



Universidade de Brasília
Faculdade de Economia, Administração, Contabilidade e Ciência
da Informação e Documentação
Departamento de Economia

MESTRADO EM GESTÃO ECONÔMICA DE NEGÓCIOS

**MODELAGEM ESTOCÁSTICA EM RISCO OPERACIONAL
APLICANDO TEORIA DOS VALORES EXTREMOS**

JOÃO VAGNES DE MOURA SILVA

Brasília – DF

2006



Universidade de Brasília
Faculdade de Economia, Administração, Contabilidade e Ciência
da Informação e Documentação
Departamento de Economia

MESTRADO EM GESTÃO ECONÔMICA DE NEGÓCIOS

MODELAGEM ESTOCÁSTICA EM RISCO OPERACIONAL APLICANDO TEORIA DOS VALORES EXTREMOS

JOÃO VAGNES DE MOURA SILVA

Dissertação apresentada à Universidade
de Brasília, Departamento de Economia,
para obtenção do título de Mestre em
Gestão Econômica de Negócios

Orientador: Prof. Dr. Paulo César Coutinho

Brasília – DF

2006

FICHA CATALOGRÁFICA

SILVA, João Vagnes de Moura

Modelagem Estocástica em Risco Operacional – Aplicando Teoria dos Valores Extremos. Brasília, UnB, Programa de Pós-Graduação em Economia, 2006. 74 p.

Dissertação: Mestrado em Gestão Econômica de Negócios (Área Economia)

Orientador: Dr. Paulo César Coutinho

1. Risco Operacional 2. Teoria de Valores Extremos

I – Universidade de Brasília

II - Título

Cessão de Direitos

NOME DO AUTOR: João Vagnes de Moura Silva

TÍTULO DA DISSERTAÇÃO DE MESTRADO EM GESTÃO ECONÔMICA DE NEGÓCIOS: Modelagem Estocástica em Risco Operacional – Aplicando Teoria dos Valores Extremos.

GRAU/ANO: Programa de Pós-Graduação em Economia, 2006.

O autor reserva direitos de publicação e nenhuma parte desta dissertação de mestrado pode ser reproduzida sem autorização por escrito do autor.

JOÃO VAGNES DE MOURA SILVA

Matrícula 05/25855

**MODELAGEM ESTOCÁSTICA EM RISCO OPERACIONAL
APLICANDO TEORIA DOS VALORES EXTREMOS**

Dissertação aprovada como requisito para a obtenção do título de Mestre em Gestão Econômica de Negócios do Programa de Pós-Graduação em Economia – Departamento de Economia da Universidade de Brasília, por intermédio do Centro de Investigação em Economia e Finanças.

Brasília (DF), 26 de outubro de 2006.

BANCA EXAMINADORA

Prof. Paulo César Coutinho, PhD

Orientador

Prof. Tarcísio Barroso da Graça, PhD

Prof. Rodrigo Silveira Veiga Cabral, Dr

À Tatiana, Tiago e Natália.
Presentes que Deus me deu.

Agradecimentos

À minha esposa pelo incentivo, confiança e paciência.

Ao Banco do Brasil pelo apoio institucional.

Ao Dr. Paulo Coutinho pela confiança e orientação deste trabalho.

A todos os colegas do curso de Mestrado em Gestão Econômica de Negócios que durante dois anos partilharam do desafio acadêmico.

Ao amigo Ricardo Rangel pela convivência e contribuições singulares, importantes para meu desenvolvimento acadêmico e profissional.

Especial agradecimento à Lígia Guedes pelo apoio e contribuições no estudo de risco operacional.

Aos professores do curso de Mestrado em Gestão Econômica de Negócios.

À Márcia Nalu pelo atendimento sempre cortês e pela prestatividade.

Por fim, a todos que participaram deste trabalho com contribuições diretas ou indiretas.

Resumo

Nos últimos anos, a modelagem de risco operacional tornou-se fator crítico de sucesso para o gerenciamento de riscos em instituições financeiras.

Com o advento do novo acordo de capital – Basiléia II, diversos modelos de mensuração de risco operacional têm sido discutidos.

Este trabalho preocupa-se em discutir métodos avançados e modernos na mensuração de riscos operacionais, em especial avaliando a aplicabilidade da Teoria dos Valores Extremos – EVT, de forma a se medir o Valor em risco – VaR operacional – e quantificar adequadamente a necessidade de alocação de capital.

Será demonstrado que a modelagem estocástica aplicada a dados empíricos de severidade de perdas operacionais, utilizando-se a EVT como ferramenta, produz resultados em gestão de risco que atendem aos requisitos de robustez, confiabilidade e segurança definidos pelo Comitê de Basiléia.

Finalmente, justificando a abordagem quantitativa, o trabalho traz aplicação da Teoria dos Valores Extremos em gerenciamento de risco operacional, modelando duas séries de dados referentes à categoria de perdas com fraudes eletrônicas de um grande banco brasileiro, no período de janeiro de 2004 a dezembro de 2005, além da análise dos resultados obtidos.

Palavras-chave: Risco operacional. Teoria de valores extremos. VaR operacional.

Abstract

In recent years, the modeling of operational risk became a critical factor of success to an efficient risk management in financial institutions.

Many operational risk models have been reviewed with the advent of the New Agreement of Capital – Basel II.

This work is focused on advanced and modern method of operational risk measurement, in particular, the applicability of Extreme Value Theory – EVT as a consistent model to evaluate Operational Value at Risk and to quantify the allocation of capital necessity.

It will be demonstrated that EVT as a stochastic modeling tool is suitable to empirical data of operational losses severity, producing robust, security and reliable risk management results as expected by the Basel Committee.

Finally, justifying the quantitative approach, the study brings application of EVT in management of operational risk, shaping two data series of electronic frauds losses of a great Brazilian bank, from January 2004 to December 2005, in addition to the results analysis.

Key-words: Operational risk. Extreme value theory - EVT. Operational value at risk.

LISTA DE TABELAS

Tabela 4.1	– Valores significativos e críticos para a estatística K–S	39
Tabela 4.2	– Valores significativos e críticos para a estatística C–M	39
Tabela 5.1	– Funções de Distribuição de Probabilidade – Processo de Contagem	41
Tabela 6.1	– Categoria de Perdas – Distribuição	45
Tabela 7.1	– Estatística Descritiva – Severidade – Perdas Tipo 1	52
Tabela 7.2	– Estimação de Momentos e Constante c – GEV– Perdas Tipo 1	53
Tabela 7.3	– Estimação de Parâmetros – GEV – Perdas Tipo 1	53
Tabela 7.4	– Quantis – Análise Comparativa	56
Tabela 7.5	– Estatística Descritiva – Freqüência – Perdas Tipo 1	57
Tabela 7.6	– Estimação de Parâmetros – Freqüência – Perdas Tipo 1	57
Tabela 7.7	– AIC e BIC – Freqüência – Perdas Tipo 1	59
Tabela 7.8	– VaR Operacional e Capital Econômico – Perdas Tipo 1	61
Tabela 7.9	– Estatística Descritiva – Severidade – Perdas Tipo 2	62
Tabela 7.10	– Estimação de Momentos e Constante c – GEV– Perdas Tipo 2	64
Tabela 7.11	– Estimação de Parâmetros – GEV – Perdas Tipo 2	64
Tabela 7.12	– Estatística Descritiva – Freqüência – Perdas Tipo 2	65
Tabela 7.13	– Estimação de Parâmetros – Freqüência – Perdas Tipo 2	66
Tabela 7.14	– AIC e BIC – Freqüência – Perdas Tipo 2	67
Tabela 7.15	– VaR Operacional e Capital Econômico – Perdas Tipo 2	67

LISTA DE FIGURAS E GRÁFICOS

Figura 6.1	– Abordagens EVT	46
Figura 6.2	– Agregando Frequência e Severidade	48
Figura 6.3	– Capital Econômico	50
Gráfico 7.1	– Severidade – Perdas Tipo 1	51
Gráfico 7.2	– Máximo de Blocos – Perdas Tipo 1	53
Gráfico 7.3	– Distribuição GEV – Análise Comparativa	54
Gráfico 7.4	– Análise Comparativa – GEV versus Exponencial	55
Gráfico 7.5	– Histograma – Frequência – Perdas Tipo 1	56
Gráfico 7.6	– Análise Comparativa – Distribuições de Frequência – Perdas Tipo 1	58
Gráfico 7.7	– Análise Consolidada – Distribuições de Frequência – Perdas Tipo 1	59
Gráfico 7.8	– Severidade – Perdas Tipo	62
Gráfico 7.9	– Máximo de Blocos – Perdas Tipo 2	63
Gráfico 7.10	– Histograma – Frequência – Perdas Tipo 2	65
Gráfico 7.11	– Análise Consolidada – Distribuições de Frequência – Perdas Tipo 2	66

LISTA DE ABREVIATURAS E SIGLAS

AIC	–	Critério de Informação de Akaike
AMA	–	Modelos de Mensuração Avançada
BIC	–	Critério de Informação de Bayesian
BIS	–	Bank for International Settlements
C–M	–	Teste de Crámer – von Mises
CDF	–	Função de Distribuição Acumulada
EVT	–	Teoria de Valores Extremos
GEV	–	Distribuição Generalizada de Valor Extremo
GPD	–	Distribuição Generalizada de Pareto
K–S	–	Teste de Kolmogorov-Smirnov
MDA	–	Domínio Máximo de Atração
ML	–	Máxima Verossimilhança
NACB	–	Novo Acordo de Capitais de Basiléia
PDF	–	Função de Distribuição de Probabilidade
POT	–	Picos Sobre o Limite
PWM	–	Momentos de Probabilidade Ponderada
TMS	–	Taxa Média Selic
VaR	–	Valor em Risco

SUMÁRIO

AGRADECIMENTOS	vi
RESUMO	vii
ABSTRACT	viii
LISTA DE TABELAS	ix
LISTA DE FIGURAS E GRÁFICOS	x
LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS	xi
CAPÍTULO 1 INTRODUÇÃO	
1.1 – Contextualização	15
1.2 – Relevância do Estudo	17
1.3 – Objetivos	18
1.4 – Divisão da Dissertação	19
CAPÍTULO 2 MODELAGEM DE BANCO DE DADOS	
2.1 – Visão Geral	21
CAPÍTULO 3 TEORIA DOS VALORES EXTREMOS – EVT	
3.1 – Introdução	24
3.2 – Teoria Geral	25
3.3 – Distribuição Generalizada de Valor Extremo – GEV	27
3.4 – Distribuição Generalizada de Pareto – GPD	30
CAPÍTULO 4 ESTIMAÇÃO DE PARÂMETROS	
4.1 – Métodos de Estimação dos Parâmetros	33
4.2 – Teste de Qualidade de Ajuste	35

SUMÁRIO

CAPÍTULO 5 MODELAGEM DE FREQUÊNCIA E PERDAS AGREGADAS

5.1	– Análise da Frequência dos Eventos	40
5.2	– Seleção de Modelos	41
5.3	– Agregando Distribuição de Frequência e Severidade	42

CAPÍTULO 6 PROCEDIMENTOS METODOLÓGICOS

6.1	– Série de Dados	45
6.2	– Abordagem para Aplicação da EVT	46
6.3	– Estimação de Parâmetros	47
6.4	– Teste de Qualidade dos Ajustes	47
6.4.1	– Severidade	47
6.4.2	– Frequência	48
6.5	– Procedimentos Finais	48

CAPÍTULO 7 RESULTADOS OBTIDOS

7.1	– Banco de Dados – Perdas Tipo 1	51
7.1.1	– Análise de Severidade das Perdas	51
7.1.2	– Análise de Frequência das Perdas	56
7.1.3	– Agregação de modelos de Severidade e Frequência	60
7.2	– Banco de Dados – Perdas Tipo 2	60
7.2.1	– Análise de Severidade das Perdas	60
7.2.2	– Análise de Frequência das Perdas	63
7.2.3	– Agregação de modelos de Severidade e Frequência	66

SUMÁRIO

CAPÍTULO 8	CONCLUSÕES	67
	REFERÊNCIAS BIBLIOGRÁFICAS	71
	APÊNDICE	73

CAPÍTULO 1

INTRODUÇÃO

1.1 Contextualização

O Comitê da Basileia sobre a Supervisão Bancária é responsável por estabelecer diretrizes que garantam a convergência internacional sobre regulamentação da supervisão bancária, em especial, ao controle da adequação de capital de bancos com atividades internacionais.

O principal objetivo do Comitê da Basileia sobre a Supervisão Bancária ao revisar o Acordo de 1988 foi desenvolver uma estrutura que tem como prioridade proporcionar a solidez e a estabilidade do sistema bancário internacional e, ainda, manter consistência suficiente para que a regulação de adequação de capital não seja uma fonte de desigualdade competitiva entre os bancos.

Com o Novo Acordo de Capital de Basileia – NACB, constante da publicação *Basel Committee (2004) - Basel II: International Convergence of Capital Measurement and Capital Standards: a Revised Framework*, de junho de 2004, o Comitê buscou atingir exigências de capital mais sensíveis ao risco a fim de permitir que os bancos e as autoridades de supervisão selecionem abordagens que sejam mais adequadas às suas operações e à infra-estrutura do seu mercado financeiro.

O principal desafio do NACB é a inserção do risco operacional no ambiente financeiro e a proposta de metodologia para mensuração desse tipo de risco.

Em função da complexidade dos eventos geradores das perdas operacionais, bem como da diversidade de suas causas, o NACB propõe três abordagens de metodologias para o cálculo de alocação de capital, as quais possuem níveis crescentes de complexidade e expectativa de níveis decrescentes de requerimento de capital: o Método de Indicador Básico, o Método Padronizado e o Método de Mensuração Avançada (*Advanced Measurement Approaches – AMA*), que é um modelo de mensuração interna. Com isso, o Comitê busca oferecer uma estrutura de incentivos à utilização de modelos internos para gerenciamento de risco operacional das instituições financeiras e cálculo de capital regulamentar.

