - (1-w) \sum_{i=1}^n 1/U_{i,2} - w \sum_{i=1}^n 1/U_{i,1} + n \right]^{-1} \quad (7.15)$$

pode ser utilizada para superar esse defeito. Pickands propôs um método para fazer isso geralmente convergir substituindo A_n pelo seu “minorant” convergente. Mas sua construção é difícil e envolve uma função que contém uma dependência implícita complicada nos dados. Portanto, é natural que procuremos outros caminhos para estimar $A(\cdot)$.

Estimativas não paramétricas foram feitas por Caperaa *et al.* (1997) são baseadas na seguinte representação.

Seja (Y_1, Y_2) distribuído de acordo com (3.18). Seja $U_1 = G_{*1}(Y_1), U_2 = G_{*2}(Y_2)$ e $Z = \log(U_1)/\log(U_1 U_2)$. E a fd da variável aleatória Z é dada por

$$P(z) = P(Z \leq z) = z + z(1-z)D(z), \quad (7.16)$$

onde $D(z) = A'(z)/A(z)$ e $A'(z)$ denota a derivada de A para todo $0 \leq z < 1$. Uma prova dessa representação pode ser encontrada em Ghoudi *et al.* (1998). Como consequência dessa representação, pode-se obter

$$\frac{A(t)}{A(s)} = \exp \left\{ \int_s^t \frac{P(z) - z}{z(1-z)} dz \right\} \quad (7.17)$$

para escolhas arbitrárias de $0 \leq s \leq t \leq 1$. Como $A(0) = A(1) = 1$, pode-se escrever

$$A(t) = \exp \left\{ \int_0^t \frac{P(z) - z}{z(1-z)} dz \right\} = \exp \left\{ - \int_t^1 \frac{P(z) - z}{z(1-z)} dz \right\}. \quad (7.18)$$

Seja P_n a fd empírica de Z_1, \dots, Z_n para uma amostra aleatória $\{(Y_{i,1}, Y_{i,2}), i = 1, \dots, n\}$ vinda de (3.18). Substituindo P por P_n nas expressões acima, acontecem dois possíveis estimadores não paramétricos para A , denotados por

$$A_n^0(t) = \exp \left\{ \int_0^t \frac{P(z) - z}{z(1-z)} dz \right\}, \quad A_n^1(t) = \exp \left\{ - \int_t^1 \frac{P(z) - z}{z(1-z)} dz \right\}. \quad (7.19)$$

Caperaa et al. (1997) mostra que $\log A_n^0$ é um estimador não viesado e uniforme, fortemente consistente, de $\log A$, ou seja, $E(\log A_n^0) = \log A$ para todo $n \geq 1$ e

$$\sup_{t \in [0,1]} |\log A_n^0(t) - \log A(t)| \rightarrow 0$$

quase certamente. Adicionalmente, o processo $n^{1/2}(\log A_n^0 - \log A)$ é assintoticamente normal com média zero e matriz de covariância

$$\Gamma_0(s, t) = \int_0^s \int_0^t \frac{P(\min(u, v)) - P(u)P(v)}{uv(1-u)(1-v)} dvdu. \quad (7.20)$$

Como $P(\min(u, v)) \geq P(u)P(v)$, deve-se observar que $\Gamma_0(t, t)$ é monótona e crescente em t , assim, a respeito das suas propriedades de atração, $\log A_n^0$ é um estimador “increasing unreliable” de $\log A(t)$ quando $t \rightarrow 1$. Uma análise similar feita por Caperaa et al. (1997) mostra que $n^{1/2}(\log A_n^1 - \log A)$ é assintoticamente normal com média zero e matriz de covariância

$$\Gamma_1(s, t) = \int_s^1 \int_t^1 \frac{P(\min(u, v)) - P(u)P(v)}{uv(1-u)(1-v)} dvdu, \quad (7.21)$$

sendo a variância de $\log A_n^1$ é uma função decrescente de t . Isso sugere que um estimador combinado seja preferido para cada $\log A_n^i$, como segue:

$$\log A_n(t) = p(t) \log A_n^0(t) + \{1 - p(t)\} \log A_n^1(t), \quad (7.22)$$

onde $p(t)$ é uma função “bounded weight” em $[0,1]$ que proporciona, comparativamente, maior peso a $\log A_n^i$ na vizinhança de i . Novamente, Caperaa et al. (1997) mostra que $\log A_n^0$

é um estimador não viesado e uniforme, fortemente consistente, de $\log A$. Adicionalmente, o processo $n^{1/2}(\log A_n - \log A)$ é assintoticamente normal com média zero e função variância

$$\Gamma(t) = p^2(t)\Gamma_0(t, t) + \{1 - p(t)\}^2\Gamma_1(t, t) + 2p(t)\{1 - p(t)\}C(t), \quad (7.23)$$

onde

$$C(t) = - \int_0^t \int_t^1 \frac{P(u)\{1 - P(v)\}}{uv(1 - u)1 - v} dvdu \leq 0 \quad (7.24)$$

