

Natalia Raquel Pires Nava

Incorporação do elemento BDE na modelagem de Risco Operacional e seu impacto no VaR

Brasil
2015, Julho

Natalia Raquel Pires Nava

Incorporação do elemento BDE na modelagem de Risco Operacional e seu impacto no VaR

Métodos de incorporação da Base de Dados Externa à Modelagem Avançada de Risco Operacional e seu respectivo impacto no cálculo de alocação de capital.

Universidade de Brasília – UnB

Departamento de Estatística

Programa de Pós-Graduação

Orientador: Eduardo Yoshio Nakano

Brasil

2015, Julho

Natalia Raquel Pires Nava

Incorporação do elemento BDE na modelagem de Risco Operacional e seu impacto no VaR/ Natalia Raquel Pires Nava. – Brasil, 2015, Julho-
79 p. : il. (algumas color.) ; 30 cm.

Orientador: Eduardo Yoshio Nakano

Dissertação (Mestrado) – Universidade de Brasília – UnB
Departamento de Estatística
Programa de Pós-Graduação, 2015, Julho.

CDU xx:xxx:xxx.x

Natalia Raquel Pires Nava

Incorporação do elemento BDE na modelagem de Risco Operacional e seu impacto no VaR

Métodos de incorporação da Base de Dados Externa à Modelagem Avançada de Risco Operacional e seu respectivo impacto no cálculo de alocação de capital.

Trabalho aprovado. Brasil, julho de 2015:

Eduardo Yoshio Nakano
Orientador

Sergio Wechsler
Convidado 1

André Luiz Fernandes Cançado
Convidado 2

Brasil
2015, Julho

Dedico este trabalho a todos os meus amigos e familiares que torceram e me deram forças para que eu chegasse até aqui. Sou feliz por ter vocês no meu caminho.

Agradecimentos

Este trabalho não teria sido possível sem o apoio incondicional dos meus colegas de trabalho Jorge Mazzo e Luiz Avando, que contribuíram efetivamente para que eu ingressasse, tivesse tempo para cursar as disciplinas da pós-graduação e me dedicasse a escrever a dissertação, respectivamente. À eles e à empresa que me acolheu, meu muito obrigada! Serei sempre grata.

Agradeço também ao meu orientador, o professor Nakano, por acreditar no meu potencial e ingressar comigo nesse assunto tão novo e abrangente que é a modelagem do risco operacional. Obrigada por dividir seus chocolates comigo e por me dar força nos momentos difíceis.

Ao meu marido, Vicente, pelo apoio e paciência, por não deixar que eu baixasse a cabeça nos momentos mais difíceis e por me dar sempre muito amor, eu tenho que agradecer todos os dias. Um dia nossos filhos irão saber o quanto você foi importante neste momento.

Por último, mas não menos importante, agradeço a meus pais, Carmita e Antonio, e a meus irmãos, Henrique e Diego, pela alegria com que encaram minhas aventuras e pelo orgulho que tem de mim. Por vocês eu chegaria à Lua.

“Dispensar sempre as fórmulas sociais criadas ou mantidas por convencionalismos ou tradições que estanquem o progresso. Toda complexidade atrasa o relógio da evolução.”
(Conduta Espírita. Waldo Vieira)

Resumo

O Banco Central do Brasil estabeleceu, por meio da Circular Nº 3.647, de Março de 2013, os requisitos mínimos para a utilização de Abordagem Avançada (AMA), baseada em modelo interno, no cálculo de alocação de capital para risco operacional. A incorporação da Base de Dados Externos (BDE) ao modelo AMA é um dos pontos abordados na referida Circular e será o foco deste trabalho. São sugeridos dois métodos de incorporação da BDE, em conjunto com modelos para escalonamento e filtragem dos dados, e avaliado o respectivo impacto do uso deste elemento na alocação de capital para risco operacional utilizando o método LDA (Loss Distribution Approach) para construção da Distribuição Agregada de Perdas (DAP) e cálculo do VaR (Value at Risk).

Palavras-chaves: Basileia. AMA. DAP. LDA. Escalonamento.

Abstract

Central Bank of Brazil has established, through Circular No. 3.647, of March 2013, the minimum requirements for the use of Advanced Measurement Approach (AMA), based on internal model, in the calculation of capital allocation for operational risk. The incorporation of External Database (BDE) to the AMA model is one of the points raised at that Circular and it was the focus of this work. We suggested two methods of incorporation of BDE, together with models for scaling and filtering data, and assessed the impact of the use of this element in capital allocation for operational risk using the Loss Distribution Approach (LDA) for construction of the Aggregate Losses Distribution (DAP) and calculating the Value at Risk (VaR).

Key-words: Basileia. AMA. DAP. LDA. Scalling.

Lista de ilustrações

Figura 1 – Distribuições de Probabilidade Estimadas - Categoria A	52
Figura 2 – Distribuições de Probabilidade Estimadas - Categoria B	52
Figura 3 – Distribuições de Probabilidade Estimadas - Categoria C	53
Figura 4 – Distribuição da BDE por Região	54
Figura 5 – Distribuição do Logaritmo das perdas da BDE	55
Figura 6 – Distribuições Agregadas de Perda considerando a incorporação da BDE para a Categoria A	60
Figura 7 – Distribuições Agregadas de Perda considerando a incorporação da BDE para a Categoria B	61
Figura 8 – Distribuições Agregadas de Perda considerando a incorporação da BDE para a Categoria C	63

Lista de tabelas

Tabela 1 – Linhas de Negócio	26
Tabela 2 – Descrição dos Tipos de Evento de Perda	28
Tabela 3 – Estatísticas Descritivas das Frequências e Severidades - Categorias A, B e C	53
Tabela 4 – Estimativas dos parâmetros dos modelos Poisson e Binomial Negativa .	54
Tabela 5 – Estimativas dos parâmetros dos modelos Weibull, Normal e Mistura de Normais	54
Tabela 6 – DQM dos ajustes das distribuições de frequência e logaritmo das perdas	54
Tabela 7 – Descritiva BDE - Categoria A	55
Tabela 8 – Descritiva BDE - Categoria B	56
Tabela 9 – Descritiva BDE - Categoria C	56
Tabela 10 – Estimativas do $VaR_{99,9}$ para Categoria A (R\$mil)	59
Tabela 11 – Estimativas do $VaR_{99,9}$ para Categoria B (R\$mil)	61
Tabela 12 – Estimativas do $VaR_{99,9}$ para Categoria C (R\$mil)	62

Lista de abreviaturas e siglas

AMA	Abordagem Avançada
ASA	Abordagem Padronizada Alternativa
BCBS	Comitê de Supervisão Bancária de Basileia
BDE	Base de dados externos
BDI	Base de dados internos
BEICF	Indicadores de Controle
BIA	Abordagem do Indicador Básico
BIS	Banco de Compensações Internacionais
CR	Categoria de Risco
DAP	Distribuição Agregada de Perdas
DQM	Desvio Quadrático Médio
EP	Erro Padrão
IMA	Internal Measurement Approach
IP	Intervalo Percentil
LDA	Loss Distribution Approach
LN	Linha de Negócio
MQO	Minimos Quadrados Ordinários
PE	Perda Esperada
PI	Perda Inesperada
POA	Perda Operacional Agregada
PRE	Patrimônio de Referência Exigido
RIF	Receita de Intermediação Financeira
RPS	Receita de Prestação de Serviço

TP	Tipo de Perda
TSA	Abordagem Padronizada
VaR	Value at Risk

Sumário

Introdução	23
1 Revisão de Literatura	25
1.1 Risco Operacional	25
1.1.1 Modelo BIA	26
1.1.2 Modelo ASA	26
1.1.3 Modelo AMA	27
1.1.3.1 BDI	29
1.1.3.2 BDE	30
1.1.3.3 BEICF	30
1.1.3.4 Cenários	31
1.2 Principais distribuições utilizadas na modelagem de Risco Operacional	31
1.2.1 Distribuições de Frequência	32
1.2.1.1 Distribuição Poisson	32
1.2.1.2 Distribuição Binomial Negativa	33
1.2.2 Distribuições de Severidade	33
1.2.2.1 Distribuição Weibull	33
1.2.2.2 Distribuição Log-normal	34
1.2.2.3 Mistura de Log-normal	34
2 Técnicas Estatísticas utilizadas na modelagem de Risco Operacional	37
2.1 LDA	37
2.1.1 Distribuição Agregada de Perdas	38
2.2 Medida de Qualidade de Ajuste	39
2.3 VaR	39
2.4 Bootstrap	40
3 Incorporação da BDE ao modelo AMA	43
3.1 De-Para	43
3.2 Filtros	44
3.3 Modelo de Escalonamento	45
3.4 Métodos de Incorporação	47
3.4.1 Método da União	47
3.4.2 Método da Interseção	48
4 Aplicações e Resultados	51

4.1	Dados internos	51
4.2	Dados Externos	54
4.2.1	Modelo de Escalonamento	56
4.3	Construção da DAP	58
4.3.1	Categoria A	58
4.3.2	Categoria B	59
4.3.3	Categoria C	60
Conclusão		65
Referências		67
Anexos		69
ANEXO A Códigos de Programação em R		71
A.1	Ajuste de Distribuições	71
A.1.1	Ajuste de Frequencia	71
A.1.2	Ajuste de Severidade	72
A.2	Escalonamento e Filtro BDE	75
A.3	LDA	76
A.3.1	BDI	76
A.3.2	$BDI \cup BDE$	77
A.3.3	$BDI \cap BDE$	78

Introdução

O Comitê de Supervisão Bancária de Basileia (BCBS), ligado ao Banco de Compensações Internacionais (BIS), estabeleceu em 1988 o primeiro Acordo de Capital de Basileia com o objetivo de criar exigências mínimas de capital para cobertura do risco de crédito incorrido pelas instituições financeiras. Em 1996, o Comitê incorporou ao Acordo a exigência de capital para risco de mercado, além de abrir a possibilidade da utilização de modelos internos para cálculo de alocação de capital - que, até então só era possível por modelos padronizados indicados pelo Comitê.

Em 2004 foi divulgada uma revisão do Acordo, conhecido como Acordo de Basileia II (Basileia, 2004), com o objetivo de tornar mais eficiente a Gestão de riscos das instituições financeiras, trazendo mais solidez e segurança para o sistema bancário internacional. Basileia II está estruturado em três Pilares mutuamente complementares, quais sejam:

I Requerimento de capital

II Supervisão Bancária

III Transparência e Disciplina de Mercado

No que tange o Pilar I, o documento aborda a importância dos modelos de alocação de capital para os Riscos de Mercado e de Crédito (já contemplados no documento de 1988) e adiciona o Risco Operacional no rol dos riscos incorridos pelas instituições financeiras que necessitam de reserva de capital como forma de salvaguarda-las de situações adversas.

Os Pilares II e III vem reforçar o papel do supervisor no que diz respeito a autonomia para validação de estrutura de gestão de riscos e o papel das instituições no compromisso de contribuir com boas práticas para manutenção de um sistema financeiro sólido e seguro, respectivamente. No Brasil, o Banco Central é o órgão que atua como regulador das instituições financeiras, visando a solidez do sistema financeiro nacional e internacional.

Ainda no âmbito do Pilar I, o Acordo de Basileia II propõe diferentes abordagens para o cálculo de alocação de capital para Risco Operacional, são elas, em ordem de complexidade: a Básica (BIA), a Padronizada (TSA) e a Avançada (AMA), sendo as duas primeiras baseadas em modelos padronizados e a AMA baseada em modelos internos.

A Abordagem Avançada dispõe de quatro elementos básicos, a saber: Base de Dados Internos (BDI), Base de Dados Externos (BDE), Indicadores de Controle (BEICF) e Análise de Cenários. O foco desta dissertação será a integração dos dados da BDE na

modelagem de dados de perda relativos a Risco Operacional utilizando modelos internos para cálculo de alocação de capital sob a Abordagem Avançada (AMA).

O trabalho está dividido em quatro capítulos. O primeiro capítulo trata das definições e requisitos regulatórios acerca do Risco Operacional, bem como da evolução dos modelos propostos pelo Comitê de Basileia. O segundo capítulo aborda técnicas estatísticas utilizadas na modelagem de alocação de capital para RO, com foco na construção da distribuição agregada de perdas utilizando o método Loss Distribution Approach (LDA). No capítulo seguinte, serão propostos métodos de incorporação dos dados externos de perda ao modelo, além da utilização de técnica de escalonamento de dados. Por último, serão apresentados os resultados obtidos a partir da aplicação das métricas abordadas neste trabalho, seguido das discussões e conclusões.

1 Revisão de Literatura

Este capítulo apresenta a revisão de literatura dos principais conceitos de risco operacional e a evolução do tema no que diz respeito a modelagem para alocação de capital.

1.1 Risco Operacional

O risco operacional é definido pelo Comitê de Basileia como o risco associado a possíveis perdas decorrentes de falhas, deficiências ou inadequação de processos internos, pessoas e sistemas, ou a eventos externos e deve incluir risco legal.

No Brasil, o tema ganhou notoriedade quando, em Junho de 2006, o Banco Central do Brasil (Bacen), atuando como regulador do sistema financeiro nacional, publicou a Resolução N°3380 ([Banco Central, 2006](#)) que dispõe sobre a implementação de estrutura de gerenciamento do risco operacional nas instituições financeiras. Neste documento, o Bacen estabeleceu as diretrizes que compõem a gestão deste risco, estabelecendo um prazo para a implementação de uma estrutura de gestão de riscos condizente com a natureza e complexidade das mesmas

Em 2008, o Bacen publicou a Circular N°3383 ([Banco Central, 2008](#)), estabelecendo os procedimentos para cálculo da parcela do Patrimônio de Referência Exigido (PRE) referente ao Risco Operacional (P_{OPR}) com base em três metodologias: Abordagem do Indicador Básico (BIA ¹), Abordagem Padronizada Alternativa (TSA ²) e Abordagem Padronizada Alternativa Simplificada (ASA ³), em grau crescente de sofisticação e complexidade.

Os modelos de alocação dispostos na Circular N°3383 atuam como proxy da exposição ao risco operacional, uma vez que não estão diretamente relacionados aos eventos de perda associados a este risco e sim ao resultado bruto (RB) da instituição, dado pela equação a seguir.

$$RB = RIF + RPS - DIF, \quad (1.1)$$

em que,

RB = Resultado Bruto

RIF = Receita de Intermediação Financeira

RPS = Receita de Prestação de Serviço

DIF = Despesa de Intermediação Financeira

¹ Basic Indicator Approach

² Trade Standardized Approach

³ Alternative Standardized Approach

1.1.1 Modelo BIA

A Abordagem do Indicador Básico, o modelo mais simples entre os listados pelo Comitê, sugere uma relação direta entre a alocação de capital e o RB da Instituição. O capital requerido é obtido a partir da multiplicação de um percentual fixo (15%), estabelecido pelo Comitê de Basileia, pelo RB. Assim, quanto maior o volume de negócios de uma instituição, teoricamente maior será sua exposição ao risco operacional e então, maior deve ser sua alocação de capital. Seja RB_t o Resultado Bruto da instituição no semestre t . O modelo BIA pode ser expresso como:

$$BIA = \sum_{t=1}^6 \frac{0.15 \times RB_t}{3} \quad (1.2)$$

1.1.2 Modelo ASA

A segunda Abordagem proposta, a TSA, leva em consideração as características inerentes às diferentes Linhas de Negócios (LN) estabelecidas na Circular N° 3.383, atribuindo pesos distintos que variam entre 0.12 e 0.18, de acordo com o risco associado a cada uma. Na tabela 1 estão dispostas as diferentes Linhas de Negócio e seus respectivos pesos no resultado final do cálculo de alocação de capital.