Assim, embasando-se nas abordagens propostas pelo Comitê, as instituições financeiras estão desenvolvendo mecanismos de identificação e avaliação de riscos operacionais inerentes a processos, produtos e serviços do âmbito de sua atividade.

Espera-se que instituições financeiras com exposições significativas a riscos operacionais implementem métodos mais sofisticados, seguindo as definições do Comitê para o AMA e buscando menor requerimento de capital em relação às outras duas abordagens.

No caso do AMA, o Comitê admite significativa flexibilidade, não especificando o método ou as premissas de distribuições de probabilidade a serem usadas para gerar a mensuração de risco operacional para fins de capital regulamentar. No entanto, as instituições financeiras devem demonstrar a robustez do método, provando que é capaz de registrar eventos de perdas potencialmente críticos.

1.2 Relevância da Pesquisa

O gerenciamento do risco operacional não é uma prática recente encontrada no sistema financeiro mundial. Na verdade, a administração existente desse tipo de risco é correntemente importante para os bancos no que se refere à prevenção de fraudes, manutenção da integridade de controles internos, redução de erros no processamento de transações dentre outras falhas. No entanto, o que é relativamente novo é a visão da gestão do risco operacional como uma prática comparável à gestão dos riscos de crédito e de mercado, onde o risco operacional passa a ser abordado como uma classe distinta de riscos e que também deve ser administrado, controlado e mensurado.

A importância da pesquisa reside, portanto, em discutir métodos avançados e modernos na mensuração de riscos operacionais de forma a se medir o Valor em risco – VaR operacional – e quantificar adequadamente a necessidade de alocação de capital.

O VaR pode ser definido como a perda máxima esperada ao longo de um horizonte de tempo predeterminado, considerando um intervalo de confiança definido (Jorion, 2000).

Há de se destacar que os métodos estatísticos e econométricos tradicionais para cálculo de VaR apresentam inconvenientes, uma vez que, geralmente, supõem normalidade. Mostram, assim, boa aderência nas regiões onde a maioria dos dados

se concentra, mas são mal comportados em relação ao ajuste das caudas, onde as observações extremas se concentram.

Na análise de risco operacional, um banco de dados típico de perdas operacionais não apresenta uma distribuição gaussiana. Em geral, é composto de poucos eventos muito grandes e vários eventos menores. Do ponto de vista de gerenciamento desses riscos, o interesse reside onde poucas observações se concentram – nas caudas (perdas máximas) – e o uso de métodos tradicionais de cálculo de VaR não são apropriados.

Em outras palavras, o baixo nível de α -percentil é a chave que abre a porta para a entrada da Teoria dos Valores Extremos – EVT no estudo de riscos operacionais, sendo usada para estudar perdas não esperadas de alto impacto e baixa frequência.

Cruz (2002) defende que a pergunta que deverá ser respondida pelo gerente de risco é: “Quanto capital econômico devo alocar a uma empresa particular a fim de protegê-la contra uma possível catástrofe operacional?”. Tal questionamento encontra solução na análise de distribuições de valores extremos.

1.3 Objetivos

Esta dissertação tem como objetivos gerais estudar a Teoria dos Valores Extremos e avaliar sua aplicabilidade em gestão de riscos operacionais, explicitando a sua importância em modelagem de perdas operacionais.

Visa, ainda, demonstrar que a modelagem estocástica aplicada a dados empíricos de severidade de perdas operacionais, utilizando-se a Teoria dos Valores Extremos como ferramenta, com o intuito de apurar o requerimento de alocação de capital ótimo, produz resultados em gestão de riscos operacionais que atendem aos requisitos de robustez, confiabilidade e segurança definidos pelo Comitê para o uso do AMA.

Além disso, procurar-se-á discutir uma forma prática de mensurar o valor em risco para perdas operacionais, com o intuito de apurar o capital econômico necessário para fazer frente a tais perdas.

1.4 Divisão da Dissertação

Esta dissertação está dividida em oito capítulos, apresentados de forma a possibilitar o entendimento da teoria de valores extremos, desde sua base teórica até a modelagem estocástica, visando gerenciar riscos operacionais.

No Capítulo 2, destaca-se a importância dos bancos de dados para análise e gestão de risco operacional.

No Capítulo 3, concentra-se a base teórica da modelagem de valores extremos, descrevendo-se diversos tipos de distribuições de valores extremos e seus métodos de modelagem.

As técnicas de estimação de parâmetros para as distribuições de valores extremos estão detalhadas no Capítulo 4. Neste capítulo, também são detalhados testes de qualidade do ajuste entre dados empíricos e distribuições teóricas a fim de verificar se o modelo é apropriado.

No Capítulo 5, discute-se o processo estocástico discreto seguido pelos eventos de perdas operacionais, esclarecendo a necessidade de estudar modelos de contagem estocásticos que informem sobre a distribuição do número de eventos em um período de tempo. Além disso, neste capítulo, demonstra-se como combinar os processos de severidade e frequência em uma distribuição de perdas agregadas que permita calcular um valor máximo para as perdas operacionais com certo grau de confiabilidade (VaR).

Os procedimentos metodológicos utilizados no trabalho estão detalhados no Capítulo 6 e os resultados obtidos na pesquisa são apresentados no Capítulo 7.

Finalmente, no Capítulo 8 são apresentadas as conclusões sobre os resultados alcançados, detalhados no Capítulo 7, bem como comentários sobre requisitos computacionais para modelagem de risco operacional, além de sugestões para continuidade de pesquisas sobre modelagem estocástica em risco operacional.

CAPÍTULO 2

MODELAGEM DE BANCO DE DADOS

2.1 Visão Geral

Em termos formais, o BIS (2001) definiu o risco operacional como “o risco de perda resultante de pessoas, sistemas e processos internos inadequados ou deficientes, ou de eventos externos”.

O rastreamento de dados de eventos de perdas internas é um pré-requisito para o desenvolvimento e funcionamento de uma sistemática de mensuração de risco operacional confiável. Segundo Marshall (2002), as fontes de dados internas devem ser o começo de qualquer análise de perdas. São cruciais para relacionar as estimativas de risco à real experiência de perda.

Cruz (2002) entende que estabelecer um banco de dados útil e robusto é uma fase crítica na modelagem e gerenciamento do risco operacional. Para ele, o analista deve ser bastante cuidadoso nesse estágio para determinar os tipos de perdas que serão de interesse na análise. Deve-se destacar que os dados de perdas internas são mais pertinentes à medida que estão claramente relacionados às atividades comerciais, aos processos tecnológicos e aos procedimentos de administração de riscos da instituição. Portanto, procedimentos documentais para avaliar a relevância

contínua dos dados de perdas internas históricas têm alta relevância na formação do banco de dados.

O NACB (2004) estabelece que:

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, quer os dados sejam usados diretamente para criar a mensuração de perdas, quer para validá-las.

Concluí-se que, para qualquer pesquisa alcançar êxito, é necessário um banco de dados suficientemente grande para que se possa aplicar a Teoria dos Valores Extremos com sucesso, sendo a confiabilidade dos dados outro fator crítico que pode comprometer os resultados da pesquisa. Estes são os verdadeiros obstáculos para a modelagem de risco operacional e não a complexidade da modelagem por EVT.

Nesta pesquisa, são utilizados dois bancos de dados de perdas, que serão intitulados Tipo 1 e Tipo 2.

- Tipo 1 - refere-se às perdas em transações com cartão de débito em terminais POS;
- Tipo 2 – refere-se às perdas em transações nos terminais de auto-atendimento - TAA.

Tanto no banco de dados Tipo 1 quanto no Tipo 2, a quantidade de eventos é apurada diariamente (n eventos por dia), formando o banco de dados de frequência.

Já o valor das perdas (severidade) é apurado por evento, atualizado até uma data focal pela taxa média Selic – TMS, compondo, assim, o banco de dados primário. A partir deste banco de dados, cria-se o banco de dados secundário que representará um vetor de máximos e será utilizado na modelagem de distribuição de valores extremos.

Ambos compreendem o período de janeiro de 2004 a dezembro de 2005. Apesar de não conter cinco anos, como sugere o NACB, mas apenas dois anos, há nestes casos, uma amostra bastante expressiva, pois muitos eventos ocorrem no mesmo dia, gerando um banco de dados que possibilita uma boa modelagem.

CAPÍTULO 3

TEORIA DOS VALORES EXTREMOS – EVT

3.1 Introdução

O uso de EVT permite estimar probabilidades e quantis com razoável precisão nos limites dos valores encontrados dentro da amostra, e mesmo além deles. A explícita modelagem dos valores extremos corresponde a um ajuste apenas da cauda da distribuição dos dados originais, ressaltando o caráter pouco informativo que pode ter o interior da distribuição para as ocorrências de baixíssima frequência.

A EVT já foi aplicada em diversas áreas. Gumbel (1958) aplicou em engenharia naval, geologia e engenharia aeronáutica. Destaca, por exemplo, que para a construção de aeronaves, as velocidades das rajadas de vento são de interesse particular, pois rajadas de alta velocidade, eventos estes extremos, podem causar danos à estrutura das aeronaves. Embrechts et al. (1997) aplicou a seguros e finanças e sua abordagem é atualmente uma referência padrão sobre EVT.

Para McNeil (1999) o risco de eventos extremos está presente em todas as áreas de gerenciamento de riscos e a EVT pode contribuir para solucionar problemas concretos de gerenciamento de riscos, seja por conta de movimentos

adversos de mercado, *default* em operações de crédito ou problemas operacionais imprevistos.

Medova (2000 e 2001) procura alinhar as expectativas do Comitê de Basiléia, inclusive antes da revisão de 2004, à aplicação de EVT, justificando que o real interesse dos reguladores é a ocorrência de eventos de baixa probabilidade e alta severidade, os quais podem produzir grandes perdas suficientes para ameaçar a saúde de uma instituição financeira. Nos seus trabalhos, apresenta as técnicas de EVT como melhores alternativas para alcançar a mensuração do capital a ser alocado devido às perdas operacionais.

Mais recentemente, Chavez-Demoulin e Embrechts (2004) apresentaram técnicas avançadas de EVT para modelagem de risco operacional, com adaptações para os casos de não-estacionariedade (estrutura dependente em relação ao tempo).

Observa-se, pois, que a aplicação da EVT como teoria que sustenta as distribuições de eventos raros é conhecida, mas ainda está num estágio inicial em gestão de riscos operacionais.

3.2 Teoria Geral

A fundamentação teórica de EVT aqui apresentada, necessária para modelagem de risco operacional, é detalhada em Embrechts et al. (1997).

Suponha que X denote as perdas operacionais no banco de dados. Seja $\chi_n = \{X_1, X_2, \dots, X_n\}$. Chama-se de estatísticas de ordem os elementos de χ_n ordenados de forma que $X_1 \leq X_2 \leq \dots \leq X_n$.

Os extremos são definidos como o máximo e o mínimo das n variáveis aleatórias ordenadas X_1, X_2, \dots, X_n . Denominando-se o máximo amostral por:

$$M_1 = X_1 \text{ e } M_n = \text{Máx}(X_1, X_2, \dots, X_n) \text{ para } n \geq 2$$

Supondo χ_n i.i.d, pode-se escrever-se distribuição exata do máximo como:

$$Pr\{M_n \leq x\} = Pr\{X_1 \leq x, \dots, X_n \leq x\} = Pr\{X_1 \leq x\} \times Pr\{X_2 \leq x\} \times \dots \times Pr\{X_n \leq x\} = F^n(x)$$

Onde $F(x) = Pr\{X \leq x\}$.

O apresentado acima sugere uma lei de potência para a distribuição do máximo. Para encontrar uma distribuição limitadora que não seja degenerada¹, a variável aleatória máxima pode ser padronizada por parâmetros de localização, escala e forma, escolhidos para dar uma distribuição apropriada de extremos padronizados. Portanto, deve-se concentrar no comportamento assintótico dos extremos. O resultado central da EVT é enunciado pelo teorema a seguir:

¹ Se X_1, X_2, \dots, X_n são variáveis aleatórias independentes com função de densidade de probabilidade comum $p_{x_j}(x) = f(x)$, $j = 1, 2, \dots, n$, então a função de distribuição cumulativa de $X'_n = \max(X_1, X_2, \dots, X_n)$ é $F_{X'_n}(x) = [F(x)]^n$, onde $F(x) = \int_{-\infty}^x f(x) dx$. Com n tendendo a infinito, para algum valor fixo de x , tem-se: $\lim_{n \rightarrow \infty} F_{X'_n}(x) = \begin{cases} 1 & \text{se } F(x) = 1 \\ 0 & \text{se } F(x) < 1 \end{cases}$, que é uma distribuição degenerada (Kotz e Nadarajah, 2000, págs. 4 e 5).