é a covariância assintótica de $n^{1/2} \log A_n^0(t)$ e $n^{1/2} \log A_n^1(t)$. A escolha de $p(t)$ que minimiza a variância assintótica, $\Gamma(t)$, é:

$$p(t) = \frac{\Gamma_1(t, t) - C(t)}{\Gamma_0(t, t) + \Gamma_1(t, t) - 2C(t)}. \quad (7.25)$$

Agora, deve-se definir o novo estimador em termos de A_n . Se $Z_{(1)}, \dots, Z_{(n)}$ se referirem aos Z_i 's ordenados e se

$$Q_i = \left\{ \prod_{k=1}^n \frac{Z_{(k)}}{1 - Z_{(k)}} \right\}^{1/n}, \quad 1 \leq i \leq n \quad (7.26)$$

Sendo assim, Caperaa et al. (1997) mostra que A_n pode ser escrita na forma fechada

$$A_n(t) = \begin{cases} (1 - t)Q_n^{1-p(t)} & \text{se } 0 \leq t \leq Z_{(1)}, \\ t^{i/n}(1 - t)^{1-i/n}Q_n^{1-p(t)}Q_i^{-1} & \text{se } Z_{(i)} \leq t \leq Z_{(i+1)} \quad (1 \leq i \leq n - 1), \\ tQ_n^{-p(t)} & \text{se } Z_{(n)} \leq t \leq 1, \end{cases} \quad (7.27)$$

sugerindo que os Z_i 's são distintos. Este estimador satisfaz $A_n(0) = A_n(1) = 1$, providenciando que $p(0) = 1 - p(1) = 1$. Este é um estimador assintoticamente não viesado de A e é geralmente uniforme, fortemente consistente.

8 Aplicação Prática

8.1 Introdução

Os dados utilizados nesse trabalho referem-se a duas variáveis aleatórias de perdas operacionais reais de uma instituição financeira brasileira. Por motivo de segurança da informação, os dados foram mascarados e os verdadeiros nomes dos eventos de perdas operacionais foram omitidos, bem como o nome da instituição que concedeu estes dados. Doravante denotaremos tais variáveis por FLT e DNN, e a instituição concedente por IF.

Nesse capítulo faremos uma análise descritiva dos dados, em busca de uma melhor compreensão do seu comportamento e relações. Estimaremos os parâmetros da GEV para cada variável, bem como os parâmetros das cópulas extremas de Gumbel e Galambos, em busca de modelos que se ajustem aos dados empíricos. E, por fim, analisaremos a estrutura de dependência destas variáveis e calcularemos o VaR, comparando o VaR obtido sem considerarmos a estrutura de dependência e o VaR obtido quando a consideramos, para cada cópula.

8.2 Análise Descritiva

A base de dados original é constituída por observações mensais de perdas, onde, em cada mês, podem ocorrer diferentes quantidades de perdas para cada evento de risco operacional. Utilizando a abordagem *Block Maxima*, citada na seção 4.2.2, foram extraídos os máximos mensais para cada uma das variáveis de perdas contempladas nesse estudo, FLT

e DNN, através do software SAS. Portanto, a base de dados final utilizada consiste em observações máximas mensais dos eventos de perdas operacionais: FLT e DNN.

Os registros de perdas operacionais feitos pela Instituição Financeira e incorporados a esse estudo datam de janeiro de 2004 até abril de 2012, o que nos proporcionou dados de 100 meses e, portanto, 100 observações para cada variável de perdas operacionais.

Para um melhor entendimento inicial do comportamento de tais variáveis separadamente e possíveis relações, segue uma análise descritiva, através de média, desvio padrão, máximo, mínimo e alguns quantis, e uma análise associativa, através dos coeficientes de correlação de Pearson, Spearman e Kendall. Além dos gráficos de séries temporais, dispersão e histogramas.

Tabela 8.1 – Estatísticas Descritivas para os Conjuntos de Dados FLT e DNN

Variável	Média	Desvio Padrão	Q1	Mediana	Q3	Mínimo	Máximo
FLT	3.351.324,28	4.444.922,54	946.531	1.530.610	3.371.907	395.514	24.231.976
DNN	440.865,31	1.769.145,4	15.797,43	37.548,87	107.498,64	2.296,74	13.897.950,3

Na variável FLT podemos perceber uma proximidade entre a média e o quantil 0,75, o que evidencia uma forte influência de valores extremos sobre o conjunto de dados, mesmo se tratando de uma distribuição de máximos. Isso se torna ainda mais evidente quando observamos a variável DNN, onde sua média é mais que quatro vezes a observação que supera 75% das observações em valor monetário desse conjunto. Isso significa que, para estes dados, em um mês a IF pode sofrer determinada perda e em um outro mês uma perda muito maior ou muito menor do que aquela observada anteriormente. Esse fato implica em uma dificuldade, por parte do gestor, em saber exatamente quando despenderá futuramente para cobrir essas perdas.

Uma medida que nos auxilia a verificar melhor essa heterogeneidade entre as observações é o coeficiente de variação, cujos valores são 132,63 e 401,29 para FLT e DNN, respectivamente. Valores altíssimos, pois suas interpretações nos permitem afirmar que, em média, os desvios de FLT relativamente à sua média atingem 13.263% do valor desta, a mesma análise pode ser feita para DNN, onde chegamos ao valor astronômico de 40.129%; mas que indicam uma possível adequação das distribuições destas variáveis às distribuições de cauda pesada, onde temos muitos valores pequenos e poucos valores grandes. Tal relação pode ser observada através dos gráficos de séries temporais abaixo.