Tabela 1 – Linhas de Negócio

i	Linha de Negócio	Peso(α)
1	Varejo	0.12
2	Comercial	0.15
3	Finanças Corporativas	0.18
4	Negociação e Vendas	0.18
5	Pagamentos e Liquidações	0.18
6	Serviços de Agente Financeiro	0.15
7	Administração de Ativos	0.12
8	Corretagem de Varejo	0.12

A Abordagem Padronizada Alternativa Simplificada difere da TSA apenas na composição das Linhas de Negócio Varejo e Comercial, em lugar do RB, pode ser utilizado o Valor médio dos últimos três anos das carteiras ativas de empréstimos e adiantamentos referentes a essas linhas de negócio, multiplicadas por um fator de 0,035.

Seja $RB_{LN_{it}}$ o Resultado Bruto da Linha de Negócio i no ano t . O Modelo TSA (e o modelo ASA) pode ser expresso como:

$$TSA = \sum_{t=1}^3 \sum_{i=1}^8 \frac{\alpha_i \times RB_{LN_{it}}}{3}, \quad (1.3)$$

em que α_i é o peso da Linha de Negócio i .

Em 2013, a Circular N°3383 foi revogada, passando a valer as orientações para cálculo do capital requerido para o risco operacional estabelecidas na Circular N°3640 ([Banco Central, 2011](#)), contudo, os modelos estabelecidos para cálculo não foram alterados.

1.1.3 Modelo AMA

O grande marco de Basiléia II no que tange ao risco operacional, foi prever a alocação de capital utilizando modelos internos baseados nas perdas geradas pelo risco operacional propriamente dito. A Abordagem Avançada (AMA ⁴) foi uma evolução representativa para a indústria financeira porque permite que os bancos adaptem seus modelos de alocação de capital ao seu perfil de riscos, uma vez que o requerimento de capital deixa de ser uma proxy relacionada ao resultado e passa a refletir diretamente os controles internos e de gestão de riscos da instituição, influenciando inclusive a tomada de decisões por parte da alta administração.

No Brasil, a Circular N°3.647, do Banco Central ([Banco Central, 2013](#)), é o documento que regula os requerimentos mínimos a serem comprovados pelas instituições financeiras para que estas possam ser autorizadas a utilizar modelos internos para alocação de capital para cobertura do risco operacional.

Os modelos internos baseados na abordagem AMA, além da segregação das perdas por Linhas de Negócios, deverão permear também a classificação dos eventos em oito diferentes Tipos de Evento de Perda (TP), definidos na Resolução N°3380 do Bacen, com base em Basiléia II, conforme Tabela 2.

As classificações geradas a partir das combinações dos Tipos de Evento de Perda com as respectivas Linhas de Negócio formam as Categorias de Risco (CR) que nada mais são do que conjuntos de perdas que, por estarem ligadas a fatores geradores de risco comuns, podem ser agrupadas, para fins de modelagem.

Os valores da mensuração de risco de cada categoria de risco devem ser adicionados para fins do cálculo do valor da parcela de alocação de capital para risco operacional, podendo ser utilizadas medidas de dependência entre as categorias desde que as hipóteses de dependência sejam conservadoras e suportadas por evidências empíricas e julgamento de especialistas ([Banco Central, 2013](#)).

Para apuração do valor de capital a ser alocado para risco operacional, o Banco Central, em linha com o Comitê de Basiléia, estabeleceu que as instituições financeiras devem estimar de forma razoável as perdas inesperadas de cada CR combinando dados relevantes de perdas, tanto internas como externas, análise de cenários e indicadores de negócio e de controles internos que são específicos a instituição ([Basileia, 2004](#)).

⁴ Advanced Model Approach

Tabela 2 – Descrição dos Tipos de Evento de Perda

Tipo de Evento de Perda	Descrição
Fraude Interna	Perdas derivadas de ações de funcionários ou contratados com objetivos de fraude, apropriação de bens ou ação contra as leis ou políticas empresariais
Fraude Externa	Perdas derivadas de ações de terceiros com objetivos de fraude, apropriação de bens ou ação contra as leis
Demandas Trabalhistas	Perdas advindas de atuações incompatíveis da instituição com leis trabalhistas ou acordos de trabalho relacionados a segurança, higiene, respeito a diversidade, entre outros casos que envolvam funcionários ou contratados resultantes de reclamações (internas ou externas).
Falhas nos Negócios	Perdas decorrentes de ação involuntária ou negligente de uma obrigação profissional frente a clientes, de modelagem de processos ou da própria natureza de um produto.
Danos ao Patrimônio Físico	Perdas relacionadas a danos ou prejuízos a ativos materiais como consequências de desastres naturais ou outros acontecimentos
Falhas em Processos	Perdas decorrentes de erros no processamento de operações ou na gestão de processos, assim como na relação com contrapartes comerciais ou fornecedores
Falhas de Sistemas	Perdas derivadas de falhas em sistemas
Interrupção das Atividades	Perdas decorrentes da interrupção voluntária ou involuntária das atividades por parte da instituição

Para estimar as perdas inesperadas, o Regulador recomenda a construção de uma Distribuição Agregada de Perdas (DAP) de onde se possa calcular, com 99,9% de confiança, o valor de alocação de capital que deve conter, ainda, as perdas esperadas pela instituição.

Considerando que a alocação de capital deve ser calculada para cada CR, a DAP deve ser construída para cada uma destas considerando os quatro elementos base na modelagem do risco operacional via AMA, a saber: Base de Dados Interna (BDI), Base de Dados Externa (BDE), Indicadores de Negócios (BEICF) e Análise de Cenários, os quais estão relacionados a seguir.

1.1.3.1 BDI

A base de dados interna é o principal elemento do modelo AMA. É da BDI que são extraídas informações acerca das ocorrências de perdas operacionais sofridas pela instituição que servem de insumo para medidas de gestão internas das instituições financeiras. Pelo fato de fornecer estatísticas sobre os diferentes tipos de evento de perda da instituição de forma segura e bem estruturada, os ensaios de modelos internos para alocação de capital com base apenas na BDI está em fase avançada nas instituições financeiras.

A BDI inclui informações contábeis sobre os eventos de perda tais como data de ocorrência, descoberta e contabilização, local em que se verificou a perda e seu respectivo valor bruto e valor recuperado, devendo abranger um período mínimo de 5 anos de informações para utilização no modelo de alocação de capital.

O uso da BDI no modelo AMA deve seguir alguns critérios, entre eles:

1. Independência de eventos

Múltiplas perdas operacionais relacionadas a um mesmo evento devem ser agrupadas segundo critérios consistentes, o que envolve entendimento do real risco do evento ocorrer. O conceito está diretamente relacionado à concepção da Base de Cálculo - base que guarda as informações que serão utilizadas diretamente no cálculo de alocação de capital, com o valor da perda dado pelo saldo final dos registros contábeis de pagamentos, provisões e recuperações.

2. Suficiência de dados

Relacionado com a homogeneidade das categorias de risco, o critério de suficiência permite o agrupamento de dados com características comuns visando minimizar possíveis problemas relacionados a pequena quantidade de dados dentro das categorias de risco. Isto se dá porque a granularidade da informação que é encontrada ao agrupar os dados em 56 categorias de risco é inversamente proporcional a quantidade de dados presente em cada uma, o que pode vir a ser um problema para a modelagem.

3. Limite para registro de perdas operacionais na base

O uso de limite é facultativo, desde que não altere o valor de alocação de capital. Tal condição sugere que, estabelecida uma metodologia para cálculo de alocação de capital, sejam testados diferentes pontos de corte de modo a encontrar limite que não altere a distribuição dos dados e conseqüentemente o valor de alocação de capital.

Vale ressaltar que os dados da BDI são insumo para construção dos BEICF.

1.1.3.2 BDE

A base de dados externa guarda informações sobre ocorrência de perdas operacionais oriundas de fontes externas à instituição, obtidas de fontes públicas, privadas e de consórcios de dados, entre outros (Circular nº3.647).

Ao fazer parte de um consórcio de dados, a instituição financeira se beneficia de alguns aspectos, entre eles: classificação padronizada das perdas provenientes de instituições de países diferentes em Linhas de Negócio e Tipos de Evento de Perda; Valor piso para envio de informações (*threshold*), informações acerca da Região onde a perda ocorreu, porte da instituição em que está classificada a perda, descrição de perdas relevantes, entre outros que variam de acordo com o consórcio.

Por serem oriundas de fontes externas, estas perdas precisam ser escalonadas antes de serem incorporadas ao modelo, de modo que sejam compatíveis com a classificação de perdas da instituição e com o perfil de risco da mesma.

A BDE deve ser incorporada no modelo com os objetivos de agregar informações acerca de perdas severas e plausíveis, complementar os registros da BDI, contribuir na construção de cenários e na apuração do cálculo de capital.

O foco desta dissertação é incorporar este elemento ao modelo AMA. Para isto, os dados da BDE serão filtrados, escalonados e utilizados no cálculo de alocação de capital sob duas abordagens diferentes, sendo avaliado o impacto da incorporação destes dados no resultado final do modelo.

1.1.3.3 BEICF

Os indicadores de risco operacional podem ser divididos em dois grupos: indicadores de ambiente de negócios - referentes aos riscos inerentes às atividades bancárias e ao mercado no qual a instituição atua; e indicadores de controles internos - devem refletir a qualidade dos processos internos da instituição relacionados à prevenção e mitigação do risco operacional.

A incorporação dos indicadores ao modelo AMA objetiva ([Banco Central, 2013](#)):

1. refletir a qualidade dos controles da instituição e do ambiente de operações;
2. contribuir para a avaliação de necessidade de capital e para o gerenciamento do risco operacional;
3. agregar avaliações de caráter prospectivo do risco operacional; e
4. reconhecer a melhoria e a deterioração dos controles internos e do ambiente de negócios na avaliação da necessidade de capital para risco operacional.

Apesar de ser um elemento muito utilizado na gestão do risco operacional, existe pouca literatura sobre como incorporar as informações geradas pelos indicadores no cálculo de alocação de capital pelo AMA.

(Davies et al., 2006) aborda o assunto com discussões detalhadas sobre a identificação, especificação, seleção e implementação deste elemento na estrutura de gerenciamento de riscos das instituições.

1.1.3.4 Cenários

A análise de cenários tem como objetivo agregar ao modelo AMA eventos de risco operacional raros e de alta severidade, porém considerados plausíveis, de modo que se possa fornecer informações sobre o risco potencial a que a instituição está exposta.

Para construção dos cenários devem ser considerados os dados da BDI, BDE, BEICF, além do conhecimento de especialistas em riscos. Este elemento deve ainda incorporar a análise de impacto da ocorrência simultânea de múltiplos eventos de risco operacional e, assim como os demais elementos, contribuir para o cálculo da parcela de capital.

Um dos métodos utilizados na indústria bancária nacional é o método Delphi (Dalkey et al., 1969), que consiste na aplicação de questionários sobre temas específicos, direcionados a especialistas no assunto abordado, com o objetivo de avaliar a opinião dos mesmos em relação à probabilidade de ocorrência de um evento. O resultado é avaliado sob a ótica do grupo: havendo convergência de opinião entre os especialistas, pode-se utilizar o método proposto por Franco 2007 para calcular probabilidades de ocorrência simultânea de múltiplos eventos, calculadas a partir de métodos bayesianos, considerando probabilidades a priori.

(Scandizzo, 2006) aborda a construção e uso do elemento cenário no modelo AMA.

1.2 Principais distribuições utilizadas na modelagem de Risco Operacional

O Comitê de Basileia sugere duas metodologias para a construção da distribuição agregada de perdas a partir dos dados de risco operacional: a IMA (*Internal Measurement Approach*) e a LDA (*Loss Distribution Approach*). A primeira metodologia consiste em construir uma distribuição de probabilidade para as informações de perda ao longo do tempo. A segunda abordagem usa simulação de Monte Carlo para gerar, a partir das distribuições de frequência e severidade dos dados, novas observações possíveis para a distribuição agregada de perdas.

Por ser a mais utilizada na literatura bancária sobre o tema, este trabalho terá foco na metodologia LDA para construção da distribuição agregada de perdas, detalhada na Seção 2.1 deste trabalho.

Para isto, serão construídas distribuições empíricas e paramétricas para os dados de perda baseada nas funções de probabilidade mais adequadas para as características dos dados de perda operacional, conforme segue.

1.2.1 Distribuições de Frequência

Os dados de frequência podem ser definidos por uma variável aleatória discreta e positiva, denotada por N , que representa o número de vezes que determinado evento de perda ocorre em um determinado período de tempo.

Neste trabalho serão ajustadas as quantidades (contagens) de perdas mensais ocorridas dentro de cada categoria de risco (CR).

As distribuições mais utilizadas para modelar frequência dos dados de perdas operacionais são as de Poisson e Binomial Negativa.

1.2.1.1 Distribuição Poisson

A distribuição de Poisson é uma distribuição discreta que modela o número de ocorrência de um evento de interesse em um determinado intervalo de tempo e é uma distribuição muito utilizada na modelagem de dados de contagem.

Segundo Johnson e Kotz 1969, uma variável aleatória X tem distribuição Poisson com parâmetro λ , ($\lambda > 0$) se sua função de probabilidades for da forma:

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, 2, \dots \quad (1.4)$$

Para o ajuste do modelo segundo o método de Máxima Verossimilhança (Schwarz et al., 1978), considere que X_1, X_2, \dots, X_n é uma amostra aleatória de tamanho n da distribuição Poisson. O logaritmo da função de verossimilhança é dado por:

$$l(\lambda|\mathbf{X}) = -n\lambda + \log \lambda \sum_{i=1}^n x_i - \sum_{i=1}^n \log x_i! \quad (1.5)$$

em que $\lambda > 0$ é o parâmetro a ser estimado.

O estimador de máxima verossimilhança é obtido maximizando (1.5) com respeito a λ , que resulta em

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (1.6)$$

1.2.1.2 Distribuição Binomial Negativa

A distribuição Binomial Negativa, também conhecida como distribuição de Pascal, possui dois parâmetros e também é muito utilizada na modelagem de dados de contagem. Sua função de probabilidades pode ser dada por:

$$P[X = x] = \frac{\Gamma(x + r)}{\Gamma(r)x!} p^r (1 - p)^x, x = 0, 1, 2, \dots \quad (1.7)$$

com $r > 0$ e $0 < p < 1$, em que $\Gamma(a) = \int_a^\infty u^{a-1} e^{-u} du$ é a função Gama.