Teorema de Fisher-Tippett: Lei de Limite para o Máximo: Suponha χ_n uma seqüência de variáveis aleatórias i.i.d.. Se existem constantes de padronização $c_n > 0$ e $d_n \in \mathfrak{R}$, e alguma distribuição não-degenerada H tal que:

$$c_n^{-1} \times (M_n - d_n) \xrightarrow{d} H,$$

então, H pertence a um dos três seguintes tipos de distribuição:

$$\begin{aligned} \text{Fréchet} &\Rightarrow \Phi_\alpha(x) = \begin{cases} 0 & \text{se } x \leq 0 \\ \exp(-x^{-\alpha}) & \text{se } x > 0 \end{cases} \quad \alpha > 0 \\ \text{Weibull} &\Rightarrow \Psi_\alpha(x) = \begin{cases} \exp(-x^\alpha) & \text{se } x \leq 0 \\ 0 & \text{se } x > 0 \end{cases} \quad \alpha > 0 \\ \text{Gumbel} &\Rightarrow \Lambda(x) = \exp(-e^{-x}), x \in \mathfrak{R} \end{aligned} \quad (3.1)$$

Outro conceito bastante útil na EVT é a forma reversa de se observar o teorema de Fisher-Tippett. Diz-se que uma variável aleatória X ou a distribuição de X pertence ao domínio de atração do máximo de H (MDA – Maximum Domain of Attraction) se existirem as constantes $c_n > 0$ e $d_n \in \mathfrak{R}$ que produzam

$$c_n^{-1} \times (M_n - d_n) \xrightarrow{d} H.$$

3.3 Distribuição Generalizada de Valor Extremo (GEV)

Uma representação conveniente para as três distribuições citadas anteriormente é dada pela Distribuição Generalizada de Valor Extremo (GEV). Essa distribuição de três parâmetros – localização, forma e escala, μ, ξ e ψ , respectivamente –, surge como a distribuição limite normalizada do máximo de

variáveis aleatórias i.i.d. e o problema recai na estimação do parâmetro de interesse que dará o formato da distribuição. Dessa forma:

$$H_{\xi} = \begin{cases} \Phi_{\frac{1}{\xi}} & \text{se } \xi > 0 \\ \Lambda & \text{se } \xi = 0 \\ \Psi_{-\frac{1}{\xi}} & \text{se } \xi < 0 \end{cases}$$

Para qualquer F pertencente ao MDA de H_{ξ} , a aproximação $F^n(c_n x + d_n) \approx H_{\xi}$ para a distribuição do extremo é válida.

Mais precisamente, pode-se definir H_{ξ} como:

$$H_{\xi}(x) = \begin{cases} \exp\left[-(1 + \xi x)^{-\frac{1}{\xi}}\right] & \text{se } \xi \neq 0 \\ \exp[-\exp(x)] & \text{se } \xi = 0 \end{cases} \quad \text{onde } (1 + \xi x) > 0 \quad (3.2)$$

Além disso, é possível substituir x na expressão acima pela transformação de escala e localização $\frac{x - \mu}{\psi}$, com $\mu \in \mathfrak{R}$ e $\psi > 0$. Assim, chega-se à distribuição:

$$H_{\xi, \mu, \psi}(x) = \begin{cases} \exp\left[-\left(1 + \xi \frac{x - \mu}{\psi}\right)^{-\frac{1}{\xi}}\right] & \text{se } \xi \neq 0 \\ \exp\left[-\exp\left(\frac{x - \mu}{\psi}\right)\right] & \text{se } \xi = 0 \end{cases} \quad \text{onde } \left(1 + \xi \frac{x - \mu}{\psi}\right) > 0 \quad (3.3)$$

denominada GEV, onde ξ é o parâmetro de forma.

Fazendo $\xi \rightarrow 0$, temos a distribuição de Gumbel; $\xi > 0$ a distribuição de Fréchet e $\xi < 0$ a distribuição de Weibull.

Para se modelar uma função de distribuição que se adeque à GEV, o melhor caminho é o método dos Máximos de Blocos.

Suponha que se pretende estimar uma distribuição F_{M_n} do máximo entre as variáveis aleatórias X_1, \dots, X_n , pertencentes a χ_n . A solução natural é dividir as N observações em blocos de tamanho n . Dessa forma, $M_n^{(j)} = \max(X_1^{(j)}, X_2^{(j)}, \dots, X_n^{(j)})$ é o máximo de n observações em cada j -ésimo bloco.

Após este procedimento, constrói-se a distribuição de $M_n^{(1)}, \dots, M_n^{(k)}$. Esta função designada por \hat{F}_{M_n} é tomada como a estimativa desejada de F_{M_n} .

Portanto, o método de máximo de blocos da EVT consiste em usar métodos de estatística paramétrica para obter estimativas dos parâmetros da distribuição com base na amostra $M_n^{(1)}, \dots, M_n^{(k)}$. Este método resulta em uma aproximação da função de distribuição desconhecida F_{M_n} por H_ξ , que é uma função do tipo GEV.

Quando o problema em gerenciamento de risco operacional é determinar a quantidade de capital econômico a ser alocado para cobrir as perdas operacionais inesperadas, uma maneira simples de fazê-lo é apurar o quantil, que é o valor correspondente à probabilidade de algum evento ocorrer. Segundo Embrechts et al (1997), para a GEV, o quantil \hat{x}_p é dado por:

$$\hat{x}_p = \hat{\mu} - \frac{\hat{\Psi}}{\hat{\xi}} \left[1 - (-\ln p)^{-\hat{\xi}} \right] \quad (3.4)$$

3.4 Distribuição Generalizada de Pareto (GPD)

Outra importante distribuição para medir eventos raros e extremos é a Distribuição Generalizada de Pareto – GPD, que possui dois parâmetros e tem função de distribuição dada por:

$$G_{\xi, \beta}(x) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\beta}\right)^{-\frac{1}{\xi}} & \xi \neq 0 \\ 1 - \exp\left(-\frac{x}{\beta}\right) & \xi = 0 \end{cases} \quad \text{onde } x \in \begin{cases} [0, \infty] & \xi \geq 0 \\ \left[0, -\frac{\beta}{\xi}\right] & \xi < 0 \end{cases} \quad (3.5)$$

ξ é o importante parâmetro de forma da distribuição e β é um parâmetro de escala.

Neste caso, tem-se:

- Se $\xi > 0$, então $G_{\xi, \beta}$ é uma versão reparametrizada da distribuição ordinária de Pareto;
- Se $\xi = 0$, $G_{\xi, \beta}$ corresponde a uma distribuição exponencial;
- Se $\xi < 0$, a distribuição é conhecida como uma distribuição de Pareto Tipo II (Beta).

Segundo Medova e Kyriacou (2002), em perdas operacionais os extremos tendem a incrementar sem saltos todo o tempo antes de agrupar próximo a limite superior bem definido. Isto sugere que a estimativa de forma para a GPD pode ser esperada como não negativa.

Por isso, o primeiro caso é o mais relevante para gerenciamento de risco operacional, uma vez que a GPD apresenta-se como uma distribuição de caudas grossas quando $\xi > 0$.

A GPD foi apresentada inicialmente por Pickands (1975). Posteriormente, diversos estudos aplicaram a GPD em sinistros de seguros e estudos de confiabilidade. Chavez-Demoulin e Embrechts (2004) resumem bem a metodologia de análise da EVT baseada no método do processo de ponto, em que se escolhe um limite e ignora-se o momento em que os eventos acontecem. Tal abordagem, conhecida como picos sobre o limite (Peaks Over Threshold – POT), é apresentada abaixo. Um maior detalhamento pode ser encontrado em Embrechts et al (1997).

O método POT permite o ajuste de modelos GPD mesmo com parâmetros que dependam do tempo (escala, forma e localização). Os excessos correspondentes além de um determinado limite u são independentes e têm uma GPD.

Suponha que $X_{1,n}, \dots, X_{k,n}$ são i.i.d com função de distribuição $F \in \text{MDA}(G_\xi)$ para algum $\xi \in \mathfrak{R}$. Então, selecione um limite u . O número de excessos ao limite

pode ser indicado por $M_1, \dots, M_n, M_n < X_{k,n}$. O excesso da função de distribuição de x pode ser descrito como:

$$F_u(y) = P[X - u \leq y \mid X > u] \quad (3.6)$$

para $0 \leq y < x_0 - u$ onde $x_0 \leq \infty$

O seguinte teorema explica a importância da GPD:

Teorema: Para uma extensa classe de distribuições subjacentes, pode-se encontrar uma função $\beta(u)$ tal que:

$$\lim_{u \rightarrow x_0} \sup_{0 \leq y < x_0 - u} |F_u(y) - G_{\xi, \beta(u)}(y)| = 0 \quad (3.7)$$

Isto significa que, para inúmeras distribuições F , com limite u progressivamente aumentando, o excesso da função de distribuição F_u converge para uma Pareto Generalizada, ou seja:

$$F_u(y) \approx G_{\xi, \beta}(y) \quad (3.8)$$

A primeira dificuldade para aplicação desta abordagem é a escolha do limite u . Para McNeil (1999) a escolha do limite é basicamente um compromisso entre escolher um limite suficientemente alto tal que o teorema acima possa ser considerado essencialmente exato e escolher um limite suficientemente baixo tal que se tenha material (dados) suficientes para estimar os parâmetros da distribuição (ξ e $\beta = \beta(u)$).

CAPÍTULO 4

ESTIMAÇÃO DE PARÂMETROS

4.1 Métodos de Estimação de Parâmetros

Cruz (2002) ressalta que o parâmetro mais importante de distribuição de valores extremos é o parâmetro de forma ξ , que descreve o peso da cauda da distribuição. Se os dados se ajustarem a uma distribuição extrema, ξ provavelmente será significativo. Caso contrário, distribuições ordinárias como log-normal ou exponencial poderão ser suficientes.

Além do parâmetro de forma, devem-se determinar os parâmetros de escala ψ e localização μ tais que a distribuição de extremos padronizados não seja degenerada. Os métodos mais utilizados para estimar parâmetros de distribuição são: momentos, momentos de probabilidade ponderada (PWM) e máxima verossimilhança (ML).

O método dos momentos não é de interesse neste trabalho, pois, apesar de simples, provoca vieses altamente significativos.

Hosking et al. (apud Kotz e Nadarajah, 2000) detalham o PWM para GEV, sugerindo os seguintes estimadores para a GEV:

$$\begin{aligned}
\hat{\mu} &= \hat{\beta}_0 + \frac{\hat{\Psi}}{\hat{\xi}} [1 - \Gamma(1 - \hat{\xi})] \\
\hat{\Psi} &= \frac{(2\hat{\beta}_1 - \hat{\beta}_0)\hat{\xi}}{\Gamma(1 - \hat{\xi}) \times (2^{\hat{\xi}} - 1)} \\
\hat{\xi} &= -78,590c - 2,9554c^2
\end{aligned} \tag{4.1}$$

Sendo:

$$\begin{aligned}
\hat{\beta}_r &= (r+1)^{-1} \left\{ \mu - \frac{\Psi}{\xi} \left[1 - \frac{\Gamma(1-\xi)}{(1+r)^{-\xi}} \right] \right\}; \quad \xi < 1, \quad \xi \neq 0 \\
c &= \frac{(2\hat{\beta}_1 - \hat{\beta}_0)}{3\hat{\beta}_2 - \hat{\beta}_0} \frac{\ln 2}{\ln 3}
\end{aligned} \tag{4.2}$$

O PWM para a GPD produz os seguintes estimadores de forma e escala:

$$\begin{aligned}
\hat{\xi} &= 2 - \frac{\hat{\beta}_0}{\hat{\beta}_0 - 2\hat{\beta}_1} \\
\hat{\Psi} &= \frac{2\hat{\beta}_0\hat{\beta}_1}{\hat{\beta}_0 - 2\hat{\beta}_1}
\end{aligned} \tag{4.3}$$

A estimação por máxima verossimilhança é uma maneira mais popular de se estimar parâmetros. Neste caso, uma função de verossimilhança para um conjunto de n observações independentes é estimada como $L(\theta) = \prod_{j=1}^n L_j(\theta)$, sendo preciso encontrar o parâmetro que maximiza esta função de verossimilhança.

Segundo Kotz e Nadarajah (2000), para uma GEV,

$$L(\theta) = \sum [-\log(\psi) - (1-\theta)y_i - \exp(-y_i)] \quad (4.4)$$

A estimação do parâmetro de forma ξ , como destacado anteriormente, é de fundamental importância no estudo dos extremos, já que indica uma conformação da cauda da distribuição, especialmente se for utilizada uma distribuição pertencente ao MDA da GEV. O parâmetro de forma está relacionado ao parâmetro $\gamma = \frac{1}{\xi}$, denominado índice de cauda.

Hill (1975) propôs o seguinte estimador:

$$\hat{\gamma}_n = \frac{1}{k} \sum_{i=1}^k [\log(X_{n-i+1}) - \log(X_{n-k})] \quad (4.5)$$

Portanto, $\hat{\xi} = \hat{\gamma}^{-1}$

4.2 Teste de Qualidade de Ajuste

É importante avaliar a qualidade do ajuste entre os dados e a distribuição teórica verificando se o modelo é apropriado. Há testes formais e gráficos que devem ser utilizados em tal avaliação.

Os testes gráficos de qualidade de ajuste mais populares são os QQ – Plots e PP – Plots.

Dado um conjunto de dados a serem analisados, inicia-se o teste gráfico a partir da construção de um histograma e um ou mais box-plot, um plot para a função de distribuição empírica e outro para a função de distribuição teórica. O problema, então, é encontrar um função de distribuição F que seja um bom modelo para variáveis aleatórias i.i.d X_1, \dots, X_n .

Conforme detalhado em Embrechts et al (1997), considere X_1, \dots, X_n i.i.d com função de distribuição F . Seja U_1, \dots, U_n variáveis aleatórias uniformemente distribuídas em $[0,1]$ e denote-se $U_{n,n} < \dots < U_{1,n}$ a correspondente estatística de ordem. Então, espera-se o seguinte resultado:

$$F^{\leftarrow}(U_1) \stackrel{d}{=} X_1$$

$$\text{Para todo } n \in \mathbb{N}, (X_{1,n}, \dots, X_{n,n}) \stackrel{d}{=} (F^{\leftarrow}(U_{1,n}), \dots, F^{\leftarrow}(U_{n,n})).$$

$$(F(X_{k,n}))_{k=1, \dots, n} \stackrel{d}{=} (U_{k,n})_{k=1, \dots, n} \text{ em } [0,1], \text{ se e somente se } F \text{ é uma função contínua.}$$

Disto segue que:

$$EF(X_{k,n}) = \frac{n-k+1}{n+1}, \quad k = 1, \dots, n.$$

Note que $F_n(X_{k,n}) = \frac{n-k+1}{n}$, onde F_n suporta a função de distribuição

empírica de F . O gráfico $\left\{ \left(F(X_{k,n}), \frac{n-k+1}{n+1} \right) : k = 1, \dots, n \right\}$ é chamado de *probability plot*

(PP – Plot).