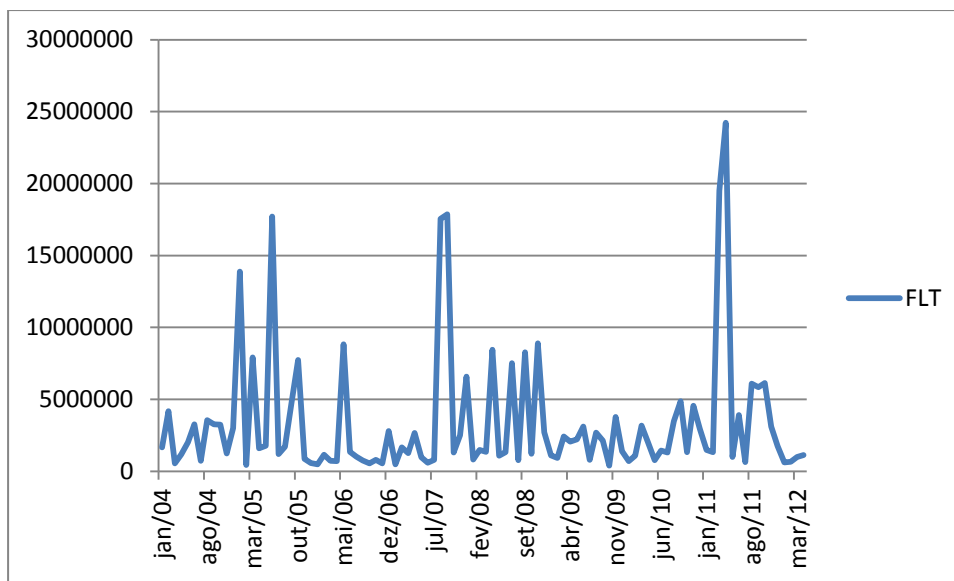


Figura 8.1: Série temporal dos dados de máximos mensais FLT

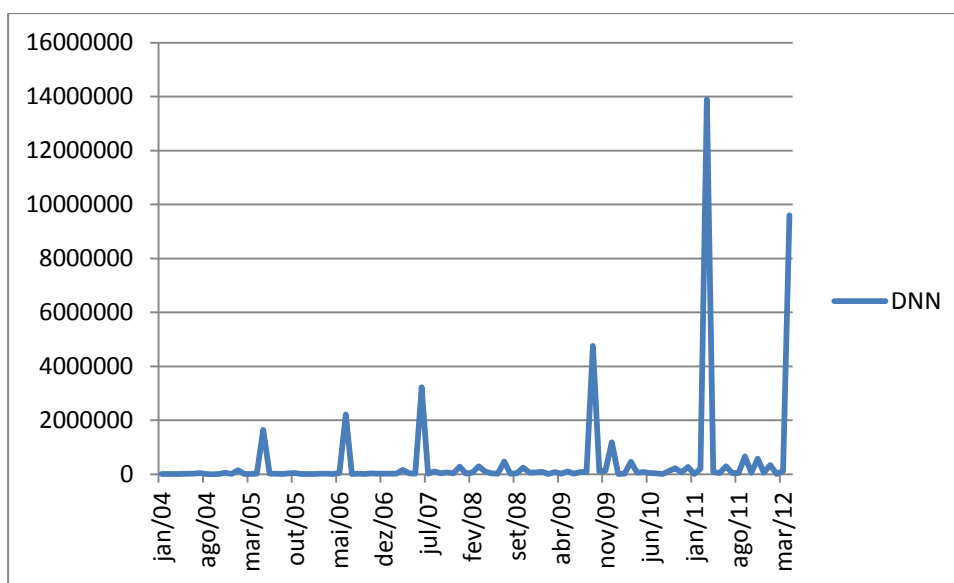


Figura 8.2: Série temporal dos dados de máximos mensais DNN

Através das figuras 8.1 e 8.2, podemos visualizar a alta volatilidade mensal das observações de máximos de perdas operacionais. Também podemos identificar a pequena quantidade de valores muito grandes e a alta quantidade de valores pequenos, característica intrínseca às distribuições de cauda pesada. Tal fato fica bem fácil de ser percebido através da análise dos histogramas destas variáveis, basta percebermos a alta densidade no lado

esquerdo dos gráficos. Para uma melhor compreensão da distribuição das variáveis estudadas devemos observar seus histogramas.