O logaritmo da função de verossimilhança do modelo Binomial Negativo é dado por

$$l(r, p | \mathbf{X}) = \sum_{i=1}^n \log \Gamma(x_i + r) - \sum_{i=1}^n \log x_i! - n \log \Gamma(r) + nr \log(p) + \log(1 - p) \sum_{i=1}^n x_i, \quad (1.8)$$

em que x_i é a i -ésima observação da amostra, $i = 1, 2, \dots, n$ e $r > 0$ e $0 < p < 1$ são os parâmetros a serem estimados.

Os estimadores de máxima verossimilhança são obtidos maximizando (1.8) com respeito a r e p . Esses estimadores não podem ser obtidos analiticamente. No entanto suas estimativas podem ser facilmente obtidas através de métodos de maximização numérica. Neste trabalho, os estimadores de máxima verossimilhança da distribuição Binomial Negativa foram obtidos através do comando *optim* do programa R (R Core Team, 2015).

1.2.2 Distribuições de Severidade

Dados de severidade representam o montante de perdas em valor monetário ocorridos na instituição. Diferente dos dados de frequência, os dados de severidade não tem relação com o tempo e podem ser definidos a partir de uma variável aleatória contínua, positiva definida, com característica assimétrica na cauda da distribuição.

Para modelagem dos dados de severidade, as distribuições mais utilizadas são a Log-Normal, Weibull e Mistura de Log-Normais.

1.2.2.1 Distribuição Weibull

Popular em aplicações na área de Análise de Sobrevivência, a distribuição Weibull (Weibull, 1951) é usualmente utilizada para modelar o tempo até a ocorrência de um evento de interesse. No entanto, por possuir um suporte positivo, essa distribuição também pode modelar outros tipos de dados com valores positivos. Sua função densidade de probabilidades é dada por:

$$f(x) = \alpha \lambda^\alpha x^{\alpha-1} e^{-\left(\frac{x}{\lambda}\right)^\alpha}, x > 0. \quad (1.9)$$

O logaritmo da função de verossimilhança do modelo Weibull é dado por:

$$l(\alpha, \lambda | \mathbf{X}) = n \log(\alpha) + n\alpha \log(\lambda) + (\alpha - 1) \sum_{i=1}^n \log x_i - \lambda^{-\alpha} \sum_{i=1}^n x_i^{\alpha}, \quad (1.10)$$

em que x_i é a i -ésima observação da amostra, $i = 1, 2, \dots, n$ e $\alpha > 0$ e $\lambda > 0$ são os parâmetros a serem estimados.

Os estimadores de máxima verossimilhança são obtidos maximizando (1.10) com respeito a α e λ . Esses estimadores não podem ser obtidos analiticamente. No entanto suas estimativas podem ser facilmente obtidas através de métodos de maximização numérica. Os estimadores de máxima verossimilhança da distribuição Weibull foram obtidos através do comando *optim* do programa R (R Core Team, 2015).

1.2.2.2 Distribuição Log-normal

Assim como a distribuição Weibull, uma outra distribuição de probabilidade bastante utilizada na modelagem de dados de sobrevivência é a distribuição log-normal. Uma variável aleatória X tem distribuição log-normal se $Y = \log(X)$ tem uma distribuição Normal. A função densidade de probabilidades do modelo log-normal é definida por:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp \left[- \left(\frac{\log x - \mu}{2\sigma^2} \right)^2 \right], x > 0. \quad (1.11)$$

O logaritmo da função de verossimilhança do modelo log-normal é dado por:

$$l(\mu, \sigma^2 | \mathbf{X}) = - \sum_{i=1}^n \log x_i - n \log \sigma - \frac{n}{2} \log(2\pi) - \frac{1}{2\sigma^2} \sum_{i=1}^n (\log x_i - \mu)^2, \quad (1.12)$$

em que x_i é a i -ésima observação da amostra, $i = 1, 2, \dots, n$ e $-\infty < \mu < \infty$ e $\sigma > 0$ são os parâmetros a serem estimados.

O estimador de máxima verossimilhança é obtido maximizando (1.12) com respeito a μ e σ , que resulta em

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n \log x_i$$

e

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (\log x_i - \hat{\mu})^2 \quad (1.13)$$

1.2.2.3 Mistura de Log-normal

Os modelos de mistura de distribuições são utilizados em estudos cuja população em estudo é formada por várias subpopulações (observáveis ou latentes). Situações como essa podem ser detectadas quando os dados apresentam um comportamento multimodal.

Considerando o caso contínuo, em que as densidades componentes de uma mistura pertencem a família paramétrica $f_j(\cdot)$, tem-se que a função densidade de probabilidades de uma mistura de k componentes é dada por:

$$f(x) = \sum_{j=1}^k p_j f_j(x), \quad (1.14)$$

em que $f_j(x)$ representam as j densidades componentes da mistura e p_j são os respectivos pesos, com $\sum_{j=1}^k p_j = 1$.

Em particular, quando $k = 2$ e $f_j(\cdot)$ pertence a família de distribuições log-normais, tem-se que a densidade de probabilidades de uma mistura de duas log-normais é dada por:

$$f(x) = \frac{p}{x\sigma_1\sqrt{2\pi}} \exp\left[-\frac{(\log x - \mu_1)^2}{2\sigma_1^2}\right] + \frac{1-p}{x\sigma_2\sqrt{2\pi}} \exp\left[-\frac{(\log x - \mu_2)^2}{2\sigma_2^2}\right], \quad (1.15)$$

em que p é o parâmetro de mistura e $-\infty < \mu_1, \mu_2 < \infty$, $\sigma_1 > 0$ e $\sigma_2 > 0$ são os parâmetros das distribuições log-normais.

O logaritmo da função de verossimilhança é dado por:

$$l(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, p | \mathbf{X}) = \sum_{i=1}^n \log \left\{ \frac{p}{x\sigma_1\sqrt{2\pi}} \exp\left[\frac{(\log x_i - \mu_1)^2}{2\sigma_1^2}\right] + \frac{1-p}{x\sigma_2\sqrt{2\pi}} \exp\left[\frac{(\log x_i - \mu_2)^2}{2\sigma_2^2}\right] \right\}. \quad (1.16)$$

Da mesma forma que em (1.8), os estimadores dos parâmetros de (1.16) não podem ser obtidos analiticamente, mas podem ser facilmente calculados através de métodos de maximização numérica. Os valores as estimativas da mistura de log-normais foram obtidos através do comando *optim* do programa R (R Core Team, 2015).

2 Técnicas Estatísticas utilizadas na modelagem de Risco Operacional

Neste segundo capítulo, será apresentado o método LDA (Loss Distribution Approach) e seu respectivo modelo de simulação para encontrar a Distribuição Agregada de Perdas (DAP), além da medida utilizada para cálculo de alocação de capital, conhecida como VaR e a métrica utilizada para avaliar a qualidade das estimativas.

2.1 LDA

O LDA (Loss Distribution Approach) é um método amplamente utilizado pela indústria financeira e tem como objetivo estimar uma função de distribuição agregada de perdas (DAP) que deve ser construída a partir das funções distribuição de frequência e severidade dos dados (Frachot et al., 2004).

Também utilizado nas abordagens de risco de crédito e de mercado, o LDA tornou-se popular por simular observações de perda de acordo com o perfil de riscos da instituição.

Considerando que as bases de dados com registros de perdas operacionais não guardam uma grande quantidade de observações, visto a recente movimentação das instituições financeiras motivadas por Basileia II, é natural a utilização do LDA especificamente por tornar possível a simulação de novos registros de perdas operacionais agregadas (POA) a partir dos dados fornecidos pela instituição, feitas por meio da técnica de simulação de Monte Carlo.

Por perda operacional agregada (POA) entende-se os valores da soma das severidades de perda em um determinado período de tempo, definida como:

$$POA = \sum_{k=1}^N S_k, \quad (2.1)$$

Aqui, S_k é o k -ésimo valor de perda ocorrido no tempo de referência, $k = 1, 2, \dots, N$ e N é o número de perdas ocorrido no tempo de referência. O LDA determina a distribuição da POA como na equação a seguir, em que $X = POA$ e $F(x)$ representa a função de distribuição de probabilidade da DAP (Diniz et al., 2014):

$$\begin{aligned} F(x) = P(X \leq x) &= \sum_{n=0}^{\infty} P(X \leq x \cap N = n) \\ &= \sum_{n=0}^{\infty} P(X \leq x | N = n) \times P(N = n) \end{aligned}$$

$$\begin{aligned}
&= \sum_{n=0}^{\infty} P\left(\sum_{k=0}^N S_k \leq x | N = n\right) \times P(N = n) \\
&= \sum_{n=0}^{\infty} P\left(\sum_{k=0}^n S_k \leq x | N = n\right) \times P(N = n), \quad (2.2)
\end{aligned}$$

com $S_0 = 0$.

A função de distribuição agregada de perdas é quem fornece o valor de alocação de capital para risco operacional, conhecido como VaR, abordado na Seção 2.3.

A metodologia IMA consiste em estimar diretamente a DAP por meio de MCMC (Markov Chain Monte Carlo) ou outro método numérico, porém, a baixa quantidade de observações não permite o uso de pressupostos assintóticos para as estimativas, o que abre espaço para a utilização do LDA.

Muitos autores utilizam o LDA para construção da DAP fazendo uso apenas da BDI, abordando outras questões de interesse como a correlação entre as categorias de risco (Requena, 2014). Neste trabalho, o LDA será utilizado para construir a DAP considerando a incorporação do elemento BDE ao AMA.

2.1.1 Distribuição Agregada de Perdas

Para construção da DAP como em (2.2), via LDA, se faz necessária a construção das distribuições de frequência e severidade dos dados de perda conforme Seção 1.2.

A premissa básica do LDA é a de que os dados de frequência e severidade são independentes, ou seja, o valor das perdas não tem relação com a quantidade de vezes que o evento ocorre. Assume-se ainda que os dados de severidade dentro de uma mesma categoria de risco são iid (independentes e identicamente distribuídos) e maiores que zero.

Tais premissas permitem a construção de valores de POA nunca antes observados mas que, partindo de distribuições de frequência e severidade ajustadas de acordo com o perfil de risco da instituição refletido nos dados utilizados para ajuste, podem acontecer a qualquer momento.

A construção da DAP, via LDA, segue o algoritmo a seguir:

Passo 1: Estime a função de probabilidade dos dados de contagem, $P(N = n)$;

Passo 2: Estime a função densidade de probabilidade dos dados de severidade das perdas, $f(s)$;

Passo 3: Sorteie uma observação da distribuição de frequência, $N = n$;

Passo 4: Sorteie n (valor da frequência observada no Passo 3) observações independentes da distribuição de severidade, S_1, S_2, \dots, S_n .

Passo 5: Some as observações de severidade para obter a POA $X = \sum_{k=1}^n S_k$

Passo 6: Repita os passos 3 a 5 um número m grande de vezes, de modo a obter a variável aleatória $X = x_1, x_2, \dots, x_k, \dots, x_m$, conforme 2.2

Para a execução dos Passos 1 e 2, recomenda-se testar mais de uma distribuição de probabilidade para avaliar qual melhor se ajusta aos dados. Comumente, seriam utilizados os testes de Qui-Quadrado (DeLong et al., 1988) e Kolmogorov-Smirnov (Smirnov, 1948) para avaliar a qualidade dos ajustes, porém, a grande quantidade de observações (no mínimo 60 para frequência e milhares para severidade) fornece estimativas viesadas dos valores dos testes. De fato, uma amostra com milhares de observações concederia poder suficiente para rejeitar qualquer ajuste. A próxima seção indica uma medida não-paramétrica alternativa para comparar a qualidade dos ajustes.

2.2 Medida de Qualidade de Ajuste

Para escolha da distribuição melhor ajustada aos dados de frequência e severidade propõe-se, em conjunto com análise gráfica da distribuição de probabilidade, o uso de estatística não-paramétrica relativa ao Desvio Quadrático Médio (DQM) dos dados de ajuste aos dados empíricos, dada por:

$$DQM = \sum_{i=1}^n \frac{(F_T(x_i) - F_E(x_i))^2}{n}, \quad (2.3)$$

em que $F_T(\cdot)$ é a função de distribuição estimada segundo o modelo teórico e $F_E(\cdot)$ é a função de distribuição empírica.

Esta medida avalia o viés médio das estimativas em relação as observações e quanto menor ela for, maior a proximidade da estimativa com os dados empíricos.

Neste trabalho, a medida DQM será utilizada para escolha das distribuições de frequência e severidade que serão utilizadas no LDA.

2.3 VaR

A estatística utilizada para estimar a necessidade de capital para risco operacional é conhecida como VaR.

O Valor em Risco (VaR) é uma medida amplamente utilizada no gerenciamento de riscos das instituições financeiras e representa o valor da pior perda esperada em um intervalo de tempo dado um nível de confiança (Jorion, 1996), satisfazendo a seguinte equação:

$$P(\text{Perda} \leq \text{VaR}_\alpha) = \alpha \quad (2.4)$$

Na abordagem AMA do risco operacional, o VaR deve medir a pior perda possível, a um nível de 99,9% de confiança, no intervalo de um ano. Sendo assim, o valor da perda que satisfaz a equação (2.4), considerando nível de confiança de 99,9%, calculado a partir da DAP, é o que determina a alocação de capital para risco operacional.

Neste contexto, vale ressaltar mais uma vez, a importância da simulação de novas observações para distribuição agregada de perdas, visto que calcular o VaR a partir das observações empíricas não extrapolaria a ocorrência de valores já esperados pela instituição, sendo que o valor de alocação de capital neste caso seria dado por, no máximo, o valor máximo de perda já observado. Além disto, a estimação precisa do percentil 99,9 requer uma quantidade razoável de dados, o que leva, mais uma vez, a preferência de uso do LDA.

Para avaliar a qualidade das estimativas do VaR geradas a partir do LDA, será utilizada a técnica de re-amostragem de Bootstrap a fim de calcular os valores de erro padrão e de outras medidas de qualidade do estimador.

2.4 Bootstrap

A ideia básica do método Bootstrap é de que, se não existe outra fonte sobre a distribuição dos dados, a amostra observada contém toda a informação acerca da distribuição e então a re-amostragem destes dados seria a melhor forma de encontrar o valor esperado do parâmetro de interesse (Efron and Tibshirani, 1994).

Cada re-amostra bootstrap i , onde $i = 1, 2, \dots, B$, deve conter n elementos gerados a partir dos dados observados, sorteados com reposição.

O sorteio com reposição pressupõe a independência dos dados e, a re-amostra obtida a partir deste sorteio, será composta por combinação de observações distinta das já observadas, gerando nova estimativa para a estatística de interesse.