Também é comum plotar o gráfico $\left\{ \left(X_{k,n}, F^{\left(\frac{n-k+1}{n+1} \right)} \right) : k = 1, \dots, n \right\}$, chamado de *quantile plot* (QQ – Plot).

Observa-se que, em ambos os casos, há aproximação linearizada dos plots que podem ser justificadas por teoremas que utilizam como base a teoria de convergência fraca de processos empíricos. A discussão destes teoremas não é escopo deste trabalho.

Há, ainda, uma série de variantes dos gráficos acima descritos. Uma delas, que será utilizada ao longo deste estudo, é detalhada abaixo:

PP – Plot:

$$\{(F(X_{k,n}), p_{k,n}) : k = 1, \dots, n\}$$

QQ – Plot:

$$\{(X_{k,n}, F^{\left(p_{k,n} \right)}) : k = 1, \dots, n\}$$

Onde:

$$p_{k,n} = \frac{k - 0,5}{n} \quad (4.6)$$

é chamada posição de plotagem.

Em resumo, as observações são classificadas em ordem crescente e encontra-se a CDF ajustada e os respectivos quantis em comparação com a CDF empírica através de uma fórmula de plotagem adequada à distribuição analisada. A verificação se o modelo é bom é feita facilmente observando se os pontos plotados estão muito próximos à linha de 45° no gráfico.

Em relação aos testes formais, detalham-se a seguir os testes de Kolmogorov-Smirnov (K–S) e Crámer – von Mises (C–M), que são baseados em testes de adequação entre a distribuição empírica e a ajustada.

Seja $x_1 < \dots < x_n$ a estatística de ordem para uma amostra de tamanho n de uma população definida por uma CDF contínua $F(x)$.

$F_0(x; \theta)$ é uma família específica de modelos que contém uma função de parâmetros θ .

O teste de adequação do modelo é feito por meio de teste de hipóteses, onde a hipótese nula é:

$$H_0 : F(x) = F_0(x; \theta)$$

A estatística K–S para o teste de adequação do modelo é:

$$D_n = \max_{1 \leq k \leq n} (\delta_k) \tag{4.7}$$

$$\delta_k = |p_{k,n} - F_0(x; \theta)|$$

Este teste se baseia na distância máxima observada entre as duas funções. Os valores significativos e críticos para a estatística de Kolmogorov – Smirnov são apresentados na tabela abaixo:

Tabela 4.1 - Valores significativos e críticos para a estatística K-S

Valor Crítico	Nível de Significância
$\frac{1,07}{\sqrt{n}}$	20%
$\frac{1,22}{\sqrt{n}}$	10%
$\frac{1,36}{\sqrt{n}}$	5%
$\frac{1,63}{\sqrt{n}}$	1%

Fonte: Cruz (2002)

A estatística C-M é representada por:

$$W_n^2 = \sum (p_{k,n} - F_0(x; \theta))^2 + \frac{1}{12n} \quad (4.8)$$

Trata-se de uma medida do desvio quadrado médio da distância entre os dados e o modelo ajustado, acrescido de uma correção baseada no tamanho da amostra. Os valores significativos e críticos para a estatística de Crámer – von Mises são apresentados na tabela a seguir.

Tabela 4.2 - Valores significativos e críticos para a estatística C-M

Valor Crítico	Nível de Significância
1,222	5%
0,338	1%

Fonte: Cruz (2002)

CAPÍTULO 5

MODELAGEM DE FREQUÊNCIA E PERDAS AGREGADAS

5.1 Análise da Frequência dos Eventos

Nos tópicos anteriores foram discutidas as técnicas de modelagem de severidade de perdas operacionais com aplicação da EVT.

Em modelagem de risco operacional também é necessário preocupar-se com a frequência dos eventos de risco operacional. Conforme relata Cruz (2002), as perdas operacionais seguem processo estocástico discreto, ou seja, podem ser contados num determinado período. Daí, pode-se dizer que um determinado evento operacional acontece n vezes por dia.

Portanto, em gerenciamento de riscos operacionais é fundamental possuir estatísticas agregadas sobre o número de eventos e o total das perdas durante um determinado período de tempo, em vez do impacto e *timing* exatos de um evento em particular. Para resolver problemas assim são necessários modelos de contagem estocásticos que nos informem sobre a distribuição do número cumulativo de eventos em um período de tempo.

Apesar de existirem técnicas matemáticas para lidar com frequência de eventos similares a EVT usada para severidade das perdas, o ajuste de distribuições

de frequência ordinárias, como Poisson, Geométrica e Binomial Negativa, cujas funções de distribuição são descritas a seguir, costumam apresentar bons resultados.

Tabela 5.1 – Funções de Distribuição de Probabilidade – Processo de Contagem

Distribuição	Função Distribuição de Probabilidade
Poisson	$f(x) = \Pr(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$ $x \in \{0, 1, 2, \dots\}$
Geométrica	$f(x) = \Pr(X = x) = p(1-p)^x$ $x \in \{0, 1, 2, \dots\}$
Binomial Negativa	$f(x) = \Pr(X = x) = p^r \binom{r+x-1}{x} (1-p)^x$ $x \in \{0, 1, 2, \dots\}$

Fonte: Meyer (1983)

5.2 Seleção de Modelos

A seleção do melhor modelo pode ser feita apurando-se o critério de Akaike (AIC – *Akaike Information Criterion*) ou o critério de Bayesian (BIC – *Bayesian Information Criterion*). Para cada distribuição analisada, podem-se determinar os valores de AIC e BIC da seguinte forma:

$$AIC = -2 \cdot \sum_{i=1}^n \ln(f(x_i, \theta)) + 2 \cdot r \quad (5.1)$$

$$BIC = -2 \cdot \sum_{i=1}^n \ln(f(x_i, \theta)) + \ln(n) \cdot r \quad (5.2)$$

Onde:

$f(x, \theta)$ é a função de densidade de probabilidade;

n é o tamanho da amostra;

r é a quantidade de parâmetros da função de distribuição analisada.

Seleciona-se o modelo que apresenta um AIC ou BIC mais baixo. Em geral, os dois critérios concordam. Se houver discordância, o BIC tende a escolher o modelo com menos parâmetros.

Em geral, é possível observar visualmente a adequação de modelos quando se plotam o histograma dos dados empíricos e o ajuste de cada distribuição teórica analisada, que é dado por $n \times \text{PDF}$ ⁽²⁾.

5.3 Agregando Distribuição de Freqüência e Severidade

Tendo calculado separadamente os processos de severidade e de freqüência, o passo seguinte é combiná-los em uma distribuição de perdas agregada que permita prever um valor para as perdas operacionais com certo grau de confiança.

Formalmente, como detalhado por Bee, M (2005), pode-se obter resultados teóricos a partir da solução analítica de uma função de distribuição de perdas agregadas. Suponha que as perdas agregadas no tempo t são dadas por $S(t) = \sum_{i=1}^{K(t)} W_i$, $0 \leq t \leq T$, onde $K(t)$ é um processo de contagem e W_1, \dots, W_k são

² PDF é a função de distribuição de probabilidade, em inglês, *probability distribution function*.

variáveis aleatórias contínuas i.i.d, também independentes de $K(t)$, que representam as perdas operacionais individuais.

A CDF correspondente é:

$$\begin{aligned}
 F_{S(t)}(w) &= Pr(S(t) \leq w) \\
 &= \sum_{k=0}^{\infty} p_k Pr(S(t) \leq w | K(t) = k) \\
 &= \sum_{k=0}^{\infty} p_k Pr\left(\sum_{i=1}^{K(t)} W_i \leq w | K(t) = k\right) \\
 &= \sum_{k=0}^{\infty} p_k F_W^{*k}(w)
 \end{aligned} \tag{5.3}$$

Onde $p_k = Pr(K = k)$ e F_W^{*k} refere-se à k -ésima convolução de F_W consigo própria, ou seja, $F_W^{*k}(w) = Pr(W_1 + \dots + W_k \leq w)$ é a função de distribuição da soma de k variáveis independentes com a mesma distribuição que W .

A resolução da fórmula acima elimina o problema de agregação de severidade e frequência de perdas. No entanto, a sua solução analítica não é trivial. Em geral, a distribuição conjunta $f_{K,W}(k, w)$ das duas origens randômicas K e W não pode ser obtida analiticamente.

Uma alternativa é obter uma solução para agregação de frequência e severidade aplicando-se simulações de Monte Carlo, que é bastante utilizado para análise de riscos. Para tanto, é requerida a decomposição da distribuição conjunta de K e S : $f_{K,S}(k, s) = f_{S|K}(s | k) \cdot f_K(k)$. Condicionando $K = k$, o modelo para as

perdas agregadas é dado pela variável aleatória $(S | K = k) = \sum_{i=1}^K W_i$, que pode ser facilmente simulado da seguinte forma:

Se $W \sim \text{EVT}$ e $K \sim \text{Processo de Contagem (Poisson } (\lambda))$, por exemplo, esta distribuição é a convolução de k variáveis aleatórias que seguem uma EVT e a geração de números aleatórios de $f_{K,S}(k,s)$ pode ser realizada de acordo com seguintes passos, conforme indicado por Di Clemente e Romano (2003):

1. Gere uma variável aleatória k da distribuição de frequência escolhida;
2. Gere k números aleatórios W_1, \dots, W_k para a EVT;
3. Calcule o cenário de perdas, somando os k valores de severidades gerados no passo 2, ou seja, faça $S = \sum_{i=1}^K W_i$;
4. Repita os passos 1 – 3 B vezes, onde B é um grande número de vezes. Repetindo B vezes os 3 passos, tem-se uma amostra aleatória de tamanho B para a densidade $f_{K,S}(k,s)$.

CAPÍTULO 6

PROCEDIMENTOS METODOLÓGICOS

6.1 Série de Dados

Este trabalho baseia-se em aplicação prática da Teoria dos Valores Extremos em gerenciamento de risco operacional, tendo, portanto, uma abordagem quantitativa.

Conforme detalhado no Capítulo 2, utilizar-se-á duas séries de dados:

- Tipo 1 – Transações com cartão de débito – Terminais POS.
- Tipo 2 – Transações em terminais de auto-atendimento – TAA.

Referidas séries de dados estão na categoria de perdas com fraudes eletrônicas de um grande banco brasileiro, cujas perdas operacionais estão categorizadas com a seguinte distribuição média:

Tabela 6.1 – Categoria de Perdas – Distribuição

CATEGORIAS	
Processos Trabalhistas	41,84%
Fraudes Eletrônicas e Roubos Externos	18,02%
Fraudes Internas	4,68%
Falhas nos Negócios	23,05%
Danos ao Patrimônio Físico	5,35%
Falhas em Processos	6,86%
Falhas em Sistemas	0,21%

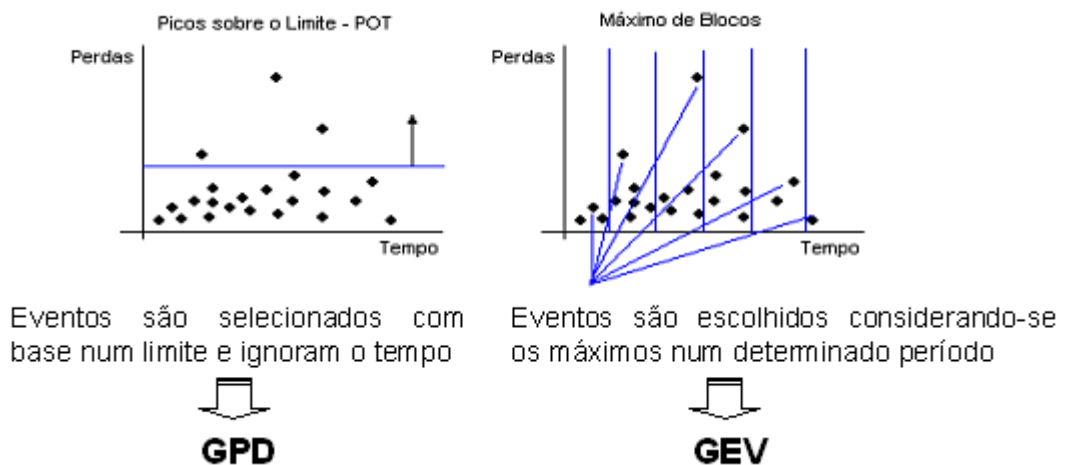
O período observado é de janeiro de 2004 a dezembro de 2005. Considerar-se-á como perdas, os valores das perdas brutas, deduzidos os valores recuperados, atualizados pela Taxa Média Selic – TMS até maio de 2006.

Tendo em vista o grau de sigilo das informações exploradas, será aplicado um fator de correção aos valores das perdas observadas.

6.2 Abordagem para Aplicação da EVT

Como pôde ser visto no Capítulo 3, há basicamente duas abordagens para análise da EVT. Uma delas baseia-se no método de ponto, onde se escolhe um limite, ignorando-se o momento em que os eventos ocorrem – abordagem POT – que em geral é ajustada pela GPD. A outra abordagem, que leva em conta o tempo, é a de máximo de blocos (eventos são escolhidos considerando-se os maiores em determinado período), que se ajusta à GEV.

Figura 6.1 – Abordagens EVT



Fonte: Cruz (2002)

Esta pesquisa enquadra-se na área de modelagem de risco operacional com vistas a melhor gerenciá-los. Neste caso, é importante entender a visão dos gerentes de riscos, que estarão preocupados em como alocar capital para fazer frente aos riscos operacionais. Dessa forma, a segunda abordagem pode ser mais apropriada, pois é difícil justificar a alocação de capital a partir de uma modelagem que ignora o momento das perdas, como no caso da abordagem POT.