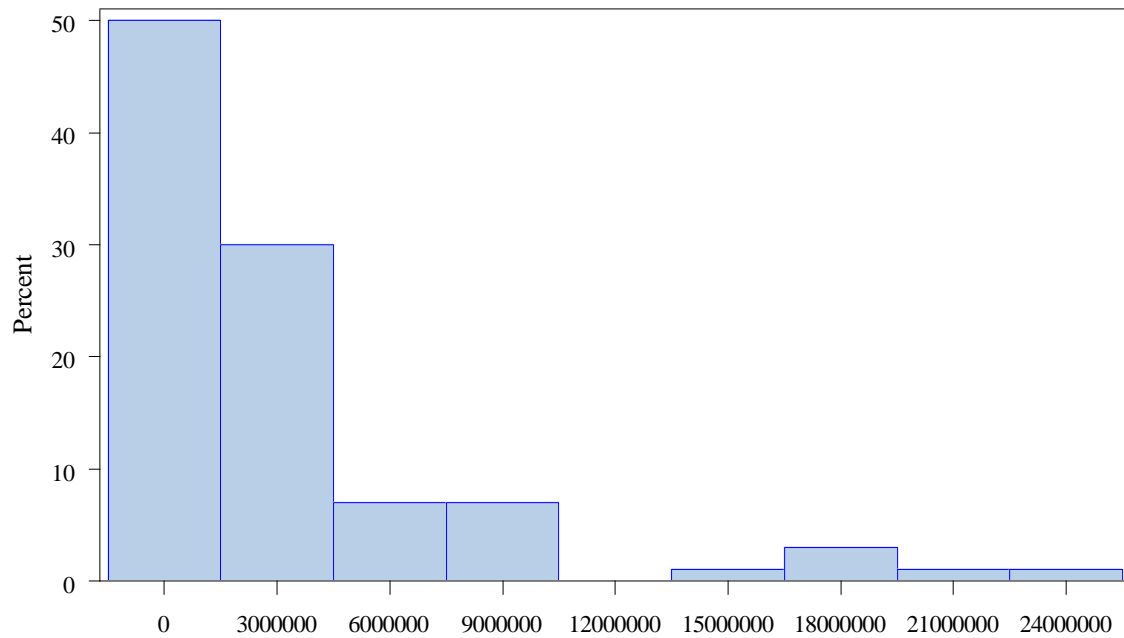


Figura 8.3: Histograma dos dados de máximos mensais FLT

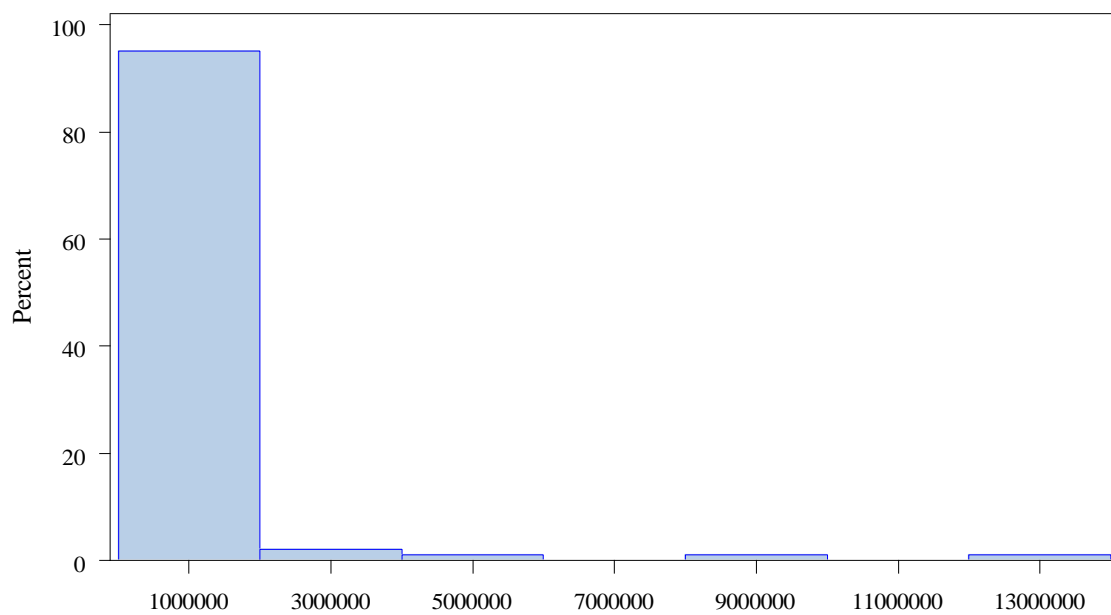


Figura 8.4: Histograma dos dados de máximos mensais DNN

Agora que temos conhecimento do comportamento individual destas variáveis, devemos estudar suas relações mútuas. Para tanto, a um nível de significância de 5%, o coeficiente de correlação linear de Pearson foi $\hat{r} = 0,2364$, o coeficiente de correlação por postos de Spearman foi $\hat{\rho}_s = 0,2530$ e o de Kendall foi $\hat{\tau} = 0,1721$; todos significativos. A figura 8.5 nos mostra o gráfico de dispersão de FLT e DNN.