Recomenda-se que sejam sorteados um número n grande de observações para compor cada re-amostra, sendo $B=100$ um número satisfatório de re-amostras quando n é suficientemente grande (Efron and Tibshirani, 1994).

Uma das avaliações de qualidade de estimadores que pode ser construída via Bootstrap, é a estimativa do erro padrão.

Considere $s(x^*)_i$ as estimativas da estatística de interesse, geradas a partir das réplicas de Bootstrap, calculadas a partir dos dados x_1, x_2, \dots, x_n .

O Erro Padrão das estimativas, utilizado para avaliar a variabilidade do estimador é dado por:

$$EP = \sqrt{\frac{1}{B-1} \times \sum_{i=1}^B (s(x^*)_i - s(\bar{x}^*))^2}, \quad (2.5)$$

em que $s(\bar{x}^*) = \frac{1}{B} \sum_{i=1}^B s(x^*)_i$ é a média das estimativas, que será utilizada, no caso de $s(x^*) = VaR_{99,9}$, como valor de alocação de capital para risco operacional.

Além do EP (2.5), pode-se calcular ainda o Intervalo Percentil (IP) do estimador, construído a partir das B réplicas da simulação Bootstrap como em (2.6).

Considere \hat{G}^* a função distribuição acumulada de $s(x^*)$.

O intervalo percentil, com coeficiente de confiança $1-2\alpha$, é definido pelos percentis α e $1-\alpha$ de \hat{G}^* :

$$IC_{(\alpha;1-\alpha)} = [\hat{G}^{-1}(\alpha), \hat{G}^{-1}(1-\alpha)] \quad (2.6)$$

As estimativas do EP (2.5) e IP (2.6) serão utilizadas para comparar as estimativas geradas para o VaR sob diferentes abordagens do LDA (considerando somente BDI ou incorporando a BDE).

Todas as técnicas apresentadas neste capítulo serão utilizadas, em conjunto com as abordagens do capítulo seguinte, para o cálculo de alocação de capital para risco operacional. Para isto, será utilizada base de dados real, devidamente descaracterizada, com informações de perdas ocorridas em um grande banco brasileiro.

O capítulo seguinte abordará a incorporação do elemento BDE ao modelo AMA, com proposta de técnicas para filtro, escalonamento e uso destes dados no método LDA para construção da DAP.

3 Incorporação da BDE ao modelo AMA

O Banco Central prevê alcançar os seguintes objetivos com a incorporação dos dados externos ao modelo AMA ([Banco Central, 2013](#)):

1. Agregar informações acerca de perdas externas severas e plausíveis;
2. Complementar os registros internos de perdas operacionais, reduzindo eventual escassez de dados;
3. Agregar informações ao processo de análise de cenários, de forma a contribuir para a avaliação prospectiva do grau de exposição da instituição ao risco operacional; e
4. Contribuir para a avaliação, mensuração e monitoramento da exposição ao risco operacional e para apuração do VaR.

A metodologia para a incorporação de dados externos ao modelo AMA deve ainda:

1. Garantir que os dados sejam compatíveis com a classificação de perdas operacionais da instituição; e
2. Prever o escalonamento, a filtragem e a adoção de ajustes qualitativos dos dados, de forma a adequá-los, sempre que possível, ao porte, à localização e ao perfil de risco da instituição, visando minimizar eventual subjetividade ou viés das fontes de informação e do processo de coleta.

Visando atender aos requerimentos do Regulador e contribuir com literatura acadêmica sobre o assunto, propõe-se aqui um processo robusto e estatisticamente fundamentado para tratamento e incorporação dos dados externos ao modelo de alocação de capital para risco operacional.

Para isso, serão abordadas metodologias para classificação ("De-Para"), filtragem e escalonamento dos dados externos, bem como duas formas de incorporação dos dados ao modelo, aqui descritos como Método da União e Método da Interseção.

3.1 De-Para

Para uso dos dados externos na gestão de riscos e cálculo de alocação de capital para risco operacional, um passo imprescindível é o processo de "De-Para", que se refere a classificação dos eventos externos nas categorias de perda internas da instituição financeira que irá utilizar a BDE.

A classificação é feita na marcação dos Tipos de Perda e Linhas de Negócio e este processo se faz necessário para minimizar alguns tipos de viés de informação, tais como viés de controle e viés de coleção referentes à diferença nas regulações de diferentes países e na forma como são armazenados os dados (no que se refere a data, correção de inflação, etc.), respectivamente.

Deve-se levar em consideração no momento da classificação do "De-Para" que existem determinados eventos, comuns a outros países, que não acontecem no Brasil ou, ainda, que são classificados em categorias de perda distintas das utilizadas no Brasil, seja pelo fato da instituição financeira que reportou a perda ter perfil de negócios diferente da instituição que irá utilizar a informação, seja porque o entendimento das instituições acerca da classificação de um determinado tipo de evento seja distinto.

A participação da instituição financeira em um consórcio de dados facilita a construção da marcação do "De-Para" pelo fato do consorcio fornecer uma pré-marcação padronizada de eventos de perda em categorias já definidas por Basiléia II, conforme Tabela 2. Esta marcação, porém, é feita em diferentes níveis de granularidade (I, II e III, geralmente), sendo que o "De-Para" deve ser feito no nível mais granular, justamente para ter a sensibilidade de marcar as observações de acordo com as atividades internas da instituição, evitando o viés de informação.

O processo de classificação é de uso interno de cada instituição e deve ser definido por analista de risco com conhecimentos sobre a base de dados e os tipos evento e linhas de negócio da instituição.

3.2 Filtros

Os filtros na BDE são utilizados com o objetivo de agregar à base interna apenas eventos prováveis de ocorrer, considerando o perfil de risco da instituição.

Assim, propõe-se que, após a classificação dos eventos externos em TP's e LN's internos, seja construída análise padronizada de dados para classificação de observações que são passíveis de exclusão (ou não-inclusão) na modelagem da BDE.

Para tanto, é indicada a construção de análise de dados via Boxplot ([Williamson et al., 1989](#)), onde são considerados *outliers* os pontos que ultrapassam o limite superior da distribuição de dados, definido como segue.

Seja X a variável aleatória que define os valores da perda da BDE.

$F(x)$ é a função distribuição acumulada de X .

No conceito utilizado para construção do Boxplot, o limite superior da distribuição dos dados é dado por:

$$LSup = F^{-1}(0, 75) + 1,5 \times [F^{-1}(0, 75) - F^{-1}(0, 25)] \quad (3.1)$$

Aqui, $F^{-1}(k)$ é o Quartil k da variável X .

Sendo assim, poderiam ser considerados para modelagem dos dados externos os dados da BDE que estivessem abaixo do limite superior do Boxplot, enquanto que os *outliers* poderiam ser utilizados para agregar informações ao processo de análise de cenários, de forma a contribuir para a avaliação prospectiva do risco operacional.

3.3 Modelo de Escalonamento

O escalonamento consiste no processo sistemático e estatisticamente fundamentado de ajuste do montante de perdas de eventos externos ao perfil de risco e às atividades da instituição.

O objetivo de escalonar os dados da BDE é ponderar aspectos específicos relacionados a cada evento de perda no valor desta, de modo a poder comparar observações que possuem ao menos um aspecto em comum e conseguir relacionar uma perda externa ao perfil de riscos interno de uma instituição.

O modelo de escalonamento pode ser construído com base nas seguintes premissas (Dahen and Dionne, 2010):

1. A perda pode ser dividida em dois componentes de risco: Comum CM e Idiossincrático (ID), mutuamente complementares. O componente de risco Comum captura questões macroeconômicas, geopolíticas e relacionadas a cultura de um país enquanto o componente Idiossincrático é composto por fatores especificamente relacionados aos diferentes TP's e LN's;
2. Os fatores que compõem o componente idiossincrático podem ser relacionados por uma função linear; e
3. Os fatores não observados no escalonamento são, supostamente, comuns a todos os bancos, não afetando o processo de escalonamento.

Para avaliar qual relação linear é a mais adequada para explicar as perdas a partir dos componentes comum e idiossincrático, propõe-se o uso de regressão linear considerando o *log* das perdas internas e externas da instituição estudada.

Dessa forma, considerando as informações disponíveis para construção de variáveis e posterior modelo para estimação dos parâmetros de regressão, será construído fator de escala para, no mínimo, ponderar a região onde as perdas operacionais ocorreram.

O fator de escala é construído a partir do seguinte raciocínio:

- Considere que as perdas são função dos componentes de risco Comum e Idiossincrático.

$$S = f(CM, ID) \quad (3.2)$$

- Assumindo a premissa de que o \log das perdas se relaciona com os componentes por uma transformação linear pode-se assumir que:

$$\log(S) = CM + \beta ID \quad (3.3)$$

- Supondo que o componente idiossincrático seja composto pelos fatores Região, Porte, Linha de Negócio e Tipo de Evento de Perda, a equação pode ser reescrita como:

$$\log(S) = CM + \beta_1 Regiao + \beta_2 Porte + \beta_3 LN + \beta_4 TP \quad (3.4)$$

- É utilizado o método MQO (Mínimos Quadrados Ordinários) para estimar os parâmetros da equação anterior, atribuindo pesos para cada uma das variáveis.

$$\log(\hat{S}) = \hat{C}M + \hat{\beta}_1 Regiao + \hat{\beta}_2 Porte + \hat{\beta}_3 LN + \hat{\beta}_4 TP \quad (3.5)$$

Considerando que os dados da instituição financeira que deseja obter o fator de escala estão contidos nos dados da BDE, é possível, após a estimação dos parâmetros da regressão linear como em (3.5), construir o fator de escala.

Para isto, é necessário definir qual(is) fator(es) deseja(m)-se escalonar. Recomenda-se construir o fator de escala para as perdas comuns ao mesmo TP e a mesma LN, ponderando apenas os outros fatores do componente Idiossincrático no processo de escalonamento.

Em outras palavras, se o objetivo é escalonar as perdas de diferentes regiões e portes, para as Linhas de Negócio (LN) e Tipos de Evento (TP) iguais, por exemplo, o fator de escala correspondente seria encontrado facilmente pela equação dado por:

$$\begin{aligned} FS &= \frac{e^{\hat{S}_{BDI}}}{e^{\hat{S}_{BDE}}} \\ &= \frac{\exp(\hat{\beta}_1 Regiao_{BDI} + \hat{\beta}_2 Porte_{BDI} + \hat{\beta}_3 LN + \beta_4 TP)}{\exp(\hat{\beta}_1 Regiao_{BDE} + \hat{\beta}_2 Porte_{BDE} + \hat{\beta}_3 LN + \hat{\beta}_4 TP)} \end{aligned} \quad (3.6)$$

Dessa forma, para obter a perda escalonada por região e porte, basta aplicar o FS encontrado em (3.7) aos dados externos de perdas operacionais, da forma que:

$$Perda_{escalonada} = Perda_{BDE} \times FS \quad (3.7)$$

Após os dados externos serem devidamente filtrados e escalonados, os mesmos podem ser incorporados ao modelo de alocação de capital para risco operacional.

3.4 Métodos de Incorporação

A incorporação dos dados externos ao modelo AMA tem como objetivo complementar os registros internos de perdas operacionais, reduzindo eventual escassez de dados.

Dados externos são utilizados com frequência em diversas áreas do conhecimento, para contribuir com a construção e análise de cenários de risco, pois constituem-se em incertezas críticas, definidas como eventos portadores de futuro, que, apesar de nunca antes terem sido observados no ambiente de controle, podem ocorrer a qualquer momento com consequências e potencialidades já conhecidas (Schwartz, 1996).

Neste trabalho, a incorporação dos dados da BDE ao modelo AMA será proposta no âmbito da construção da Distribuição Agregada de Perdas (DAP) via LDA (Loss Distribution Approach), considerando esta técnica a mais adequada para cálculo de alocação de capital na Abordagem avançada, conforme exposto na Seção 2.1.

Tal forma de incorporação foi motivada a partir da necessidade de avaliar separadamente o impacto de cada elemento (BDI, BDE, Cenários e BEICF) no valor de alocação de capital utilizando o modelo AMA, conforme estabelecido na Circular N° 3.647 (Banco Central, 2013).

Serão propostos dois métodos para utilização dos eventos da BDE na construção da DAP: Método da União e Método da Interseção, definidos como segue.

3.4.1 Método da União

A ideia desta forma de incorporação é simples e consiste em adicionar eventos externos de perda a cauda (valores acima do *threshold*) da distribuição da BDI, reduzindo eventual escassez de dados. Para isto, propõe-se que, após os tratamentos de filtro e escalonamento da BDE, os dados externos sejam considerados no processo de estimação da distribuição de severidade das perdas, via LD, juntamente com os dados da BDI.

Em outras palavras, o método da união visa aumentar a quantidade de eventos de severidade das perdas, de modo a contribuir com uma avaliação prospectiva de risco, incorporando eventos com probabilidade de ocorrência já identificada aos dados internos de perda.

Sejam

- S_{BDI} a variável aleatória que representa os montantes de perdas ocorridos na insti-tuição que estão gravados na BDI.
- S_{BDE} a variável aleatória que representa os montantes de perdas registrados na BDE.

Ainda, considerando independência entre os eventos da BDI e BDE, e independência entre os eventos da BDE, pode-se definir a variável aleatória $S_{BDI \cup BDE}$ como a união dos dados da BDI com os da BDE, com função de probabilidade dada por:

$$P(S_{BDI \cup BDE} \leq s) = P(BDI)P(S_{BDI} \leq s) + P(BDE)P(S_{BDE} \leq s) \quad (3.8)$$

Desta forma, visando a utilização desta abordagem na metodologia LDA, basta unir os dados da BDE aos dados da BDI para que seja possível avaliar o impacto da incorporação deste elemento no valor de alocação de capital sob abordagem AMA.

Para utilização deste método, o processo de construção da distribuição agregada de perdas via LDA deve ser modificado para considerar os dados da BDE no processo de estimação da distribuição de severidade das perdas.

Partindo do pressuposto que a única fonte de dados utilizada para construir a BDE é de dados de consórcio e que, sendo assim, os dados da BDI (acima do *threshold*) estão inclusos na BDE, a incorporação pelo método da união deve prever a substituição dos dados da BDI (acima do *threshold*) pelos dados da BDE. Feito isto, pode-se seguir com o algoritmo LDA para construção da DAP, conforme Seção 2.1.1.

O uso dos dados da BDE não é indicado para redefinição da distribuição de frequência das perdas visto que alteraria o perfil de riscos da instituição.

3.4.2 Método da Interseção

Este método é uma alternativa mais sofisticada ao método da União e surge a partir da necessidade do modelo ser mais sensível ao perfil de riscos da instituição.

Enquanto no Método da União os dados externos inflam a cauda da distribuição de severidade, aumentando a probabilidade de ocorrência de perdas com valores altos, no Método da Interseção, para evitar este viés para cima dos valores de perda, propõe-se estimar as distribuições de severidade da BDI e da BDE separadamente, incorporando os dados externos apenas quando, na execução do LDA, for sorteado da BDI algum evento maior do que o *threshold*.