6.3 Estimação de Parâmetros

Para análise da severidade das perdas, a estimação de parâmetros da distribuição de valores extremos será realizada utilizando o método dos momentos de probabilidade ponderada – PWM, conforme detalhado no Capítulo 4. A preferência dada ao PWM é devido a sua simplicidade e, conforme observaram Hosking et al (1985), por apresentarem vieses pequenos e que diminuem rapidamente à medida que a amostra aumenta.

Para análise da frequência de perdas, a estimação dos parâmetros será por meio do método de máxima verossimilhança, disponível no *software* utilizado (Matlab 7.0).

6.4 Teste de Qualidade dos Ajustes

6.4.1 Severidade

Os testes de verificação da qualidade dos ajustes terão como base os testes gráficos (PP – Plots e QQ – Plots) e os testes formais de Kolmogorov – Smirnov (K–S) e Crámer – von Mises (C–M), que foram detalhados no Capítulo 4.

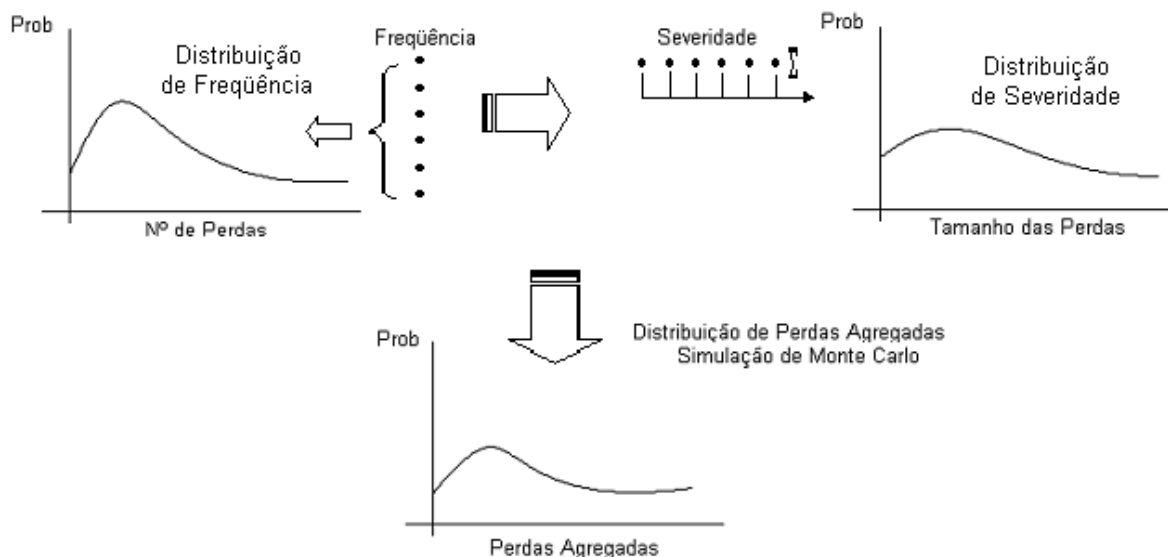
6.4.2 Freqüência

Para aferição dos modelos de contagem, a visualização gráfica já oferece bons resultados. No entanto, além dos testes gráficos serão calculados os testes de Akaike e Bayesian, conforme detalhado no Capítulo 5.

6.5 Procedimentos Finais

Após a modelagem da distribuição de severidade das perdas operacionais, procurar-se-á identificar uma distribuição de freqüência que se ajuste aos dados, para, em seguida, realizar simulações de Monte Carlo para combinar freqüência e severidade.

Figura 6.2 – Agregando Freqüência e Severidade



Os passos para realizar a simulação são descritos abaixo:

1. Gerar uma variável aleatória k da distribuição de frequência escolhida;
2. Gerar k números aleatórios W_1, \dots, W_k para a EVT;
3. Calcular o cenário de perdas e somar os k valores de severidades gerados no passo 2, ou seja, fazer $S = \sum_{i=1}^k W_i$;
4. Repetir os passos 1 a 3 B vezes, onde B é um grande número de vezes (50.000 nesta pesquisa). Repetindo B vezes os 3 passos, tem-se uma amostra aleatória de tamanho B para a densidade conjunta $f_{k,S}(k, s)$.

A partir dos resultados obtidos para S nas simulações, deve-se classificá-los em ordem decendente e verificar o valor máximo das perdas com determinado grau de confiança (VaR operacional).

Se a simulação é feita 50.000 vezes, o quinto maior valor de S representa o quantil com grau de confiança α igual a:

$$\alpha = 1 - \frac{5}{50.000} = 0,999 \quad \text{ou} \quad 99,99\%$$

O quinquagésimo valor de S , em ordem, terá α igual a:

$$\alpha = 1 - \frac{50}{50.000} = 0,9990 \quad \text{ou} \quad 99,90\%$$

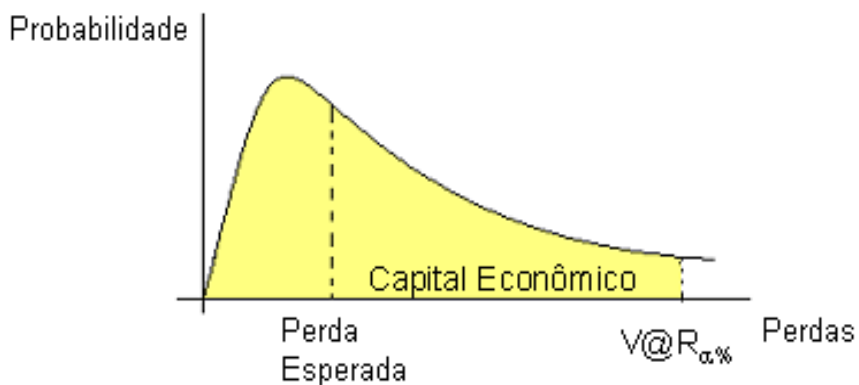
e assim sucessivamente.

Escolhido o intervalo de confiança, verifica-se o valor correspondente (VaR Operacional) e calcula-se o capital econômico necessário para fazer frente às perdas operacionais não esperadas, onde:

$$\text{Capital Econômico} = \text{VaR}_{\alpha} - \text{Perda Esperada} \quad (6.1)$$

Graficamente, tem-se:

Figura 6.3 – Capital Econômico



Os cálculos serão realizados individualmente para os dois bancos de dados de perdas operacionais, Tipo 1 e Tipo 2, observando-se os resultados univariados.

CAPÍTULO 7

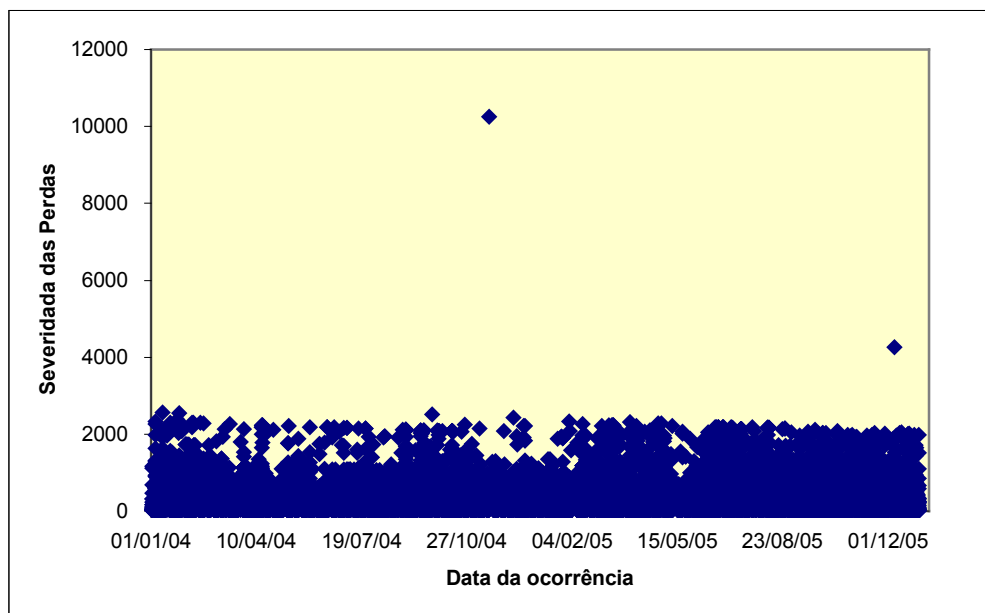
RESULTADOS OBTIDOS

7.1 Banco de Dados – Perdas Tipo 1

7.1.1 Análise de Severidade das Perdas

A partir do banco de dados dos eventos relacionados às perdas do Tipo 1, é plotado no Gráfico 7.1 todos as perdas observadas no período em estudo.

Gráfico 7.1 – Severidade – Perdas Tipo 1



Como se pode observar, há uma elevada quantidade de dados na amostra, dificultando a sua visualização. Para um melhor entendimento dos dados, apresenta-se a seguir uma tabela com as estatísticas descritivas da amostra.

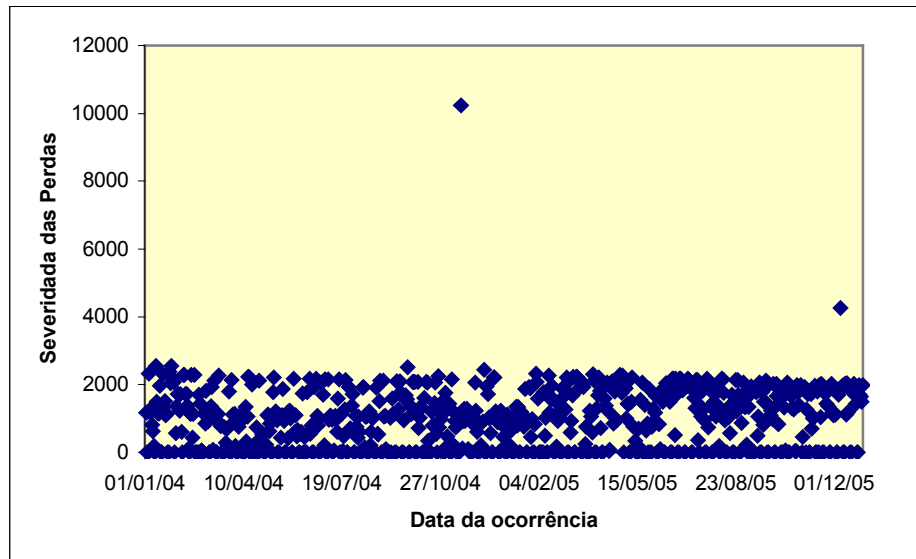
Tabela 7.1 – Estatística Descritiva – Severidade – Perdas Tipo 1

Estatística Descritiva	
Tamanho da Amostra	14.949
Média	304,54
Mediana	128,34
Desvio Padrão	426,58
Variância	181.970
Curtose	28,326
Assimetria	3,092
Mínimo	0,01
Máximo	10.247

Como dito anteriormente, a amostra tem tamanho expressivo (cerca de 15 mil ocorrências no período de dois anos). Há poucos valores acima de 4.000 (dois valores somente) e alguns acima de 2.000. A massa crítica de perdas é inferior a 2.000. Isto dificulta a modelagem quando se consideram distribuições tradicionais e abre o caminho para análise de distribuições de valores extremos, principalmente quando a preocupação é com eventos com baixa probabilidade de ocorrência, mas com alta severidade, ou seja, com a cauda da distribuição.

Na seqüência, para modelar as distribuições de valores extremos foi utilizada a abordagem de máximo de blocos, tendo sido necessário criar outro banco de dados, chamado aqui de secundário, contendo os máximos de determinado período. Optou nesta pesquisa pelo bloco de 1 dia. Portanto, o gráfico a seguir mostra os máximos deste bloco.

Gráfico 7.2 – Máximo de Blocos – Perdas Tipo 1



Esta série de dados contém 318 informações de perdas e será utilizada para modelagem de uma distribuição Generalizada de Valor Extremo – GEV, conforme detalhamento nos Capítulos 3 e 4. As Tabelas 7.2 e 7.3 trazem cálculos auxiliares para apuração dos parâmetros e as estimações dos parâmetros, respectivamente.