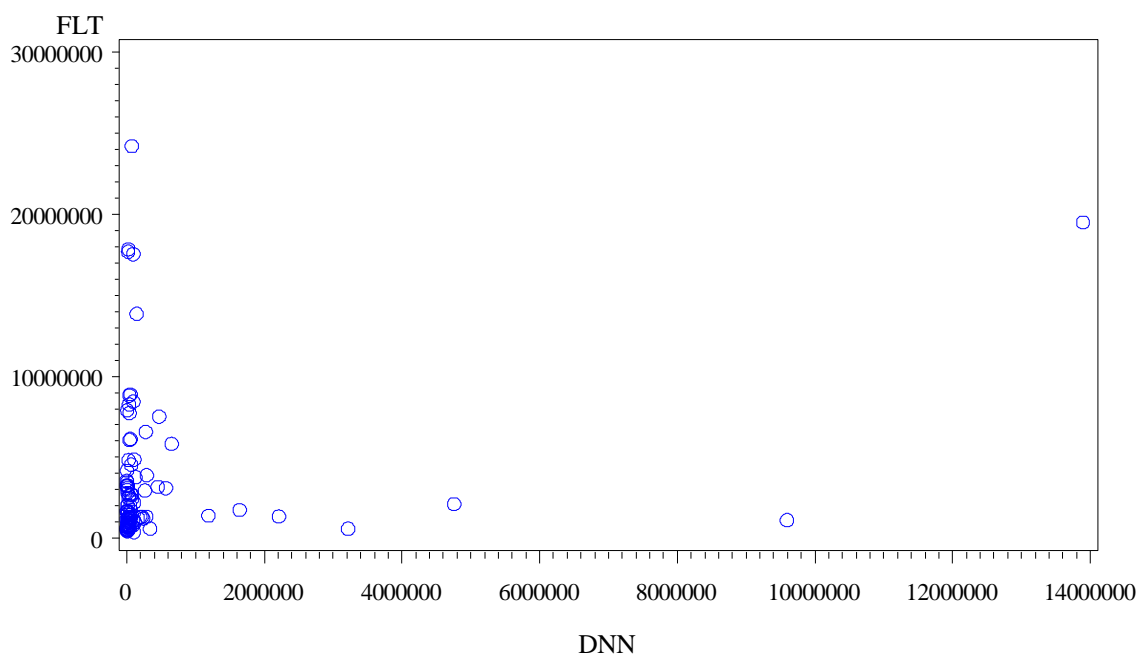


Figura 8.5: Gráfico de dispersão dos dados de máximos mensais de FLT e DNN

Tanto os coeficientes de correlação calculados quanto o gráfico de dispersão nos informam que a relação entre os máximos de FLT e DNN é fraca e positiva, ou seja, quanto maior for a perda em FLT, maior será a perda em DNN, e vice-versa. Talvez essa constatação esteja sendo enviesada pelo mês que resultou no ponto discrepante observado no canto superior direito da figura 8.5, mas não podemos excluir tal observação de nosso estudo, visto que este trabalho concerne justamente no estudo de observações que fogem da média.

Para uma melhor conclusão sobre a dependência caudal dos dados devemos utilizar medidas desenvolvidas a partir das cópulas extremas. Para tanto devemos ajustar modelos que reflitam as distribuições empíricas e conjunta. A próxima seção concerne na estimação dos parâmetros de tais modelos.

8.3 Estimação dos Parâmetros

A partir dos histogramas para as variáveis estudadas, apresentados nas figuras 8.3 e 8.4, e através dos qq-plots apresentados na figura 8.6, com base na distribuição Gumbel, podemos levantar a suspeita de que as distribuições dos máximos mensais de FLT e DNN podem ser ajustadas por uma Frechet, caso especial da GEV. A tabela 8.2 apresenta as estimativas dos parâmetros ξ , μ e σ da GEV.

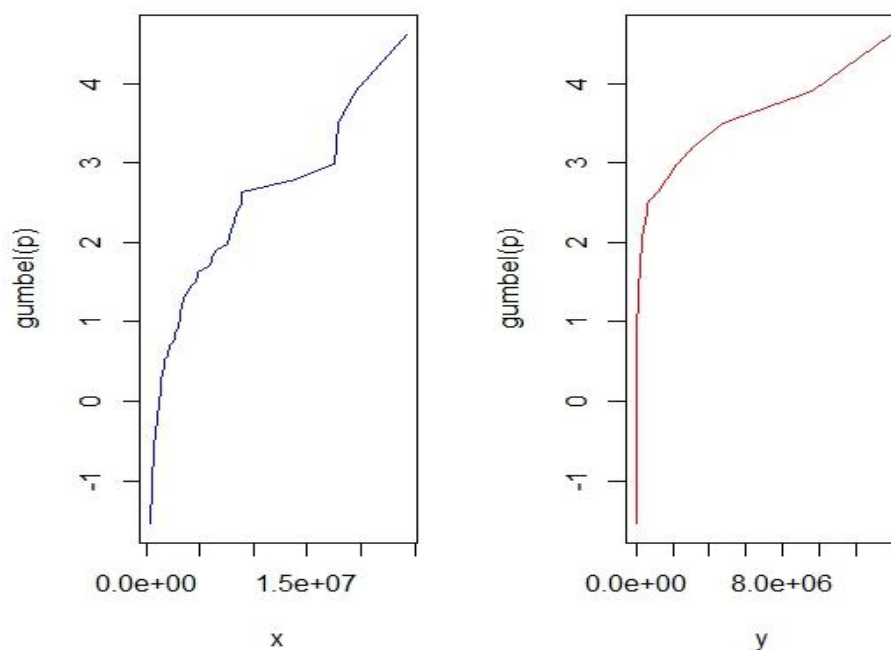


Figura 8.6: QQ-Plots de FLT (x) e DNN (y)

Tabela 8.2 – Estimativas dos Parâmetros da GEV

Variável	$\hat{\xi}$	$\hat{\mu}$	$\hat{\sigma}$
FLT	0,83508	1.697.430	3.465.688
DNN	1,33843	23.086,27	30.610,80

A estimação de tais parâmetros foi feita pelo método de maximização da função de verossimilhança, através do software R. A figura 8.7 ilustra a qualidade do ajuste das

estimativas. Podemos notar que a variável DNN teve um melhor ajuste, mas ambas foram relativamente bem ajustadas.