O algoritmo LDA deverá então ser reconfigurado para testar se a observação sorteada da BDI no momento da simulação de Monte Carlo é maior ou menor que o *threshold*, conforme algoritmo a seguir que só deve ser utilizado após o processo de filtro e escalonamento dos dados.

Passo 1: Sorteie um dado de frequência da distribuição de contagem, $N = n_i$;

Passo 2: Sorteie da distribuição de severidade da BDI n_i observações, $S_1, S_2, \dots, S_k, \dots, S_{n_i}$

Se $S_k < threshold$, $k = 1, 2, \dots, n_i$, guarde esta observação.

Se $S_k \geq threshold$, sorteie uma observação, S_k^* da distribuição de severidade da BDE e substitua S_k por S_k^*

Passo 3: Some as observações de tal forma que $POA = \sum_{k \in BDI} S_k + \sum_{k \in BDE} S_k$

Passo 4: Repita os passos de 1 a 3 um número m grande de vezes.

No capítulo a seguir serão testados os dois métodos de incorporação da BDE propostos neste trabalho, utilizando base de dados real fornecida por um grande banco brasileiro com dados internos e externos coletados em um período de cinco anos, com o objetivo de avaliar a qualidade das estimativas da distribuição agregada de perdas geradas a partir dos métodos propostos.

Serão mostrados os resultados obtidos a partir da implementação das técnicas abordadas neste trabalho, contemplando as estimativas das distribuições de frequência e severidade, bem como o uso do LDA para incorporação da BDE pelos diferentes métodos propostos. A implementação dos algoritmos foi formulada em R (([R Core Team, 2015](#)) e estão dispostas na forma de Anexo a este trabalho.

4 Aplicações e Resultados

Para análise de impacto das diferentes metodologias de incorporação da BDE ao modelo AMA, foi utilizada base de dados real cedida por um grande banco brasileiro, conveniado a um consórcio internacional de dados externos que reúne eventos de perda de diferentes países do mundo.

Os dados foram devidamente descaracterizados a fim de manter a confidencialidade acerca do apetite a risco da instituição.

Serão analisadas três categorias de risco com características distintas, mutuamente excludentes, pertencentes a linha de negócio Varejo, denotadas neste trabalho por Categoria A, B e C. Para efeito de cálculo, foram selecionados eventos ocorridos entre Jan/2010 e Dez/2014, avaliados sob a perda bruta.

O valor de alocação de capital para risco operacional será calculado considerando as abordagens Empírica e Paramétrica, ficando a cargo da instituição financeira escolher a mais adequada para uso na gestão de riscos.

4.1 Dados internos

A Categoria de Risco A se refere a um subnível de granularidade do Tipo de Evento Fraudes e Roubos Externos classificado na Linha de Negócio Varejo.

As perdas relativas a esta Categoria ocorrem com uma frequência média de 13570 eventos mensais, com coeficiente de variação de 0,45, observado o período de Jan/2010 a Jan/2014. A distribuição de probabilidade mais adequada para ajuste foi a Binomial Negativa com parâmetros $\hat{\mu} = 12554,9$ e $\hat{p} = 0,0004$, como pode ser observado na Figura 1.

No que se refere aos dados de severidade, as distribuições utilizadas para ajuste das perdas operacionais observadas nesta categoria podem ser avaliadas na Figura 1b. Foram observados 814.253 eventos distribuídos entre R\$1,00 e R\$488.100,00, com média de R\$1.400,00 de perda por evento.

Para ajuste da distribuição foi utilizado como variável aleatória o logaritmo do valor da perda. Como se pode notar, a distribuição que melhor se ajustou aos dados foi a Mistura de Normais com parâmetros $(\hat{\mu}_1 = 2,91, \hat{\sigma}_1 = 1,09, \hat{\mu}_2 = 6,59, \hat{\sigma}_2 = 1,28, \hat{p} = 0,19)$, mostrando que a distribuição dos dados é multimodal.

As estatísticas descritivas relacionadas a frequência e severidade das perdas desta e das demais categorias pode ser consultada na Tabela 3.

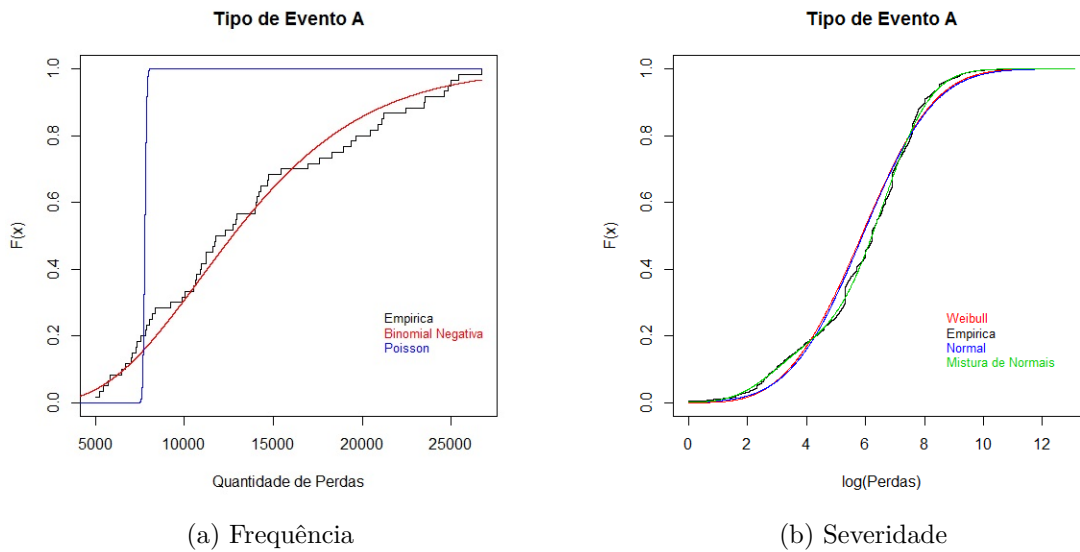


Figura 1 – Distribuições de Probabilidade Estimadas - Categoria A

Para ajuste do logaritmo das perdas da Categoria B, pertencente ao TP Perdas Trabalhistas, foram utilizados 10.363 observações distribuídas entre R\$ 1,00 e R\$ 3.312.000,00. As distribuições Binomial Negativa com parâmetros $\hat{\mu} = 185,67$ e $\hat{p} = 0,003$ e Mistura de Normais com parâmetros $\hat{\mu}_1 = 10,11$, $\hat{\sigma}_1 = 1,56$, $\hat{\mu}_2 = 4,92$, $\hat{\sigma}_2 = 1,79$, $\hat{p} = 0,12$, foram as melhores ajustadas para os dados de frequência e severidade dos dados, respectivamente, como pode ser visto na figura 2.

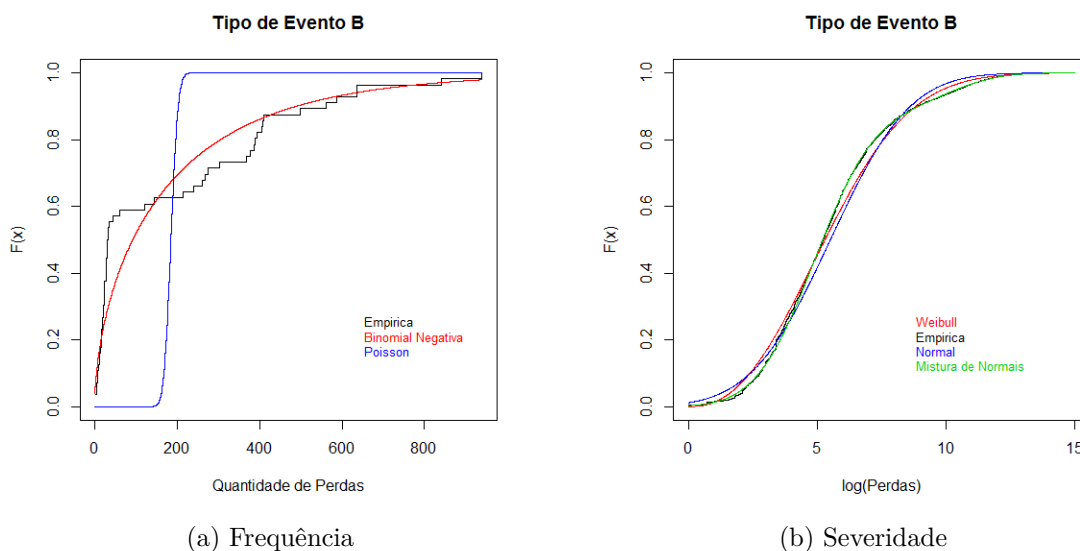


Figura 2 – Distribuições de Probabilidade Estimadas - Categoria B

Assim como para as demais categorias, além das distribuições Binomial Negativa e Mistura de Normais, foram testadas ainda a Poisson para os dados de frequência e a Normal e Weibull para os dados de severidade. Os parâmetros ajustados para as respectivas

distribuições podem ser consultados nas Tabelas 4 e 5.

Ainda, as distribuições de frequência e severidade melhor ajustadas aos dados da Categoria C, referente a perdas ocorridas no TP Falhas nos Negócios, foram a Binomial Negativa ($\mu=513,91$, $p=0,0022$) e Mistura de Normais ($\mu_1 = 9,22$, $\sigma_1 = 1,09$, $\mu_2 = 3,13$, $\sigma_2 = 1,30$, $p = 0,43$), respectivamente, conforme pode ser visto na Figura 3.

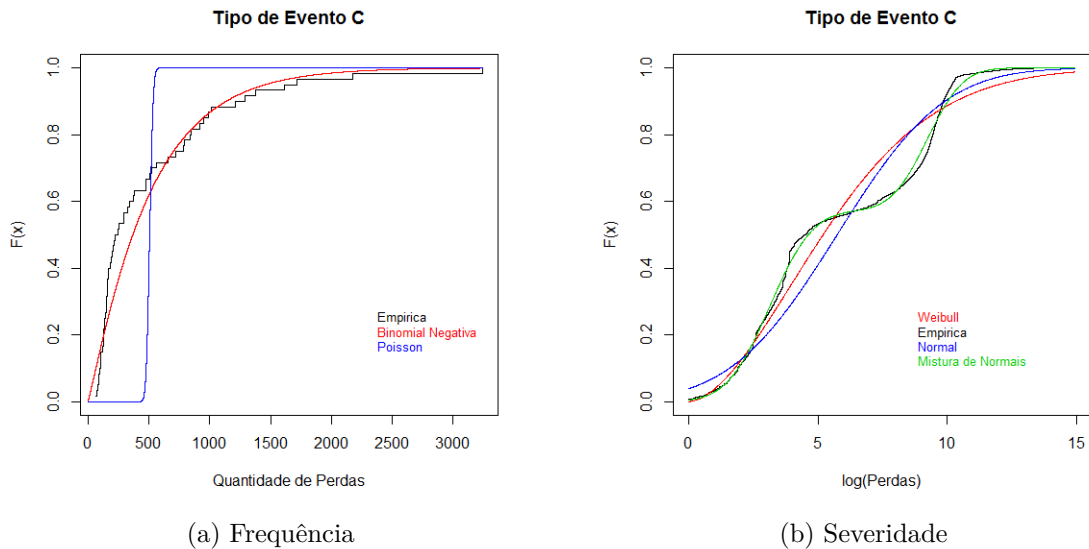


Figura 3 – Distribuições de Probabilidade Estimadas - Categoria C

A escolha das distribuições melhor ajustadas aos dados foi feita levando em consideração, além da análise gráfica, o Desvio Quadrático Médio dado em (2.3). Os resultados para análises dos ajustes de frequência e severidade dos dados internos podem ser consultados na Tabela 6.

Tabela 3 – Estatísticas Descritivas das Frequências e Severidades - Categorias A, B e C

	Categoria A	Categoria B	Categoria C
Frequência Mensal			
N	814.253	10.363	30.448
Média	13.570	172	507
Coefficiente de Variação	0,45	1,28	1,16
Severidade Mensal (R\$)			
Mínimo	1,00	1,00	1,00
1º Quartil	139,00	43,00	19,50
Média	1.401,00	9.932,00	9.079,00
Mediana	500,00	187,00	89,90
3º Quartil	1.470,00	899,00	10.290,00
Máximo	488.100,00	3.312.000,00	3.049.000,00

Tabela 4 – Estimativas dos parâmetros dos modelos Poisson e Binomial Negativa

	Poisson	Binomial Negativa	
	λ	μ	p
Categoria A	7788,53	12554,98	0,0004
Categoria B	185,05	185,67	0,003
Categoria C	507,46	513,91	0,002

Tabela 5 – Estimativas dos parâmetros dos modelos Weibull, Normal e Mistura de Normais

	Weibull		Normal		Mistura de Normais				
	α	β	μ	σ	μ_1	σ_1	μ_2	σ_2	p
Categoria A	3,45	6,52	5,89	1,91	6,59	1,28	2,91	1,09	0,19
Categoria B	2,35	6,19	5,51	2,42	10,11	1,56	4,91	1,79	0,88
Categoria C	1,75	6,41	5,73	3,25	9,22	1,09	3,13	1,30	0,57

Tabela 6 – DQM dos ajustes das distribuições de frequência e logaritmo das perdas

	Frequência		Log-Severidade		
	Poisson	Bin. Negativa	Weibull	Normal	Mistura Normais
Categoria A	0,1599	0,0015	0,0019	0,0017	0,0001
Categoria B	0,0926	0,0085	0,0006	0,0015	0,0001
Categoria C	0,1096	0,0073	0,0064	0,0077	0,0004

4.2 Dados Externos

No que compete aos registros externos de perdas operacionais da Categoria A, consta na Base de Dados Externa fornecida para estudo, 760 observações de perda, com valores acima do *threshold* de R\$50.000, ocorridas entre Jan/2010 e Dez/2014, distribuídas em 64 países do mundo, classificados em 6 diferentes regiões, conforme Figura 4.

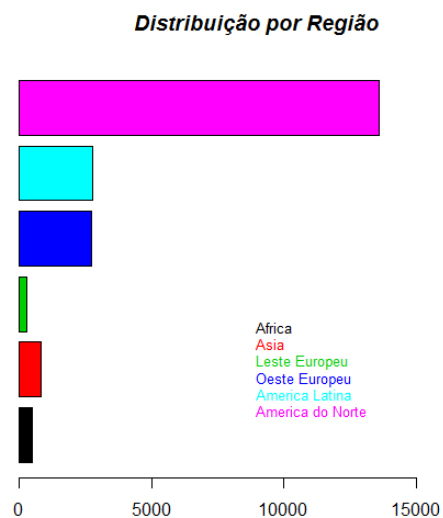


Figura 4 – Distribuição da BDE por Região

Visando agregar os dados da BDE ao modelo AMA, foram seguidos os procedimentos dispostos no Capítulo 3, começando pelo processo de filtragem, visto que o “De-Para” foi previamente realizado nas instituições financeiras estudadas.