Tabela 7.2 – Estimação de Momentos e Constante c – GEV – Perdas Tipo 1

Momentos	Estimativas
Momento ₁	723,86
Momento ₂	568,61
Momento ₃	466,23
c	-0,018385

Tabela 7.3 – Estimação de Parâmetros – GEV – Perdas Tipo 1

Parâmetros	Estimativas
Forma (ξ)	0,14349
Localização (μ)	343,88
Escala (ψ)	512,65

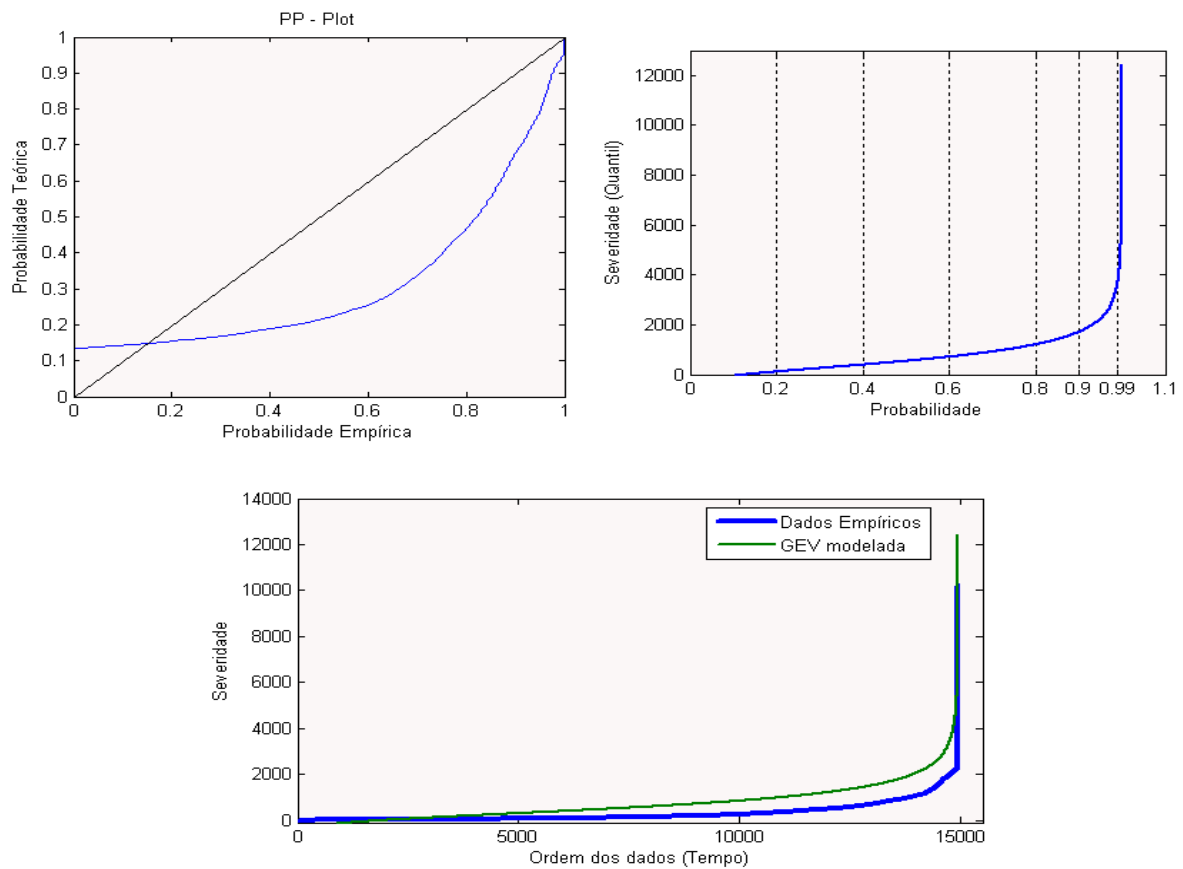
Como $\xi > 0$, temos uma distribuição de extremos Tipo 2 ou de Fréchet. Pode-se escrever a função H de acordo com a equação 3.3:

$$H_{\xi, \mu, \psi}(x) = \exp \left[- \left(1 + \xi \frac{x - \mu}{\psi} \right)^{-\frac{1}{\xi}} \right] \quad \text{onde} \quad \left(1 + \xi \frac{x - \mu}{\psi} \right) > 0$$

$$H_{\xi, \mu, \psi}(x) = \exp \left[- \left(1 + 0,14349 \frac{x - 343,88}{512,65} \right)^{-0,14349} \right] \quad \text{onde} \quad \left(1 + 0,14349 \frac{x - 343,88}{512,65} \right) > 0$$

A seguir são plotados para a distribuição GEV modelada, o PP-Plot, a probabilidade versus o quantil teórico e a comparação entre os quantis da GEV e os dados empíricos:

Gráfico 7.3 – Distribuição GEV – Análise Comparativa

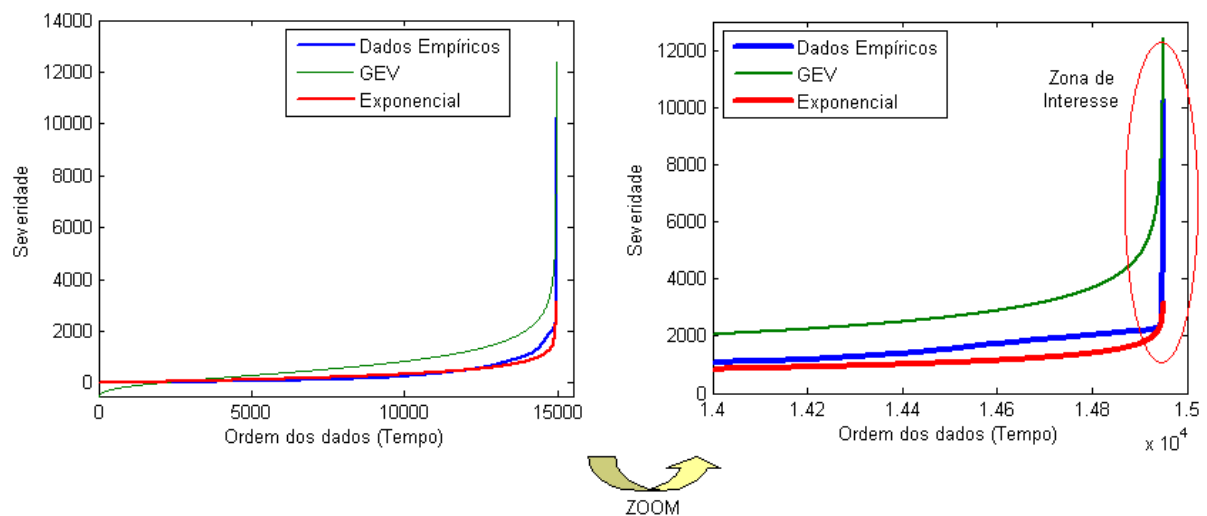


Visualizando os gráficos, pode-se observar que não há um bom ajuste para todo o domínio dos dados, pois para as baixas probabilidades ou mesmo no meio da distribuição ocorrem desvios elevados. Os testes formais K–S e C–M confirmam isto, pois nestes testes a hipótese nula é rejeitada. Entretanto, na cauda da distribuição, verifica-se que a aderência é significativa.

Comparando-se com outros tipos de distribuições, observa-se resultado semelhante, ou seja, testes formais também não demonstram adequação do modelo do ponto de vista estatístico, sendo que algumas têm melhor performance no interior da distribuição em comparação com a GEV, mas não na cauda, região de interesse.

Para exemplificar tal assertiva, apresenta-se no gráfico a seguir uma comparação entre a GEV, os dados empíricos e uma distribuição exponencial com parâmetro $\lambda = 304,54$.

Gráfico 7.4 – Análise Comparativa – GEV versus Exponencial



A maior confiabilidade da GEV para o propósito desta pesquisa fica evidente quando se está interessado em quantis que representam grau de confiança de 99,9% ou 99,99%. Comparativamente, tem-se:

Tabela 7.4 – Quantis – Análise Comparativa

Grau de Confiança	Quantil		
	Empírico	GEV	Exponencial
99,00%	2.026,0	3.684,2	1.402,5
99,90%	4.264,8	6.327,0	2.103,7
99,99%	10.247,0	10.167,0	2.804,9

Devido à característica assintótica da GEV, quantis mais elevados devem ser evitados, pois apresentariam valores muito conservadores (superestimados).

7.1.2 Análise de Frequência das Perdas

A partir do banco de dados dos eventos relacionados às perdas do Tipo 1, apresenta-se no Gráfico 7.5 o histograma com o número de ocorrências por dia e na Tabela 7.5 a estatística descritiva da amostra.

Gráfico 7.5 – Histograma – Frequência – Perdas Tipo 1

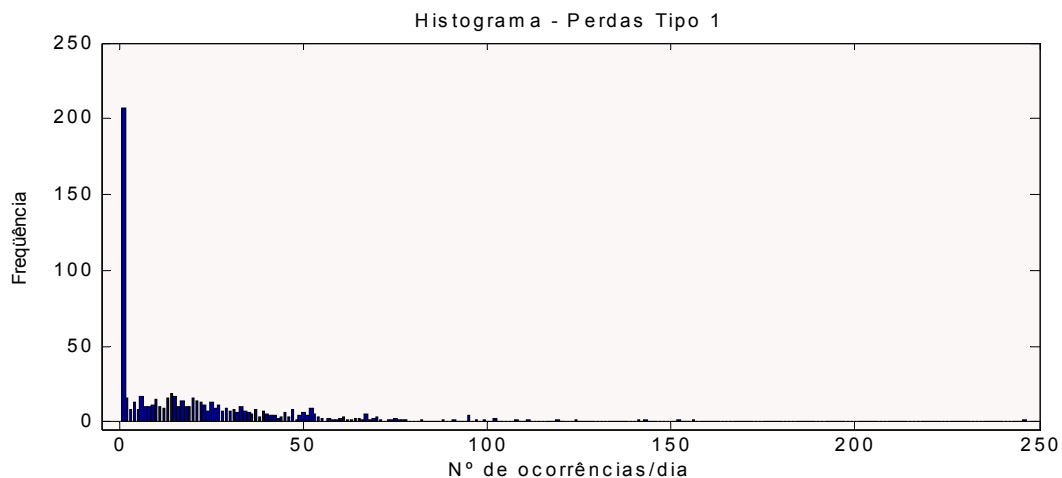


Tabela 7.5 – Estatística Descritiva – Freqüência – Perdas Tipo 1

Estatística Descritiva	
Tamanho da Amostra	731
Média	20,145
Mediana	13
Desvio Padrão	25,169
Variância	633,50
Assimetria	2,6077
Curtose	15,426
Mínimo	0
Máximo	245

Foram escolhidas três distribuições para modelar a freqüência: Poisson, Geométrica e Binomial Negativa.

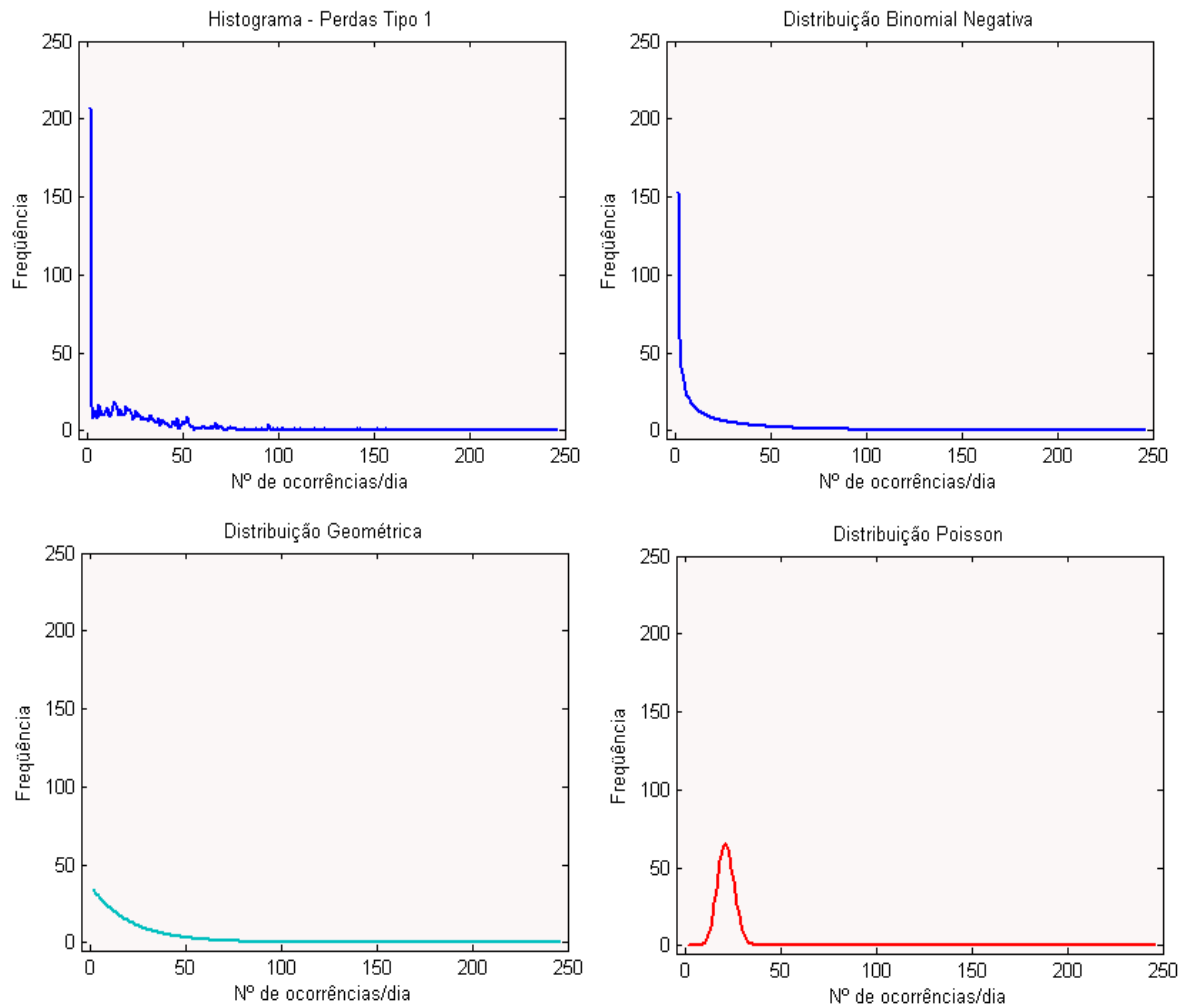
Utilizando o *software* Matlab 7.0, foram estimados os parâmetros de cada distribuição. Os resultados estão na tabela a seguir:

Tabela 7.6 – Estimação de Parâmetros – Freqüência – Perdas Tipo 1

Distribuições de Freqüência	Parâmetros
Poisson	$\lambda = 20,145$
Geométrica	$p = 0,047292$
Binomial Negativa	$r = 0,39725$ e $p = 0,019338$

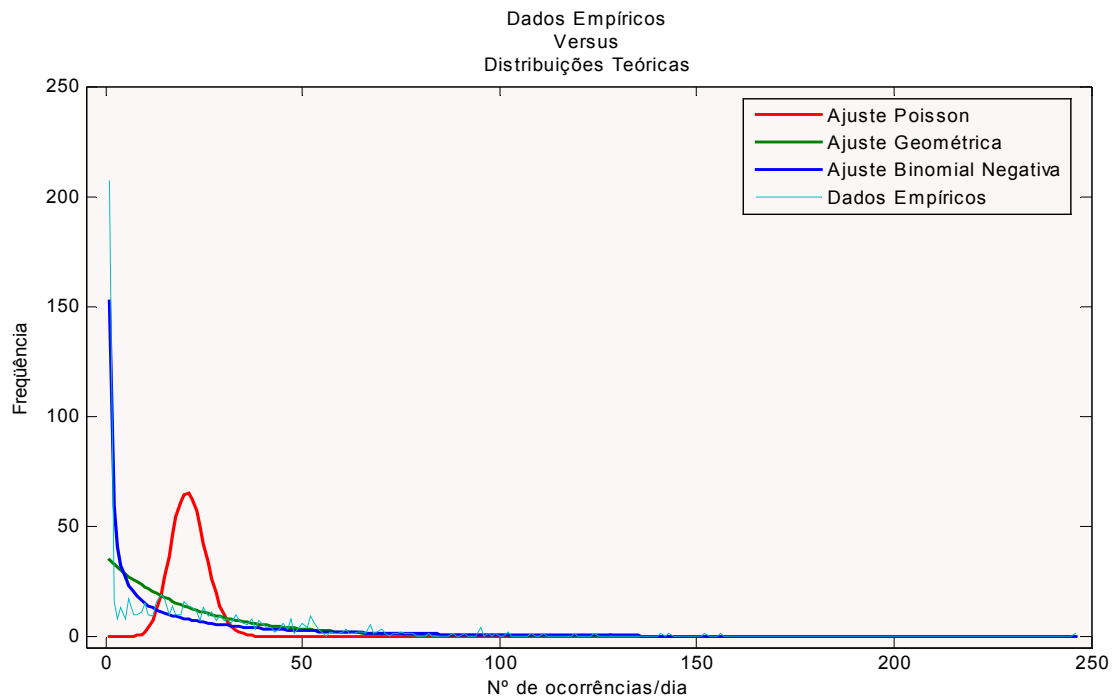
De forma a comparar os dados empíricos à modelagem com as três distribuições, apresenta-se a seguir gráficos onde se pode verificar o comportamento de cada distribuição teórica.