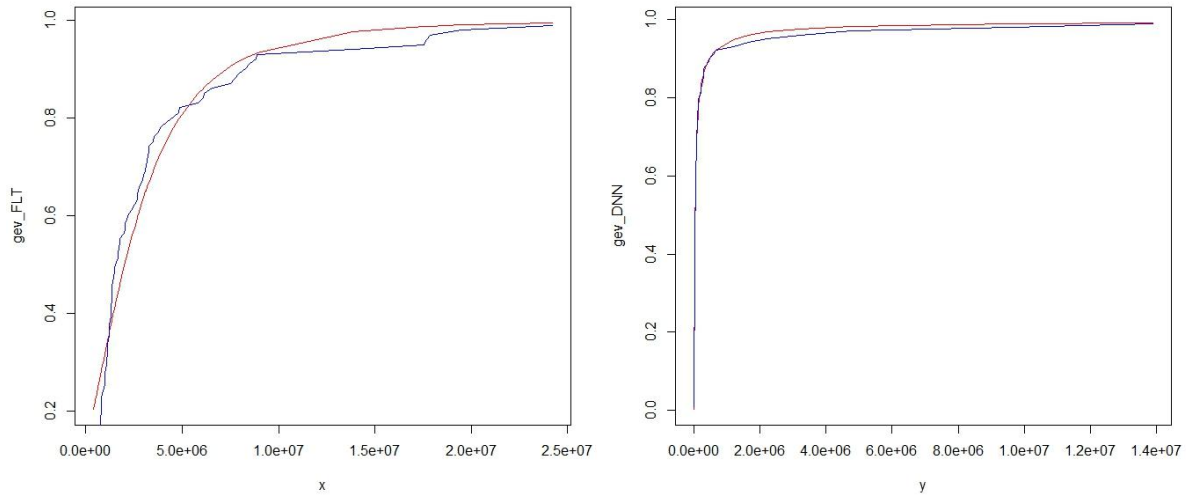


Figura 8.7: F.d. empírica (azul) x estimada (vermelho) de FLT e DNN

Agora que encontramos os modelos e os seus respectivos parâmetros que descrevem as distribuições de FLT e DNN, distribuições marginais, devemos utilizar tais estimativas para estimar os parâmetros das funções cópula.

As cópulas extremais de Gumbel e de Galambos foram escolhidas por serem amplamente utilizadas em dados financeiros e pela facilidade da implementação da maximização de sua função de verossimilhança. A tabela 8.3 apresenta a estimativa do parâmetro θ , feita pelo Método da Verossimilhança Exata em dois estágios, onde primeiro estimamos os parâmetros das marginais e os aproveitamos para estimar o parâmetro da cópula; bem como de $A(1/2)$, a partir de 6.13 e 6.15, e de λ_u , a partir de 6.21.

Tabela 8.3 - Estimativa do parâmetro da cópula, função de dependência e dependência caudal

Cópula	$\hat{\theta}$	$DP(\hat{\theta})$
Gumbel	1,22781	0,202
Galambos	0,52906	0,201

8.4 Resultados

Finalmente somos capazes de analisar as características do risco conjunto das variáveis FLT e DNN e a dependência extremal usando cópulas de valor extremo. Primeiramente calculamos o VaR para cada variável utilizando (6.19), a um nível de confiança de 99%.

Tabela 8.4 – Valor em Risco para FLT e DNN

Variável	VaR
FLT	195.019.842,91
DNN	10.800.161,56
Total	205.820.004,47

Portanto, quando não consideramos a estrutura de correlação entre FLT e DNN, o VaR total que deve ser alocado para cobrir perdas inesperadas referentes aos eventos de risco operacional FLT e DNN é de 205.820.004,47 unidades monetárias. Mas, como visto no capítulo 3, a simples soma dos VaR's individuais gera uma alocação de capital altamente conservadora. Para que consigamos obter um VaR mais apropriado, devemos estudar a estrutura de dependência da função cópula. Isso pode ser feito a partir das medidas encontradas na tabela 8.5.

Tabela 8.5 – Estrutura de dependência

Medida	Cópula	
	Gumbel	Galambos
$A(1/2)$	0,86513	0,87932
$\hat{\lambda}_u$	0,26974	0,24135
$C_0(\gamma, \gamma)$	0,98248	0,98276
$P(X > x, Y > y)$	0,00223	0,00251
$P(X > VaR(X) Y > VaR(Y))$	0,24819	0,27606

Temos que $A(1/2)$ nos dá a informação sobre a dependência caudal entre as variáveis. $A(1/2) = 1$ significa independência das duas variáveis, enquanto $A(1/2) = 0,5$ indica dependência positiva perfeita. Da Tabela 8.5, podemos ver que usando ambas as cópulas não há contraste, ambas

localizadas perto de 0,87, que mostra a dependência assintótica dos dados, que, no caso, comparado a 0,5 não é forte.