A técnica de Boxplot para filtro dos *outliers*, conforme descrito em 3.1 foi aplicado no logaritmo das perdas externas, eliminando 43 observações da Categoria A, 434 da Categoria B e 1521 da Categoria C, alterando assim a cauda da distribuição dos dados da BDE conforme pode ser apreciado na Figura 5.

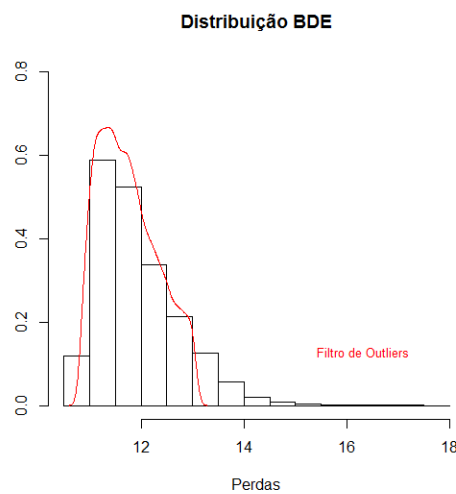


Figura 5 – Distribuição do Logaritmo das perdas da BDE

Mais estatísticas descritivas sobre os eventos externos das Categorias A, B e C podem ser apreciadas nas Tabelas 7, 8 e 9, respectivamente, que contêm ainda informações a cerca da distribuição dos dados externos após aplicação dos filtros e também dados da BDI que tem valor acima do *threshold*, para comparação.

Vale pontuar que os dados da BDI acima do *threshold*, fazem parte da base de dados externa e são utilizados no processo de filtragem e no modelo de escalonamento da BDE.

Tabela 7 – Descritiva BDE - Categoria A

	BDE	$BDE_{filtrada}$	$BDI_{>threshold}^1$
N	760	717	667
Minimo	50.170,00	50.170,00	50.040,00
1º Quartil	69.980,00	68.710,00	56.990,00
Média	166.800,00	106.800,00	84.260,00
Mediana	88.400,00	85.230,00	67.660,00
3º Quartil	135.700,00	121.300,00	93.570,00
Máximo	5.183.000,00	469.100,00	488.100,00

Tabela 8 – Descritiva BDE - Categoria B

	BDE	$BDE_{filtrada}$	$BDI_{>threshold}$
N	2.611	2.177	401
Mínimo	50.130,00	50.130,00	50.170,00
1° Quartil	84.120,00	77.210,00	69.230,00
Média	453.700,00	150.400,00	208.500,00
Mediana	137.400,00	115.900,00	107.400,00
3° Quartil	298.900,00	196.200,00	203.800,00
Máximo	50.030.000,00	470.500,00	3.312.000,00

Tabela 9 – Descritiva BDE - Categoria C

	BDE	$BDE_{filtrada}$	$BDI_{>threshold}$
N	17.065	15.544	582
Mínimo	50.010,00	50.010,00	50.090,00
1° Quartil	83.900,00	80.890,00	77.290,00
Média	229.300,00	153.600,00	207.500,00
Mediana	129.700,00	120.200,00	125.300,00
3° Quartil	236.700,00	194.500,00	227.300,00
Máximo	41.050.000,00	471.700,00	3.049.000,00

4.2.1 Modelo de Escalonamento

Após o processo de filtragem dos dados, foi construído modelo de escalonamento considerando os dados de severidade das observações. Para isto, foram utilizadas as variáveis disponibilizadas na base de dados estudada, a saber: Região de ocorrência da perda e Tipo de Evento de Perda (TP).

Para construção de modelo mais robusto no sentido de significância dos parâmetros e explicação da variável de interesse, recomenda-se a ampliação das variáveis explicativas, adicionando, no mínimo, as variáveis Porte e Linha de Negócio ao processo de estimação do modelo.

Os resultados apresentados a seguir foram gerados utilizando a função *lm* do software R, que faz uso do método MQO para estimação dos parâmetros de regressão linear, calculados a partir das variáveis disponibilizadas na base de dados.

Vale ressaltar que a base de dados fornecida para este trabalho não contém toda a informação disponível para a instituição financeira, gerando resultados pouco significativos, que serão utilizados nesta dissertação de forma conservadora para ilustrar o uso da metodologia.