Gráfico 7.6 – Análise Comparativa – Distribuições de Frequência – Perdas Tipo 1



Plotando-se todo o conjunto de curvas, como no gráfico a seguir, sugere-se que, das três distribuições analisadas, a Poisson é que menos se adequa aos dados empíricos. As outras duas distribuições, Binomial Negativa e Geométrica, têm boa aderência aos dados.

Gráfico 7.7 – Análise Consolidada – Distribuições de Frequência – Perdas Tipo 1



A fim de selecionar o melhor modelo, para cada distribuição analisada foram apurados os valores dos testes Akaike (AIC) e Bayesian (BIC), conforme tabela abaixo:

Tabela 7.7 – AIC e BIC – Frequência – Perdas Tipo 1

Distribuições de Frequência	AIC	BIC
Poisson	22.179	22.184
Geométrica	5.890	5.895
Binomial Negativa	5.578	5.587

Seleciona-se o modelo que apresenta um AIC ou BIC mais baixo. Ambos os testes, demonstram que a distribuição que melhor se ajusta aos dados empíricos é a Binomial Negativa.

7.1.3 Agregação de modelos de severidade e frequência

Para agregar os modelos de severidade (7.1.1) e frequência (7.1.2), seguindo a descrição do Capítulo 6, foi utilizado o algoritmo detalhado no Apêndice a ser aplicado no *software* Matlab 7.0.

Na sequência, o vetor de perdas “Perdas_evt” (conforme descrito no Apêndice) foi classificado em ordem decrescente para se identificar o VaR Operacional (calculado os percentis). Calculou-se, preliminarmente, a perda esperada de “Perdas_evt”, onde $PE = E(\text{Perdas_evt}) = 256.756,69$, além do Capital Econômico, conforme tabela abaixo:

Tabela 7.8 – VaR Operacional e Capital Econômico – Perdas Tipo 1

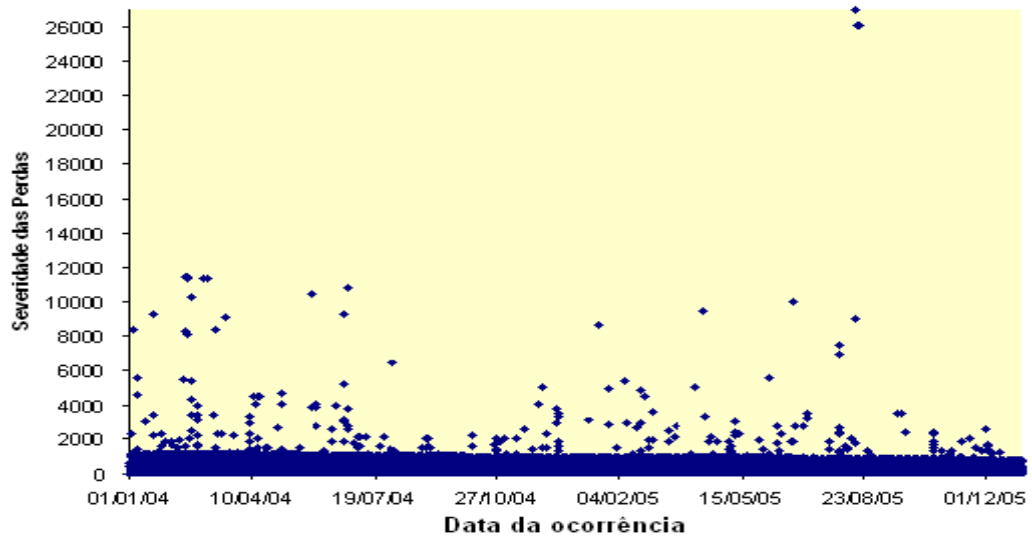
Grau de Confiança	VaR Operacional	Capital Econômico
99,00%	383.910,00	127.153,31
99,90%	402.160,00	145.403,31
99,99%	410.020,00	153.263,31

7.2 Banco de Dados – Perdas Tipo 2

7.2.1 Análise de Severidade das Perdas

A partir do banco de dados dos eventos relacionados às perdas do Tipo 2, é plotado no Gráfico 7.8 todas as perdas observadas no período em estudo.

Gráfico 7.8 – Severidade – Perdas Tipo 2



De forma semelhante ao observado na análise anterior (perda Tipo 1), há uma elevada quantidade de dados na amostra, dificultando sua visualização. Para melhor visualizar tais dados, apresenta-se a seguir uma tabela com as estatísticas descritivas desta amostra.

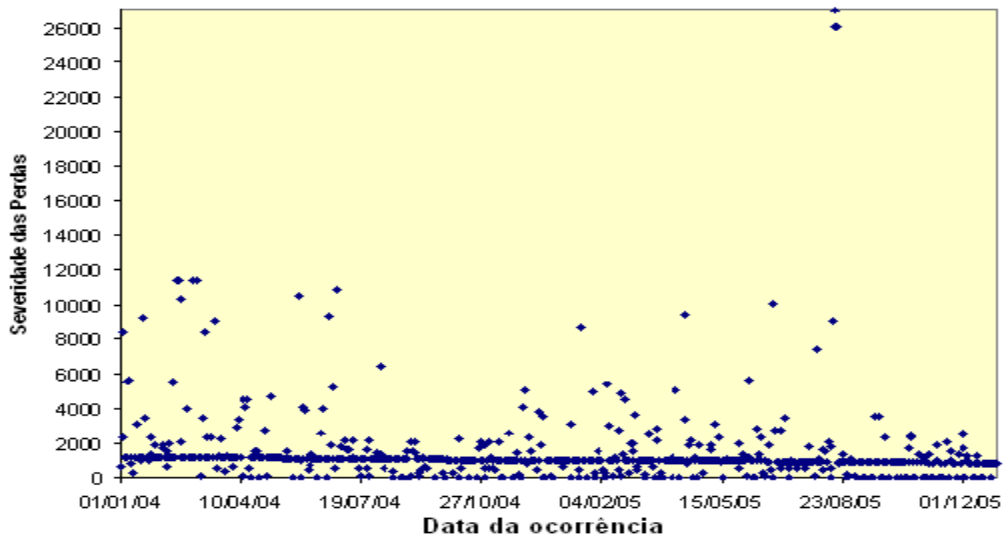
Tabela 7.9 – Estatística Descritiva – Severidade – Perdas Tipo 2

Estatística Descritiva	
Tamanho da Amostra	27.734
Média	550,42
Mediana	498,46
Desvio Padrão	588,95
Variância	346.870
Assimetria	13,247
Curtose	462,55
Mínimo	0,01
Máximo	26.986

Esta amostra é ainda maior que a anterior, com 27,7 mil eventos no período de dois anos. Há poucos valores acima de 10.000 e a massa crítica de perdas é inferior a 2.000.

Para modelar as distribuições de valores extremos foi utilizada a abordagem de máximo de blocos, tendo sido necessário criar outro banco de dados, como descrito anteriormente no caso de perdas Tipo 1. Neste caso, também optou-se pelo bloco de 1 dia. Portanto, o gráfico a seguir mostra os máximos deste bloco.

Gráfico 7.9 – Máximo de Blocos – Perdas Tipo 2



Esta série de dados contém 650 informações de perdas e será utilizada para modelagem de uma distribuição Generalizada de Valor Extremo – GEV, conforme detalhamento nos Capítulos 3 e 4. A Tabela 7.10 e 7.11 trazem cálculos auxiliares para apuração dos parâmetros e as estimações dos parâmetros, respectivamente.

Tabela 7.10 – Estimação de Momentos e Constante c – GEV – Perdas Tipo 2

Momentos	Estimativa
Momento ₁	1.167,6
Momento ₂	971,01
Momento ₃	857,45
c	-0,079624

Tabela 7.11 – Estimação de Parâmetros – GEV – Perdas Tipo 2

Parâmetros	Estimativa
Forma (ξ)	0,60703
Localização (μ)	342,5
Escala (ψ)	397,85

Neste caso $\xi > 0$, como no caso anterior. Logo, também se tem uma distribuição de extremos Tipo 2 ou de Fréchet. Pode-se escrever a função H como abaixo:

$$H_{\xi, \mu, \psi}(x) = \exp \left[- \left(1 + \xi \frac{x - \mu}{\psi} \right)^{-\frac{1}{\xi}} \right] \quad \text{onde} \quad \left(1 + \xi \frac{x - \mu}{\psi} \right) > 0$$

$$H_{\xi, \mu, \psi}(x) = \exp \left[- \left(1 + 0,60703 \frac{x - 342,5}{397,85} \right)^{-\frac{1}{0,60703}} \right] \quad \text{onde} \quad \left(1 + 0,60703 \frac{x - 342,5}{397,85} \right) > 0$$

7.2.2 Análise de Frequência das Perdas

A partir do banco de dados dos eventos relacionados às perdas do Tipo 2, apresenta-se no Gráfico 7.10 o histograma com o número de ocorrências por dia e na Tabela 7.12 a estatística descritiva da amostra.

Gráfico 7.10 – Histograma – Freqüência – Perdas Tipo 2

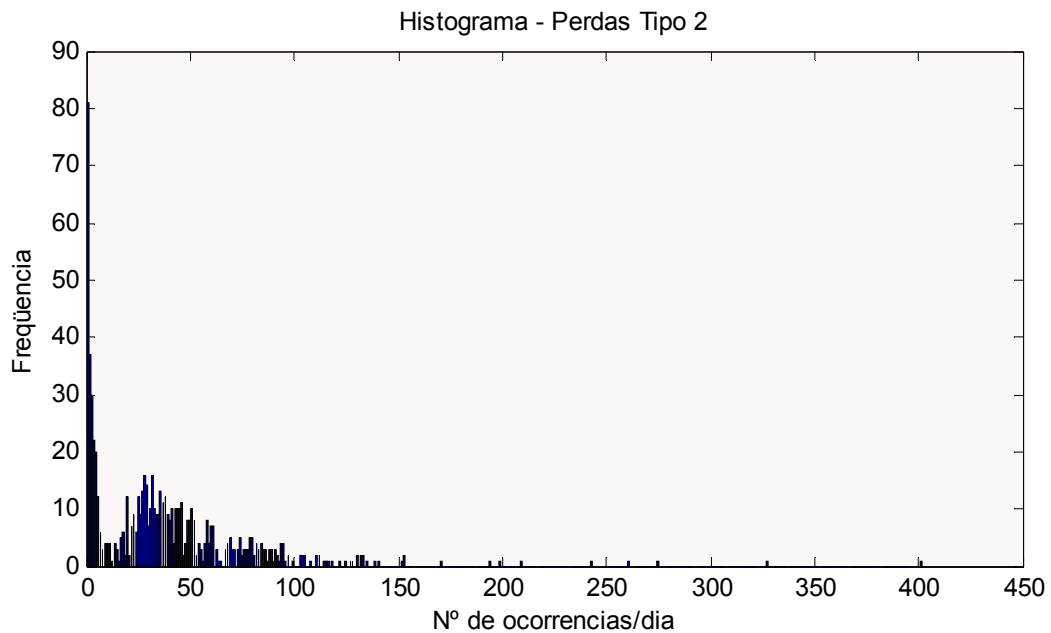


Tabela 7.12 – Estatística Descritiva – Freqüência – Perdas Tipo 2

Estatística Descritiva	
Tamanho da Amostra	731
Média	37,51
Mediana	31
Desvio Padrão	40,506
Variância	1.640,8
Curtose	2,844
Mínimo	0
Máximo	400

Foram escolhidas as mesmas três distribuições anteriores para modelar a freqüência: Poisson, Geométrica e Binomial Negativa.