Baseado em (6.21), podemos facilmente calcular a estimativa da dependência caudal superior, sendo aproximadamente 0,27 para a cópula de Gumbel e 0,24 para a cópula de Galambos. A relativa proximidade entre estas estimativas pode representar que ambos os modelos de cópulas paramétricas são abeis em representar a estrutura de dependência dos nossos dados de perdas extremas mensais.

Para mensurar o risco caudal conjunto, utilizamos (6.20) para calcular a probabilidade da cauda conjunta excedente sob os dois diferentes modelos de cópulas extremas paramétricas. Os resultados, mostrados na Tabela 8.5, informam que a probabilidade de excedentes simultâneos aos respectivos VaR's das duas variáveis de perdas extremas é demasiadamente pequena.

Usando (6.22), encontramos a estimativa da probabilidade condicional de uma variável de perda, FLT ou DNN, superar seu VaR quando a outra variável superou seu VaR correspondente.

Finalmente, a partir de (6.18) calculamos o quantil indicado pela cópula que deve ser calculado de cada marginal para constituí o VaR quando consideramos a estrutura de dependência. Tal quantil nos forneceu os VaR's ilustrados na Tabela 8.6.

Tabela 8.6 – Valor em Risco de FLT e DNN considerando a estrutura de dependência

Variável	VaR	
	Gumbel	Galambos
FLT	122.354.252,80	124.002.866,38
DNN	5.086.483,32	5.197.799,74
Total	127.440.736,12	129.200.666,12

Podemos observar uma notável redução, aproximadamente 35%, no VaR total quando comparamos o VaR calculado como a soma dos VaR's das marginais com os VaR's das cópulas, sendo estes de 127.440.736,12 unidades monetárias pela cópula de Gumbel e 129.200.666,12 unidades monetárias pela cópula de Galambos. Devemos, então, enfatizar a necessidade do estudo da estrutura de dependência para cálculo do Valor em Risco Operacional.

8.5 Observações

O software SAS foi utilizado para preparação dos dados, extração dos máximos em cada período, construção dos histogramas e do gráfico de dispersão e cálculo das estatísticas descritivas e dos coeficientes de correlação.

O software R foi utilizado para estimar os parâmetros das marginais, bem como os das cópulas, e para construção dos gráficos de quantil-quantil.

O software Excel foi utilizado para construção dos gráficos de séries temporais e implementação de vários cálculos, principalmente aqueles contidos na seção 8.4.

As programações utilizadas nos softwares SAS e R estão contidas no Apêndice.

Conclusão

Neste trabalho pudemos observar a relevância da Teoria de Valor Extremo para a modelagem de dados de perdas operacionais de instituições financeiras. Vimos que a distribuição Generalizada de Valor Extremo, contida na Teoria de Valor Extremo, tem um alto poder de se ajustar a tais dados.

Também vimos que a utilização conjunta da Teoria de Valor Extremo com a teoria de Cópulas de Valor Extremo é de alta relevância nesse tipo de estudo, proporcionando ao analista informações riquíssimas.

A partir da teoria de cópulas pudemos estabelecer pressupostos de dependência caudal entre as variáveis estudadas e, a partir da estrutura de dependência, constatamos a necessidade, ou a aprimoração que podemos obter, em estimar o Valor em Risco Operacional através desta abordagem.

Portanto, os objetivos desse trabalho foram alcançados ao fazermos uma revisão da Teoria de Valor Extremo e de Cópulas, enfatizando as extremas, aplicarmos tais teorias a dados reais de perdas operacionais, elaborarmos pressupostos apropriados de correlação entre variáveis de perdas operacionais e estimarmos o Valor em Risco Operacional considerando a estrutura de dependência, obtendo uma redução de capital que deve ser alocado à cobertura de eventos de perdas operacionais.

Referências

- [1] Cebrián, A., Denuit, M., Lambert, P., Analysis of bivariate tail dependence using extreme values copulas: An application to the SOA medical large claims database. *Belgian Actuarial Journal* 3(1), 33-41, 2003.
- [2] Comitê da Basileia sobre a Supervisão Bancária. *Convergência internacional de mensuração de capital e padrões de capital*. Bank for International Settlements, 2004.
- [3] Cruz, M.G., *Modeling, Measuring and Hedging Operational Risk*. Wiley. 2002.
- [4] Embrechts, P., Klüppelberg, C., and Mikosch, T., *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin. 1997.
- [5] Embrechts, P., McNeil, A., and Straumann, D., *Correlation and dependence in risk management: properties and pitfalls*. Ed. M.A.H Dempster. Cambridge University Press, Cambridge. 2002.
- [6] Gudendorf, G., Segers, J., *Extreme-Value Copulas*. Institut de statistique, Université catholique de Louvain, Louvain-la-Neuve, Bélgica. 2009a.
- [7] Gudendorf, G., Segers, J., *Nonparametric estimation of an extreme-value copula in arbitrary dimensions*. Institut de statistique, Université catholique de Louvain, Louvain-la-Neuve, Bélgica. 2009b.
- [8] Kotz, S., Nadarajah, S., *Extreme value distributions: theory and applications*. Imperial College Press. London. 2002.
- [9] Lu, j., Tian, W., and Zhang, P., *The extreme value copulas analysis of the risk dependence for the foreign Exchange data*. IEEE, 2008.
- [10] Madeira, F. L., *Modelos de Risco Operacional com Distribuições de Cauda Pesada*. Monografia – Departamento de Estatística, Universidade de Brasília, 2011.
- [11] Neslehová, J., Embrechts, P., Chavez-Demoulin, V., *Infinite Mean Models and the LDA for Operational Risk*, in: *Journal of Operational Risk* 1(1), 3-25, 2006.
- [12] Nelsen, R., *An introduction to copulas*, second edn. Springer Series in Statistics. Springer, New York. 2006.