Considerando o interesse de trazer as perdas externas à “realidade” do Brasil, os eventos da BDE foram escalonados de tal forma a ser possível calcular o impacto das diferentes regiões no valor das perdas, sempre que possuírem o mesmo TP.

```
lm(formula = log(perda) ~ bde$Regiao + tpext)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.97779	-0.46164	-0.07491	0.40431	1.58891

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	11.43511	0.03727	306.823	<2e-16 ***
bde\$RegiaoAsia	-0.08276	0.03690	-2.243	0.0249 *
bde\$RegiaoEastern Europe	0.01580	0.04652	0.340	0.7342
bde\$RegiaoLatin America	0.02356	0.03158	0.746	0.4557
bde\$RegiaoAmerica Norte	0.05223	0.03253	1.606	0.1084
bde\$RegiaoWestern Europe	0.03462	0.03210	1.079	0.2808
tpextC	0.31044	0.02344	13.247	<2e-16 ***
tpextB	0.28813	0.02556	11.273	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5674 on 18430 degrees of freedom

Multiple R-squared: 0.01254, Adjusted R-squared: 0.01216

F-statistic: 33.44 on 7 and 18430 DF, p-value: < 2.2e-16

Os fatores de escala para as diferentes Regiões i , (FS_i), encontrados a partir do modelo de regressão são definidos como:

$$FS_i = \frac{\exp(\hat{\beta}_{RegiaoAmericaLatina})}{\exp(\hat{\beta}_{Regiao_i})}$$

Que fornece os resultados listados a seguir.

- $FS_{Asia} = 1,112$ (Incremento de 11,2% no valor das perdas ocorridas na Ásia)
- $FS_{Africa} = 1,024$ (Incremento de 2,4% no valor das perdas ocorridas na África)
- $FS_{EaternEurope} = 1,008$ (Incremento de 0,8% no valor das perdas ocorridas no Leste Europeu)
- $FS_{WesternEurope} = 0,989$ (Desconto de 1,1% no valor das perdas ocorridas no Oeste Europeu)
- $FS_{AmericaNorte} = 0,972$ (Desconto de 2,8% no valor das perdas ocorridas na América do Norte)

As perdas ocorridas na Região América Latina tem fator de escala $1 \left(\frac{e^{0,02356}}{e^{0,02356}} \right)$, ou seja, não há escalonamento das perdas ocorridas na mesma região a que pertence a instituição financeira estudada.

4.3 Construção da DAP

Para avaliar o efeito da incorporação dos dados da BDE pelos Métodos da União e da Interseção definidos da Seção 3.4, foram construídas as Distribuições Agregadas de Perda (DAP) para os conjuntos de observações da BDI e da união ou interseção dos dados da BDE a BDI.

As DAP foram formuladas a partir das distribuições de frequência e severidade dos eventos de perda das três diferentes categorias relacionadas neste capítulo. Foram avaliadas a abordagem empírica e paramétrica para construção da DAP. Na primeira abordagem, quando do uso do LDA para simulação das perdas operacionais agregadas, os dados são sorteados diretamente do conjunto de eventos observados, tanto da BDI quanto da BDE. Já na abordagem paramétrica, são estimadas distribuições de probabilidade para os dados de frequência (exclusivamente da BDI) e de severidade, como mostrado na Seção 4.1.

No Método da União, a distribuição paramétrica da severidade é estimada para o conjunto ($BDI \cup BDE$), enquanto que no Método da Interseção, são estimadas separadamente distribuições de severidade para os dados da BDI e da BDE, sendo que no momento do sorteio de observações de severidade, a informação que fará parte da construção da POA pode ser de uma fonte ou de outra.

O método LDA foi avaliado com 1.000.000 de iterações de Monte Carlo afim de poder calcular o $VaR_{99,9}$ da DAP. Para avaliar a qualidade das estimativas, utilizou-se a técnica de re-amostragem Bootstrap com $B=100$ iterações com o objetivo de estimar o Erro Padrão das estimativas (EP) bem como o Intervalo Percentil para o $VaR_{99,9}$. As estimativas encontradas estão separadas por categoria, como segue.

4.3.1 Categoria A

A Categoria A, relativa a Fraudes e Roubos Externos é a que possui a maior quantidade de eventos na BDI. Considerado o fato de que, no Brasil, o quesito segurança - seja física ou virtual ainda deixa a desejar e que, apesar do grande esforço das instituições financeiras para tornar o ambiente das agências e dos meios de pagamento eletrônicos mais seguros, há ainda muitos criminosos especializadas em roubos e assaltos no ambiente das agências bancárias e nos meios de pagamento eletrônico, a grande massa de eventos nesta categoria é bem entendida pelas instituições financeiras do Brasil, que acompanham as informações de perda diariamente e dispõe, cada vez mais, de dispositivos para impedir

que este tipo de evento ocorra com altas severidades. Tanto que menos de 1% dos eventos desta categoria possuem valor acima do *threshold*, representando 667 eventos.

Na BDE esta categoria é pouco representativa, com poucas observações (93 eventos) provenientes de outras regiões participantes do consórcio. Desta forma, como pode ser visto na Tabela 10, a incorporação dos dados externos na modelagem das perdas operacionais não resultou em diferença significativa no valor de alocação de capital por nenhum dos dois métodos (União e Interseção).

Tabela 10 – Estimativas do $VaR_{99,9}$ para Categoria A (R\$mil)

		Empírico	Paramétrico
BDI	$VaR_{99,9}$	38.436	50.373
	$IC_{(0,05;0,95)}$	(38.417 ; 38.453)	(50.134 ; 50.657)
	EP	<0,001	<0,001
BDI \cup BDE	$VaR_{99,9}$	39.211	50.484
	$IC_{(0,05;0,95)}$	(39.193 ; 39.228)	(50.280 ; 50.746)
	EP	<0,001	<0,001
BDI \cap BDE	$VaR_{99,9}$	39.074	50.188
	$IC_{(0,05;0,95)}$	(39.058 ; 39.099)	(50.056 ; 50.478)
	EP	<0,001	<0,001

As Figuras 6a, 6b e 6c mostram as distribuições de probabilidade empírica e paramétrica nas visões BDI, $BDI \cup BDE$ e $BDI \cap BDE$, respectivamente, para a Categoria A. Note que, apesar da DAP empírica ter uma distribuição mais heterogênea, a DAP paramétrica fornece estimativas mais altas do $VaR_{99,9}$ pelo fato de possibilitar uma gama maior de valores de cauda, independente da fonte das observações (BDI ou BDI com BDE).

4.3.2 Categoria B

Esta categoria, relativa a perdas decorrentes de ações judiciais, tem a característica de agrupar várias ocorrências em um único evento, como por exemplo acontece quando várias pessoas entram com uma ação conjunta contra a instituição financeira. O pagamento é feito em várias parcelas (cada pessoa é referente a uma parcela) mas só há um evento causador da perda, o que permite considerar a severidade deste evento como a soma das parcelas. Por este motivo, há menos eventos registrados na BDI, porém a severidade deles são maiores.

Consta na BDE mais de cinco vezes a quantidade de perdas ocorridas na BDI, o que causa um grande aumento na cauda da distribuição de severidade quando os dados são incluídos pelo método da União, gerando estimativas do VaR bem superiores ao observado apenas com os dados da BDI. Em contrapartida, no método da Interseção, os valores para o VaR empírico e paramétrico são subestimados pelo fato de que a distribuição de

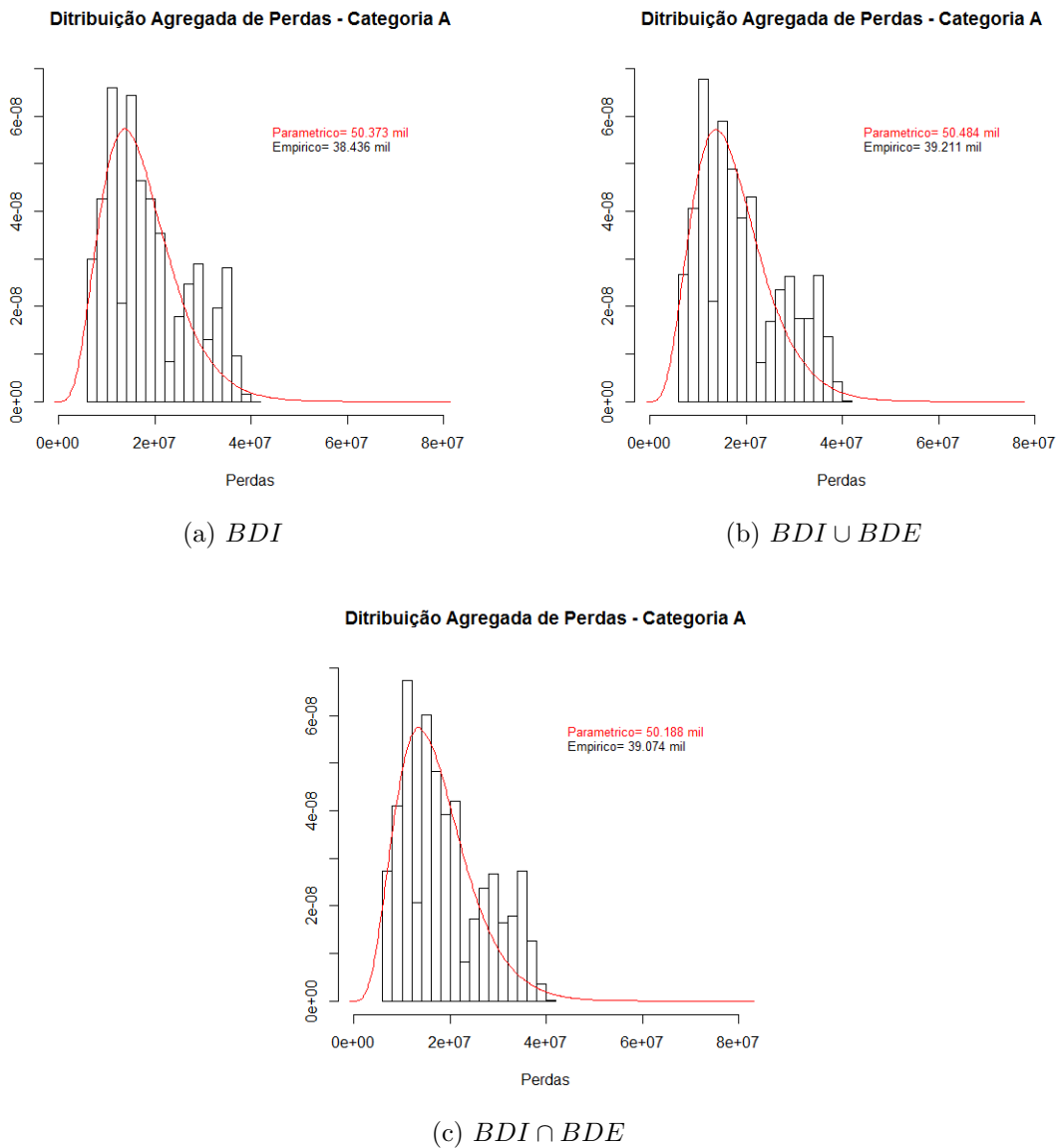


Figura 6 – Distribuições Agregadas de Perda considerando a incorporação da BDE para a Categoria A

severidade da BDE, nesta categoria, é mais dispersa do que a da BDI, com valores de quantis, após escalonamento, em geral abaixo dos observados pela instituição em estudo.

As estimativas para os diferentes métodos de incorporação e de fonte de dados podem ser apreciadas na Tabela 11. A Figura 7 mostra as diferentes curvas de distribuição de acordo com a abordagem utilizada.

4.3.3 Categoria C

Por último, a Categoria C, referente a perdas decorrentes de ação involuntária ou negligente de uma obrigação profissional frente a clientes, de modelagem de processos ou da própria natureza de um produto, é a classe de eventos que mais possui diferenças

Tabela 11 – Estimativas do $VaR_{99,9}$ para Categoria B (R\$mil)

		Empírico	Paramétrico
BDI	$VaR_{99,9}$	14.153	22.647
	$IC_{(0,05;0,95)}$	(14.072 ; 14.227)	(22.454 ; 22.883)
	EP	<0,001	<0,001
BDI \cup BDE	$VaR_{99,9}$	29.998	53.604
	$IC_{(0,05;0,95)}$	(29.954; 30.049)	(53.186 54.124)
	EP	<0,001	<0,001
BDI \cap BDE	$VaR_{99,9}$	9.026	14.381
	$IC_{(0,05;0,95)}$	(9.003 ; 9.053)	(14.260 ; 14.484)
	EP	<0,001	<0,001

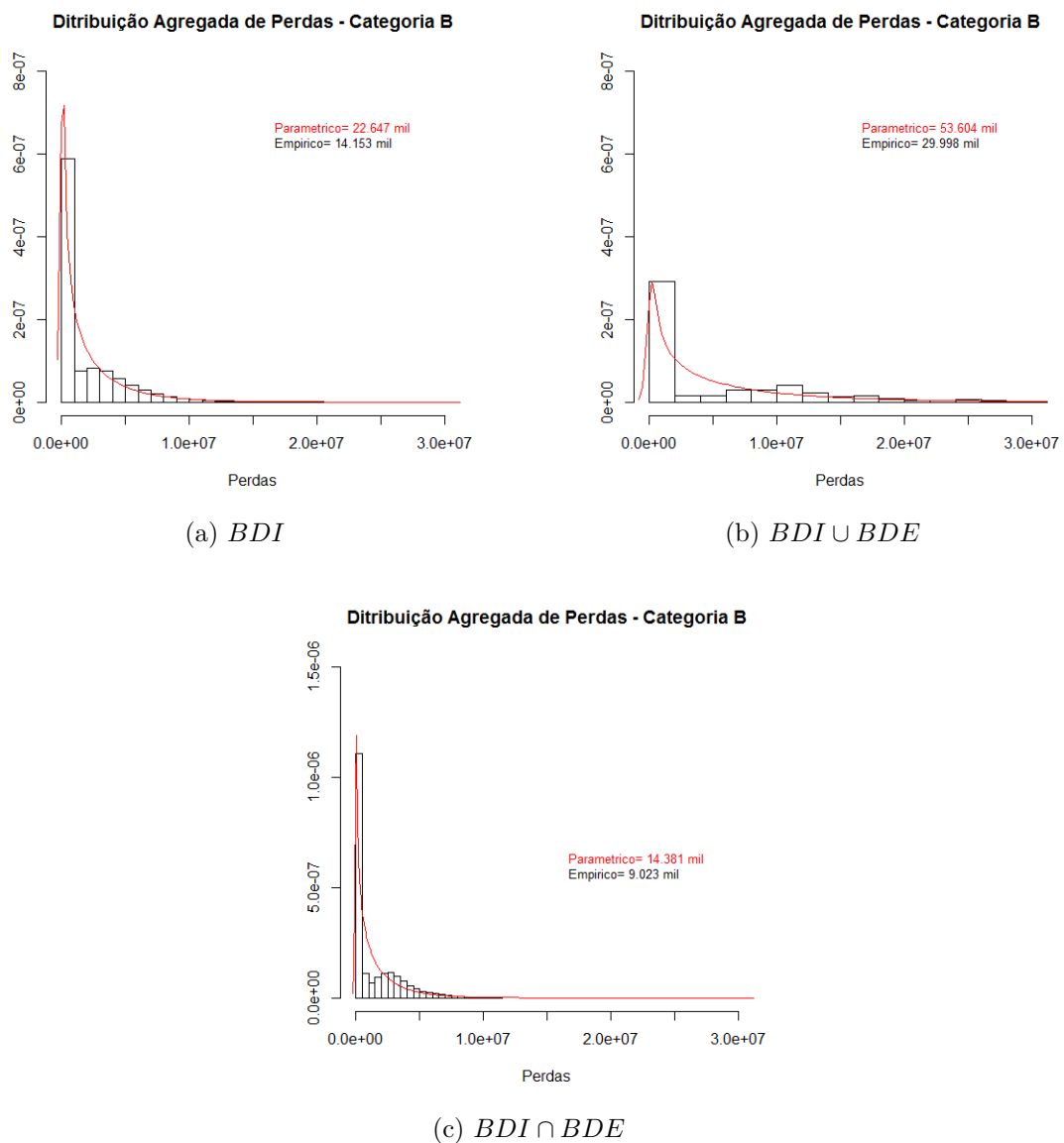


Figura 7 – Distribuições Agregadas de Perda considerando a incorporação da BDE para a Categoria B

entre os dados da BDI e da BDE, começando pela quantidade de observações acima do *threshold*, 15.544 após filtro de *outliers* da BDE contra 582 da BDI. Este fato, assim como visto na Categoria B, superestima a cauda da distribuição de severidade quando da utilização do Método da União para incorporação da BDE no método LDA para geração da DAP.

No caso do uso do método da Interseção, verificou-se que, na abordagem paramétrica, houve um incremento no valor do VaR possivelmente associado a qualidade do ajuste da distribuição paramétrica da BDE, visto que na abordagem empírica o valor médio do $VaR_{99,9}$ ficou abaixo dos dados da BDI, influenciado pela distribuição dos dados escalonados que apresentaram valores estimados dos quantis da distribuição externa, em geral, inferiores aos da distribuição dos dados internos.

Tabela 12 – Estimativas do $VaR_{99,9}$ para Categoria C (R\$mil)

		Empírico	Paramétrico
BDI	$VaR_{99,9}$	33.740	26.260
	$IC_{(0,05;0,95)}$	(33.659 ; 33.818)	(26.078 ; 26.434)
	EP	<0,001	<0,001
BDI \cup BDE	$VaR_{99,9}$	189.458	288.540
	$IC_{(0,05;0,95)}$	(189.350; 189.581)	(286.347 ; 290527)
	EP	<0,001	<0,001
BDI \cap BDE	$VaR_{99,9}$	28.430	35.703
	$IC_{(0,05;0,95)}$	(28.393 ; 28.477)	(35.386 35.963)
	EP	<0,001	<0,001

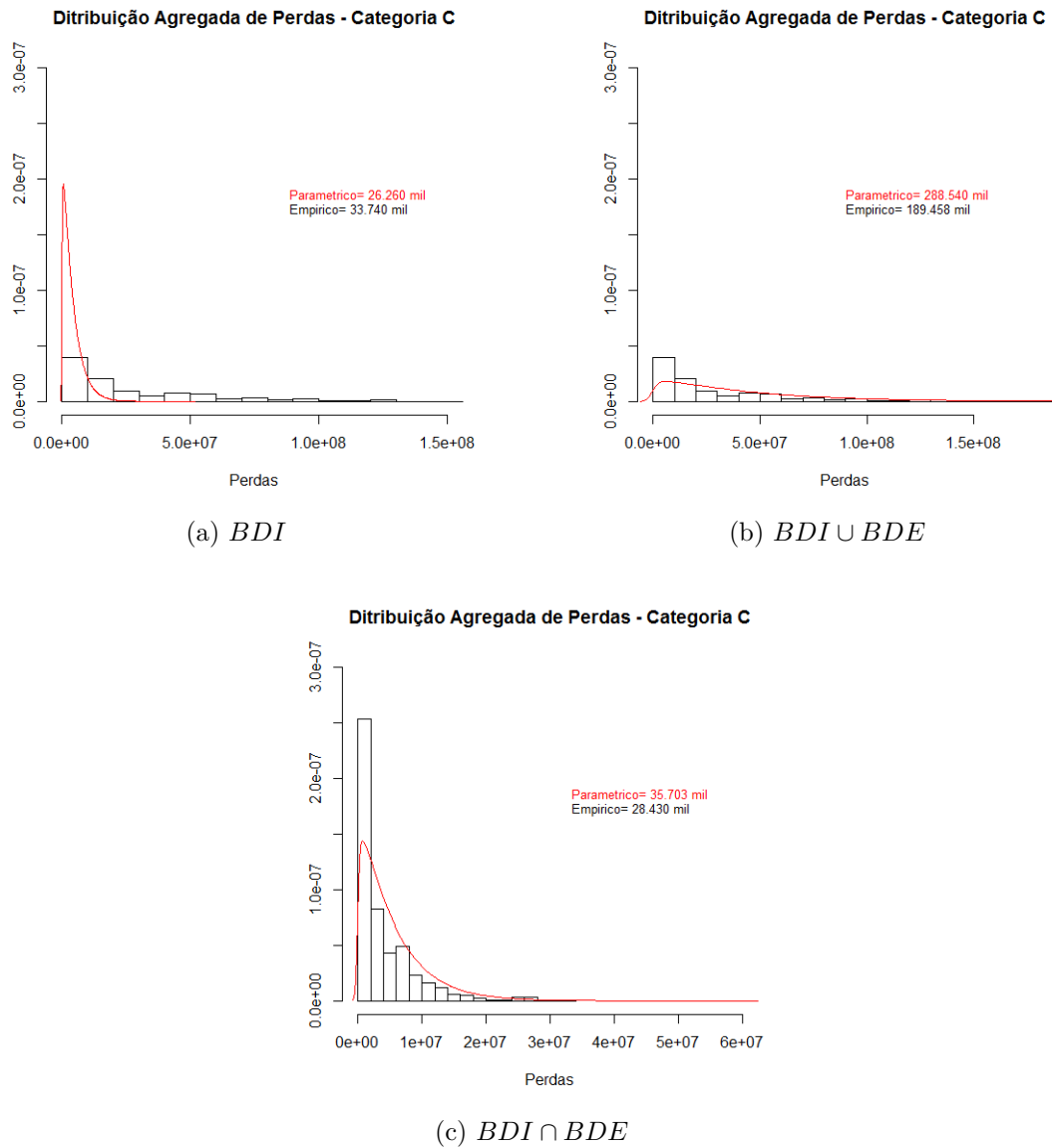


Figura 8 – Distribuições Agregadas de Perda considerando a incorporação da BDE para a Categoria C

Conclusão

Foram abordados neste trabalho as metodologias mais utilizadas na modelagem de capital para risco operacional a fim de discutir o tema e avaliar as necessidades que o permeiam. Apresentou-se a definição do risco operacional, a evolução dos modelos regulamentares para cálculo de alocação de capital até a possibilidade de uso de modelos internos para encontrar estimativa que garanta o cálculo com 99,9% de confiança do valor de alocação a partir de uma distribuição agregada de perdas.

Foi utilizada a técnica LDA que faz uso de simulação de Monte Carlo para, a partir das distribuições de frequência e severidade dos dados de perda, construir a distribuição agregada de perdas. Foram construídas distribuições empíricas e paramétricas dos dados de contagem e dos valores de perda para avaliação do comportamento do $VaR_{99,9}$ sob estas duas abordagens e se verificou que a abordagem paramétrica formula distribuições com caudas mais pesadas a direita, retornando valores de alocação de capital maiores do que as estimadas na abordagem empírica.

As estimativas do $VaR_{99,9}$, utilizadas como valor para alocação de capital, foram avaliadas sob o Erro Padrão, gerados a partir de 100 repetições de Bootstrap, as quais, para todas as abordagens utilizadas, retornaram valores abaixo de 0,001, mostrando que o LDA é adequado para estimar este parâmetro.

Como forma de contribuição à literatura do risco operacional foram propostos dois métodos de incorporação do elemento BDE à modelagem de alocação de capital, denotados por Método da União e Método da Interseção, que devem ser utilizados apenas após o devido tratamento de filtro de *outliers* e escalonamento das observações da base externa de dados.

Foi observado no tratamento de *outliers* que este deve ser mais granular, devendo ser efetuado para cada categoria de risco, visando um tratamento mais homogêneo da exclusão de observações, visto que percebeu-se que eventos com alta probabilidade de ocorrência, mais especificamente com valores abaixo dos já observados na BDI, foram excluídos. Recomenda-se que, antes da exclusão dos *outliers*, seja feita análise destes por parte de analistas de risco, a fim de avaliar a real necessidade da retirada destas observações da base.

Os resultados do VaR calculados a partir dos métodos da União e da Interseção propostos neste trabalho mostram que para a escolha de um outro método, é necessária análise exploratória dos dados da BDI que ficam acima do *threshold*, comparados aos dados da BDE que serão incorporados à base, com o objetivo de avaliar a distribuição dos dados de severidade no que compete a quantidade de observações e perfil das perdas.

Das simulações construídas a partir de base de dados reais pode-se observar que o Método da União superdimensiona o VaR nas ocasiões em que a quantidade de dados incorporadas é muito superior a frequência observada na BDI, independente do perfil das perdas, pelo fato de inflar a cauda da distribuição de severidade, atribuindo maior peso aos dados de alto valor. Já o Método da Interseção se mostrou ineficiente quando a distribuição dos dados de severidade da BDE é similar à da BDI, retornando valores menores de VaR quando os dados são incorporados. Os dois métodos apresentaram resultados similares quando as observações de perda da BDE e da BDI não distinguem substancialmente na frequência e severidade das observações.

Resultado da aplicação do LDA a dados simulados de BDE com perfil de severidade das perdas em geral maior do que os da BDI (quantis mais altos na BDE), mostraram que o Método da Interseção é mais adequado, apresentando valores mais plausíveis de resultado do *VaR*, acrescentando parcela de capital quando da incorporação dos dados externos. Tal resultado mostra a importância do processo de análise e tratamento dos dados externos, a fim de incorporar ao modelo observações plausíveis de ocorrência que adicionem informação prospectiva ao risco incorrido pela instituição, no sentido de adicionar dados com valores acima dos já observados.

Como sugestão de trabalhos futuros, indica-se, principalmente, a melhoria no processo de análise e filtragem dos dados externos, além do aprimoramento do modelo de escalonamento no que diz respeito a explicação do modelo, a fim de tornar mais apurado o fator de escala e conseqüentemente, o impacto da adição da BDE na alocação de capital.

Referências

- Banco Central, B. (2006). Resolução nº 3.380. Citado na página 25.
- Banco Central, B. (2008). Circular nº 3.383. Citado na página 25.
- Banco Central, B. (2011). Circular nº 3.640. Citado na página 27.
- Banco Central, B. (2013). Circular nº 3.647. Citado 4 vezes nas páginas 27, 30, 43 e 47.
- Basileia, C. (2004). *International convergence of capital measurement and capital standards: a revised framework*. Bank for International Settlements. Citado 2 vezes nas páginas 23 e 27.
- Dahen, H. and Dionne, G. (2010). Scaling models for the severity and frequency of external operational loss data. *Journal of Banking & Finance*, 34(7):1484–1496. Citado na página 45.
- Dalkey, N. C., Brown, B. B., and Cochran, S. (1969). *The Delphi method: An experimental study of group opinion*, volume 3. Rand Corporation Santa Monica, CA. Citado na página 31.
- Davies, J., Finlay, M., McLenaghan, T., and Wilson, D. (2006). Key risk indicators—their role in operational risk management and measurement. *ARM and RiskBusiness International, Prague*, pages 1–32. Citado na página 31.
- DeLong, E. R., DeLong, D. M., and Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, pages 837–845. Citado na página 39.
- Diniz, C., Delbem, D., and Requena, G. (2014). Modelagem estatística para risco operacional. XXI Simposio Nacional de Probabilidade e Estatística. Citado na página 37.
- Efron, B. and Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press. Citado na página 40.
- Frachot, A., Moudoulaud, O., Roncalli, T., et al. (2004). Loss distribution approach in practice. Citado na página 37.
- Franco, F. L., dos Santos Grumbach, R. J., Grumbach, R. P., and Ferreira, G. C. D. (2007). *Aplicação da teoria dos jogos à modelagem de cenários prospectivos: o método Lince*. PhD thesis, Tese (Doutorado em Ciências em Engenharia da produção)—Universidade Federal do Rio de Janeiro, Rio de Janeiro. Citado na página 31.

- Jonhson, N., Kotz, S., and Kemp, A. (1969). Discrete distributions: Distributions in statistics. Citado na página 32.
- Jorion, P. (1996). Risk2: Measuring the risk in value at risk. *Financial Analysts Journal*, 52(6):47–56. Citado na página 39.
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Citado 4 vezes nas páginas 33, 34, 35 e 49.
- Requena, G. d. L. (2014). Dependencia entre perdas em risco operacional. Master’s thesis, UFSCar. Citado na página 38.
- Scandizzo, S. (2006). Scenario analysis in operational risk management. *The Advanced Measurement Approach to Operational Risk*, pages 187–214. Citado na página 31.
- Schwartz, P. (1996). *The art of the long view: paths to strategic insight for yourself and your company*. Broadway Business. Citado na página 47.
- Schwarz, G. et al. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464. Citado na página 32.
- Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions. *The annals of mathematical statistics*, pages 279–281. Citado na página 39.
- Weibull, W. (1951). Wide applicability. *Journal of applied mechanics*, 103. Citado na página 33.
- Williamson, D. F., Parker, R. A., and Kendrick, J. S. (1989). The box plot: a simple visual method to interpret data. *Annals of internal medicine*, 110(11):916–921. Citado na página 44.

Anexos

ANEXO A – Códigos de Programação em R

A.1 Ajuste de Distribuições

A.1.1 Ajuste de Frequencia