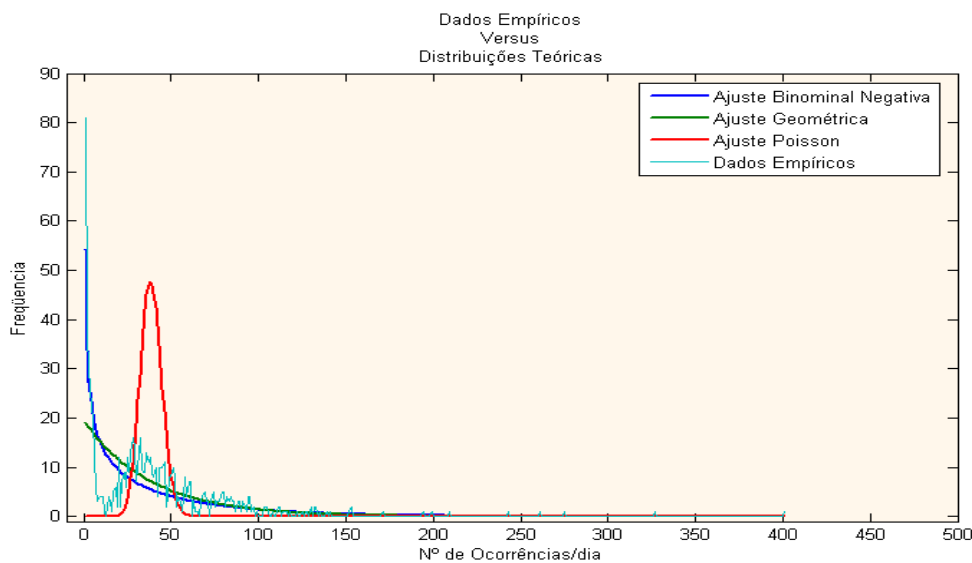
Utilizando o *software* Matlab 7.0, foram estimados os parâmetros de cada distribuição. Os resultados estão na tabela a seguir:

Tabela 7.13 – Estimação de Parâmetros – Freqüência – Perdas Tipo 2

Distribuições de Freqüência	Parâmetros
Poisson	$\lambda = 37,51$
Geométrica	$p = 0,025967$
Binomial Negativa	$r = 0,63558$ e $p = 0,016662$

Plotando-se todo o conjunto de curvas, como no gráfico a seguir, é possível verificar o comportamento de cada distribuição teórica. Visualmente, observa-se que se repetiu o resultado das perdas Tipo 1, ou seja, das três distribuições analisadas, a Poisson é que menos se adequa aos dados empíricos. As outras duas distribuições, Binomial Negativa e Geométrica, têm boa aderência aos dados.

Gráfico 7.11 – Análise Consolidada – Distribuições de Freqüência – Perdas Tipo 2



A fim de selecionar o melhor modelo, para cada distribuição analisada foram apurados os valores dos testes Akaike (AIC) e Bayesian (BIC), conforme tabela abaixo:

Tabela 7.14 – AIC e BIC – Freqüência – Perdas Tipo 2

Distribuições de Freqüência	AIC	BIC
Poisson	30.721	30.725
Geométrica	6.782,5	6.787,1
Binomial Negativa	6.699,2	6.708,3

A distribuição que melhor se ajusta aos dados empíricos é a Binomial Negativa, mesmo resultado da avaliação das perdas Tipo 1.

7.2.3 Agregação de modelos de severidade e freqüência

Para agregar os modelos de severidade (7.2.1) e freqüência (7.2.2), foi utilizado o algoritmo e descrições posteriores apresentados no Apêndice e item 7.1.3.

Os seguintes resultados foram obtidos:

$$PE = E(\text{Perdas}_{\text{evt}}) = 692.963,78$$

Tabela 7.15 – VaR Operacional e Capital Econômico – Perdas Tipo 2

Grau de Confiança	VaR Operacional	Capital Econômico
99,00%	1.109.400,00	416.436,22
99,90%	1.767.000,00	1.074.036,22
99,99%	3.267.400,00	2.574.436,22

CAPÍTULO 8

CONCLUSÕES

Recentemente, as instituições financeiras passaram a se preocupar em estudar e adotar abordagens sofisticadas para quantificar a adequada alocação de capital para fazer frente às perdas operacionais, em conformidade com o que o Novo Acordo de Capitais de Basileia (2004) chamou de Método de Mensuração Avançada – AMA. A importância dada ao gerenciamento de riscos operacionais é crescente. Com isso, a tendência é haver maior refinamento das práticas de gestão, com alocação de capital mais eficiente, o que propiciará alinhamento entre o gerencial e o regulatório.

Esta pesquisa mostra que a análise de eventos extremos é um dos caminhos para auxiliar uma boa gestão de risco operacional. A aplicação da EVT proporcionou resultados bastante satisfatórios.

Destaque-se que a preocupação de gerentes de risco é com eventos de baixa probabilidade que possam resultar em alta severidade de perdas para a instituição financeira. Isto incentivou esta pesquisa, que se concentrou na modelagem estocástica para a parte extrema da cauda da distribuição de perdas operacionais.

Na pesquisa, foram discutidos os conceitos da Teoria de Valores Extremos – EVT, sua modelagem matemática para diversos tipos de distribuições e aplicações

gerais, bem como se demonstrou a sua importância em gerenciamento de riscos operacionais, tendo sido priorizadas as técnicas associadas à distribuição Generalizada de Valor Extremo – GEV, cujo modelo melhor se adequa à abordagem baseada em máximos coletados em blocos. Isto permitiu fazer inferências mais precisas sobre as caudas e parâmetros da distribuição subjacente $H_{\xi, \mu, \psi}(x)$.

A aplicação de EVT na modelagem estocástica de risco operacional nesta pesquisa foi feita sobre bancos de dados de fraudes eletrônicas. A estimação de parâmetros resultou em parâmetro de forma (ξ) superior a zero nas duas séries de dados analisadas. Portanto, a distribuição de extremos modelada é do tipo Fréchet.

Os testes formais de Kolmogorov – Smirnov e Crámer – von Mises não certificaram bom ajuste, pois levam em consideração todo o domínio da distribuição. No entanto, quando se foca a atenção na cauda da distribuição, os resultados discutidos no Capítulo 7 evidenciaram uma boa aderência do modelo EVT, com performance na cauda melhor que outras distribuições ordinárias. Como na pesquisa interessa-se por quantis que representam grau de confiabilidade acima de 99,0%, ficou evidenciada a adequabilidade da modelagem EVT.

Na modelagem de frequência, como era esperado, as distribuições de contagem tradicionais mostraram bom ajuste aos dados empíricos. A distribuição binomial negativa se comportou bem e teve ajuste superior a outras distribuições como Poisson, bastante popular, e Geométrica, em ambas as séries de dados analisadas.

Após analisar separadamente os processos de frequência e severidade, foi utilizado um processo de simulação de Monte Carlo para agregá-los. Ressalte-se que quanto maior o número de rodadas na simulação, maior estabilidade terá a distribuição de perdas agregadas. A escolha de 50 mil rodadas é satisfatória para que haja consistência nos resultados.

Com a simulação de Monte Carlo, o que se está fazendo na prática é um tipo de análise de cenário estruturado. Cada rodada representa um dia. Nos próximos 50 mil dias, considerando a experiência passada, com, em média, k eventos por dia (seguindo uma binomial negativa) e um padrão de perdas que segue uma distribuição de eventos extremos (GEV: Tipo 2 – Fréchet), poder-se-ia prever quão severas seriam as perdas operacionais com certo grau de confiança.

A partir da simulação, chegou-se ao VaR operacional e à apuração do capital econômico a ser alocado para a instituição financeira se proteger contra perdas operacionais extremas.

Ressalte-se que o VaR operacional e o capital econômico são calculados para um dia. No Brasil, a periodicidade de revisão ainda não está definida pelo regulador, mas é provável que siga a sistemática já adotada para risco de mercado, ou seja, calcula-se o VaR diariamente e informa-se ao regulador, no entanto, apenas a informação do último dia do mês é utilizada para alocação de capital, sendo válida para todo o mês. Há modelos internos de risco de mercado que prevêem cálculo de VaR e capital econômico para dez dias que já estão sendo acompanhados pelo regulador, mas ainda não estão sendo efetivamente aplicados.

Dada a apenas recente preocupação com risco operacional, há vários caminhos a se trilhar em busca das melhores técnicas para gerenciamento destes riscos. Para exemplificar, observa-se que esta pesquisa modelou série de dados de forma univariada. A justificativa para, neste momento, não se preocupar com a estrutura de dependência dos tipos de perdas, o que levaria a uma análise multivariada, é que cerca de 65% das perdas operacionais de uma grande instituição financeira vem de demandas trabalhistas e fraudes, que são categorias que não devem apresentar nenhuma, ou quase nenhuma, interdependência. Outras categorias que possuem dependência entre si ou com as categorias citadas não são representativas e a relação custo/benefício de seu estudo e implementação pode não ser compensatório. No entanto, o estudo da estrutura de dependência é importante e mostra-se como campo fértil para futuras pesquisas, onde se poderia agregar o estudo de distribuições de valores extremos à análise de dependência das séries de perdas.

Outro interessante campo para pesquisa é o estudo de modelagem multifatorial em risco operacional. O fato de se ter uma cifra de VaR operacional é primordial para a adequada alocação de capital e a EVT é útil para este objetivo. No entanto, para gerenciar os riscos pode ser necessário um modelo que suplemente a análise do VaR, buscando fatores manipuláveis que expliquem as perdas agregadas e que meça a sensibilidade das perdas a mudanças nestes fatores. Portanto, o estudo de modelos causais também se configura como sugestão para outras pesquisas em risco operacional.

REFERÊNCIAS BIBLIOGRÁFICAS

Basel Committee - Basel II: International Convergence of Capital Measurement and Capital Standards: a Revised Framework, June 2004.

Basel Committee on Banking Supervision. Operational Risk, January 2001.

Bee, M. Copula-based Multivariate Models with Applications to Risk Management and Insurance. University of Trento, 2005.

Chavez-Demoulin, V., Embrechts, P. *Advanced Extremal Models for Operational Risk*, 2004.

Cruz, M. *Modeling, Measuring and Hedging Operational Risk*. John Wiley Trade, 2002.

Di Clemente, A., Romano, C. A Copula-Extreme Value Theory Approach for Modelling Operational Risk, 2003

Embrechts, P., Kluppelberg, C., Mikosch, T. *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag, New York, 1997.

Embrechts, P., Lindskog, F., McNeil, A. Modelling Dependence with Copulas and Applications to Risk Management, 2001.

Embrechts, P., McNeil, A., Straumann, D. Correlation and Dependence in Risk Management: Properties and Pitfalls, 1999.

Gumbel, E.J. *Statistics of Extremes*. Columbia University Press. New York, 1958.

Hill, B. A Simple General Approach to Inference About the Tail of a Distribution. *Annals of Statistics*, 3, 1163 – 1173, 1975.

Hosking, J. R. M.; Wallis, J & Wood, E. *Estimation of Generalized Extreme Value Distribution by the Method of Probability Weighted Moments*. *Technometrics*, 27, 251 – 261, 1985.

Jorion, P. *Value-at-Risk: The New Benchmark for Managing Financial Risk*. McGraw-Hill, New York, 2nd edition, 2000.

Kotz, S. Nadajarah, S.. *Extreme Value Distributions: Theory and Applications*. Imperial College Press, London, 2000.

Marshall, Christopher. *Measuring and Managing Operational Risks in Financial Institutions*. John Wiley Trade, 2001.

McNeil, A. J.. *Extreme Value Theory for Risk Managers*. Zurich, 1999.

Medova, E.A. *Measuring Risk by Extreme Values*. Operational Risk Special Report, Risk Magazine, November 2000, 20-25.

Medova, E.A. *Operational Risk Capital Allocation and Integration of Risks*. Advances in Operational Risk: Firmwide Issues for Financial Institutions, RiskBooks, 2001.

Medova, E.A.; Kyriacou, M. N.. *Extremes in Operational Risk Management*. Risk Management: Value at Risk and Beyond. MAH Dempster (ed). Cambridge University Press, 2002.

Meyer, Paul L. *Probabilidade – Aplicações à Estatística*. 2^a Edição. Editora LTC, Rio de Janeiro, 1983.

Pickands, J. *Statistical Inference Using Extreme Order Statistics*. The Annals of Statistics, 3, 119 – 131, 1975.

APÊNDICE

```

%-----Estimação de Parâmetros GEV-----%
load('C:\Mestrado UnB\Dissertação\sev_POS_maximo.txt');
n=size(sev_POS_maximo,1);
x=1:n;
Posicao_plotagem=(x-0.5)/n;
U0=Posicao_plotagem';
for i=1:n;
    XU0(i,1)=sev_POS_maximo(i,1)*U0(i,1);
    U1(i,1)=U0(i,1)^2;
    XU1(i,1)=sev_POS_maximo(i,1)*U1(i,1);
    U2(i,1)=U0(i,1)^3;
    XU2(i,1)=sev_POS_maximo(i,1)*U2(i,1);
end
M0=mean(XU0);
M1=mean(XU1);
M2=mean(XU2);
c=((2*M1-M0)/(3*M2-M0))-(log(2)/log(3));
Forma=-((7.8590*c+2.9554*c^2));
Escala=(2*M1-M0)*Forma/(gamma(1-Forma)*(2^Forma-1));
Localizacao=M0+(Escala/Forma)*(1-gamma(1-Forma));
%---Monte Carlo {Freq = Binomial Negativa; Sev = GEV}----%
for t=1:50000
    A_binneg(t,1)=nbinrnd(rbinneg,pbinneg);
    if A_binneg(t,1)==0;
        Perdas_evt(t,1)=0;
    else for q=1:A_binneg(t,1)
        prob(1,q)=rand(1,1,1);
        A_evt_frechet(1,q)=Localizacao-(Escala/Forma)*(1-(-
            log(prob(1,q)))^(-Forma));
        Perdas_evt(t,1)=sum(A_evt_frechet);
    end
end
end
%-----%

```