Apêndice

Códigos de programas

Código SAS

```
libname dados 'E:\Pasta';

data dados;
set dados.abr12;
keep dt_efetiva2 nu_evento_nivel_3 no_evento_nivel_3 valor_ajustado;
run;

proc sort data=dados;
by nu_evento_nivel_3 dt_efetiva2;
run;

proc means data=dados max;
var valor_ajustado;
by nu_evento_nivel_3 dt_efetiva2;
output out=dados1 max=maximo;
run;

proc sort data=dados1;
by dt_efetiva2;
run;

proc transpose data=dados1 out=dados2;
id nu_evento_nivel_3;
by dt_efetiva2;
run;

data dados3;
set dados2;
if _name_ = 'maximo';
run;

proc corr data=dados3 kendall spearman pearson;
var FLT DNN;
run;

goptions ;
symbol1 value=circle
```



```

ci=blue;
axis1 label=('DNN');
axis2 label=('FLT');

proc gplot data = dados3;
plot FLT*DNN / haxis=axis1 vaxis=axis2 ;
run;

proc univariate data=dados3;
var FLT DNN;
histogram FLT DNN;
run;

```

```

/*****

```

Código R

```

dados <- data.frame(read.table("C:/ Pasta.txt",header=TRUE))

x <- sort(dados$FLT)
y <- sort(dados$DNN)

plot(dados$FLT,dados$DNN,xlab="FLT",ylab="DNN")

par(mfrow=c(1,1))

plot(density(dados$FLT))
plot(density(dados$DNN))

gumbel <- function(p){
y <- -log(-log(p))
return(y)
}

n <- length(x)
p <- seq(1,n)
p <- p/(n+1)

par(mfrow=c(1,2))

## QQ-PLOT DE FLT
plot(x,gumbel(p), type="l", col="blue")

## QQ-PLOT DO DNN
plot(y,gumbel(p), type="l", col="red")

##### Função GevFit pacotes: evd, evir, extRemes, ismev
h <- gev.fit(x)

```

```

v <- gev.fit(y)
print(h)
print(v)

gev_FLT <- pgev(x,xi=0.835076e+00, mu=1.697430e+06 , sigma=3.465688e+06)
gev_DNN <- pgev(y,xi=1.33843, mu=23086.27137 , sigma=30610.80403)

par(mfrow=c(1,2))
plot(x,gev_FLT,type="l",col="red")
lines(x,p,col="blue")
plot(y,gev_DNN,type="l",col="red")
lines(y,p,col="blue")

##### Estimando a Cópula de Gumbel via MLE

x2 <- dados$FLT
y2 <- dados$DNN

uh <- pgev(x2,xi=0.835076e+00, mu=1.697430e+06 , sigma=3.465688e+06)[1:n]
vh <- pgev(y2,xi=1.33843, mu=23086.27137 , sigma=30610.80403)[1:n]

cgun <- function(u,v,a){
f <- exp(-((-log(u))^a + (-log(v))^a)^(1/a))
return(f)
}

dc <- function(a,b,c){
f <- (b/a)*(log(a)/log(b))^(c-1)
return(f)
}

dc2 <- function(u,v,a){
f <- (1/cgun(u,v,a))*dc(u,cgun(u,v,a),a)*dc(v,cgun(u,v,a),a)*((1-a)*(log(cgun(u,v,a)))^(-1)+1)
return(f)
}

loglest <- function(a){
f <- sum(log(dc2(uh,vh,a)))
return(f)
}

optim(5,loglest,gr=NULL, method="BFGS", control=list(fnscale=-1),hessian=TRUE)

#####Estimando a Cópula de Galambos via MLE

cgal <- function(u,v,a){
r <- u*v*exp(((ln(u))^(-a)+(ln(v))^(-a))^(-1/a))

```

```
}
```

```
dcgal2 <- function(u,v,a){  
r <- u*v*(1+a)*((u*v)^(-a-1))*(u^(-a)+v^(-a)-1)^(-2-1/a)  
return(r)  
}
```

```
lestc <- function(a){  
f <- sum(log(dcg2(uh,vh,a)))  
return(f)  
}
```

```
optim(5,lestc,gr=NULL, method="BFGS", control=list(fnscale=-1),hessian=TRUE)
```

```
#####
```