```
#####Poisson
```

```
#### Definindo a Função de Verossimilhança
```

```
vero_p<-function(x,lambda)
```

```
{
```

```
  if (lambda>0) return (-1*sum(dpois(x,lambda,log=TRUE)))
```

```
  else return (-Inf)
```

```
}
```

```
#### Obtendo as Estimativas dos parâmetros
```

```
emv_p<-nlm(vero_p, 1, x=x)
```

```
lambda.est<-emv_p$estimate
```

```
#####Binomial
```

```
#### Definindo a Função de Verossimilhança
```

```
vero_b<-function(x,par)
```

```
{
```

```
  r<-par[1]
```

```
  p<-par[2]
```

```
  if ((r>0) && (p>0) && (p<1)) return (-1*sum(dnbinom(x,r,p,log=TRUE)))
```

```
  else return (-Inf)
```

```
}
```

```
#### Obtendo as Estimativas dos parâmetros
```

```
emv_b<-optim(c(1,.5),vero_b,x=x)
```

```
r.est<-emv_b$par[1]
```

```
p.est<-emv_b$par[2]
```

```
#####Plotar graficos#####
require(survival)
d<-rep(1,length(x))
km<-survfit(Surv(x,d)~1)
X<-km$time
F<-1-km$surv
m<-length(X)
XX<-seq(0,max(X))

# FUNÇÃO ACUMULADA

plot(X,F,type="s",ylim=c(0,1),xlab="Quantidade de Perdas",ylab="F(x)",
     main="Tipo de Evento A")

F.BN<-pnbinom(XX,r.est,p.est)
points(XX,F.BN,type="s",col=2)

F.Poisson<-ppois(XX,lambda.est)
points(XX,F.Poisson,type="s",col=4)

legend(20000,0.3,c("Empirica","Binomial Negativa","Poisson"),
      text.col=c(1,2,4),bty="n",cex=0.8)

### COMPARA ESTIMATIVAS
F.BNX<-pnbinom(X,r.est,p.est)
F.PoissonX<-ppois(X,lambda.est)
eqm_bn<-mean((F.BNX-F)^2)
eqm_p<-mean((F.PoissonX-F)^2)
```

A.1.2 Ajuste de Severidade

```
#####Weibull#####
```

```
#### Definindo a Função de Verossimilhança
```

```
vero.w<-function(x,parametro)
{
  alpha<-parametro[1]
  lambda<-parametro[2]
  if ((alpha>0)&&(lambda>0))
  return (-1*sum(dweibull(x,alpha,lambda,log=TRUE)))
  else return (-Inf)
}
```

```
#### Obtendo as Estimativas dos parâmetros
emv<-optim(c(1,1),vero.w,x=x)
alpha.est<-emv$par[1]
lambda.est<-emv$par[2]
```

```
##### Lognormal #####
```

```
vero.l<-function(x,par)
{
  alpha<-par[1]
  lambda<-par[2]
  if ((alpha>0)&&(lambda>0))
  return (-1*sum(dlnorm(x,alpha,lambda,log=TRUE)))
  else return (-Inf)
}
```

```
#### Obtendo as Estimativas dos parâmetros
emv.l<-optim(c(1,1),vero.l,x=xx)
alpha.estl<-emv.l$par[1]
lambda.estl<-emv.l$par[2]
```

```
##### Mistura de Lognormais #####
```

```
vero.ml<-function(x,par)
{
  alpha1<-par[1]
  lambda1<-par[2]
  alpha2<-par[3]
```

```

lambda2<-par[4]
p<-par[5]
if ((lambda1>0)&&(lambda2>0)&&(p>0)&&(p<1))
return (-1*(sum(log( p* (dnorm(x,alpha1,lambda1)) +
                    (1-p)*(dnorm(x,alpha2,lambda2))
                    )
                )
        )
    )
else return (-Inf)
}

#### Obtendo as Estimativas dos parâmetros
emv.ml<-optim(c(3,1,8,1,0.5),vero.ml,x=x, control = list(maxit = 1000))
alpha1.ml<-emv.ml$par[1]
lambda1.ml<-emv.ml$par[2]
alpha2.ml<-emv.ml$par[3]
lambda2.ml<-emv.ml$par[4]
p.ml<-emv.ml$par[5]

#####Plotar graficos#####
require(survival)
d<-rep(1,length(x))
km<-survfit(Surv(x,d)~1)
X<-km$time
F<-1-km$surv
m<-length(X)
XX<-seq(0.01,max(X),0.01)

# FUNÇÃO ACUMULADA

plot(X,F,type="s",ylim=c(0,1),xlab="log(Perdas)",ylab="F(x)",
main="Tipo de Evento A")

F.Weibull<-pweibull(XX,alpha.est,lambda.est)
points(XX,F.Weibull,type="l",col=2)

```

```

F.Normal<-pnorm(XX,alpha.est1,lambda.est1)
points(XX,F.Normal,type="l",col=4)

F.Mnormal<-( p.ml*pnorm(XX,alpha1.ml,lambda1.ml)
             + (1-p.ml)*pnorm(XX,alpha2.ml,lambda2.ml)
             )
points(XX,F.Mnormal,type="l",col=3)

legend(8,0.3,c("Weibull","Empirica","Normal","Mistura de Normais"),
text.col=c(2,1,4,3),bty="n",cex=0.8)

```

```

### COMPARA ESTIMATIVAS
F.W<-pweibull(X,alpha.est,lambda.est)
F.N<-pnorm(X,alpha.est1,lambda.est1)
F.MN<-(p.ml*pnorm(X,alpha1.ml,lambda1.ml)+
(1-p.ml)*pnorm(X,alpha2.ml,lambda2.ml))

eqm_w<-mean((F.W-F)^2)
eqm_n<-mean((F.N-F)^2)
eqm_mn<-mean((F.MN-F)^2)

```

A.2 Escalonamento e Filtro BDE

```

bde<-bde[which(bde$Perda>50000),]
dif_qt<-1.5*(as.numeric(Quartile(bde$Perda,probs=0.75))
            - as.numeric(Quartile(bde$Perda,probs=0.25)))
outl<-Quartile(bde$Perda,probs=0.75)+dif_qt

bde<-bde[which(bde$Perda<=outl),] #bde sem
#perda_cenario<-perda2[which(perda2>outl)]

perda<-log(bde$Perda)
tpext<-as.factor(bde$tp.ext)

z<-lm(perda~bde$Regiao+tpext)
summary(z)

```

```

Af<-as.numeric(exp(z$coeff[4])/exp(0))
Asia<-as.numeric(exp(z$coeff[4])/exp(z$coeff[2]))
Eur_L<-as.numeric(exp(z$coeff[4])/exp(z$coeff[3]))
Am_Lat<-as.numeric(exp(z$coeff[4])/exp(z$coeff[4]))
Am_N<-as.numeric(exp(z$coeff[4])/exp(z$coeff[5]))
Eur_W<-as.numeric(exp(z$coeff[4])/exp(z$coeff[6]))

bde<-transform(bde,perda_scl=bde$Perda)

bde[which(bde$Regiao=='Africa'),]$perda_scl<-
      bde[which(bde$Regiao=='Africa'),]$Perda*Af
bde[which(bde$Regiao=='Asia / Pacific'),]$perda_scl<-
      bde[which(bde$Regiao=='Asia / Pacific'),]$Perda*Asia
bde[which(bde$Regiao=='Eastern Europe'),]$perda_scl<-
      bde[which(bde$Regiao=='Eastern Europe'),]$Perda*Eur_L
bde[which(bde$Regiao=='Latin America & Caribbean'),]$perda_scl<-
      bde[which(bde$Regiao=='Latin America & Caribbean'),]$Perda*Am_Lat
bde[which(bde$Regiao=='North America'),]$perda_scl<-
      bde[which(bde$Regiao=='North America'),]$Perda*Am_N
bde[which(bde$Regiao=='Western Europe '),]$perda_scl<-
      bde[which(bde$Regiao=='Western Europe '),]$Perda*Eur_W

sevA_e<-bde[which(bde$tp.ext==A),11]
sevB_e<-bde[which(bde$tp.ext==B),11]
sevC_e<-bde[which(bde$tp.ext==C),11]

```

A.3 LDA

A.3.1 BDI

```

set.seed(555)

mlnorm<-function(n,mu1,sd1,mu2,sd2,p)
{
x<-rbinom(n,1,p)
ln1<-rlnorm(n,mu1,sd1)
ln2<-rlnorm(n,mu2,sd2)

```

```

x*ln1+(1-x)*ln2
}

M<-1000000
B<-100

amB<-dapB<-numeric(M)
qdapB_t1<-numeric(B)

for(j in 1:B) #parametrico
{
amB<-rnbinom(M, mu=185,6746, size=0.0030)
for(i in 1:M)
{
dapB[i]<-
sum(mlnorm(amB[i],mu1=4,9111,sd1=1.7995,mu2=10.1081,sd2=1.5653,p=0.8848))
cat(j,"-",i,"\n")
}
qdapB_t1[j]<-quantile(dapB, probs = 0.999)
cat(j,"-'",i,"\n")
}
VaR999<-mean(qdapB_t1)
c(quantile(qdapB_t1,probs=0.05),quantile(qdapB_t1,probs=0.95))
sqrt( (1/(B-1))* sum( qdapB_t1 - mean(qdapB_t1) )^2 )

```

A.3.2 BDI \cup BDE

```

mlnorm<-function(n,mu1,sd1,mu2,sd2,p)
{
x<-rbinom(n,1,p)
ln1<-rlnorm(n,mu1,sd1)
ln2<-rlnorm(n,mu2,sd2)
x*ln1+(1-x)*ln2
}

```

```

set.seed(555)
B<-100
N<-1000000
dapB<-numeric(N)
qdapB_e1<-numeric(B)

for(j in 1:B)#parametrico
{
for(i in 1:N)
{
am<-rbinom(1, 0.55,0.003)
dapB[i]<-sum(mlnorm(am, 5.11,1.97,11.57,0.73,0.8))
}
qdapB_e1[j]<-quantile(dapB, probs = 0.999)
cat(j, '- ', i, "\n")
}

var999_B<-mean(qdapB_e1)
c(quantile(qdapB_e1,probs=0.05),quantile(qdapB_e1,probs=0.95))
sqrt( (1/(B-1))* sum( qdapB_e1 - mean(qdapB_e1) )^2 )

```

A.3.3 $BDI \cap BDE$

```

mlnorm<-function(n,mu1,sd1,mu2,sd2,p)
{
x<-rbinom(n,1,p)
ln1<-rlnorm(n,mu1,sd1)
ln2<-rlnorm(n,mu2,sd2)
x*ln1+(1-x)*ln2
}

```

```

set.seed(555)
B<-100
N<-1000000
dap<-numeric(N)
qdap_e1<-numeric(B)

```



```
for(j in 1:B)
{
for(i in 1:N)
{
am<-rnbino(N, 5.02,0.0004)
II<-mnorm(am,mu1=2.91,sd1=1.09,mu2=6.59,sd2=1.28,p=0.19)
EE<-mnorm(am,mu1=11.21,sd1=0.21,mu2=11.85,sd2=0.44,p=0.62)
IE<-II
IE[which(II>50000)]<-EE[which(II>50000)]
dap[i]<-sum(IE)
}

qdap_e1[j]<-quantile(dap, probs = 0.999)
cat(j, '-',i, "\n")
}

var999_A<-mean(qdap_e1)
var999_A
c(quantile(qdap_e1,probs=0.05),quantile(qdap_e1,probs=0.95))
sqrt( (1/(B-1))* sum( qdap_e1 - mean(qdap_e1) )^2 )
```