

3

Base de Dados

Este capítulo é dedicado ao estudo da base de dados dos sistemas de medida e gerenciamento do risco operacional baseados na LDA.

A primeira etapa da construção da base de dados consiste na elaboração do modelo de dados baseado na definição e na estrutura de risco operacional apontadas pelo Novo Acordo da Basileia. Posteriormente são coletados dados consistentes e íntegros referentes aos metadados definidos no modelo de dados. A seguir, os dados adquiridos são armazenados no banco de dados, podendo antes necessitar de transformações para se tornarem compatíveis com as demais informações. Na última etapa os dados são analisados por especialistas por intermédio de gráficos, testes estatísticos e critérios que indiquem a validade das informações, e apontem os caminhos a serem tomados na modelagem estatística.

3.1.

Modelagem da Base de Dados

A modelagem da base de dados representa o processo pelo qual é concebido, analisado e construído o banco de dados de um sistema computacional. Logo, o modelo de dados está intimamente ligado aos requisitos que definem o sistema. No caso do risco operacional não é diferente. O modelo de dados depende da finalidade do sistema, que normalmente é destinado somente à medida, ou ao gerenciamento (qualitativo e quantitativo) do risco operacional.

Uma vez estabelecidos os objetivos do sistema, deve ser realizado um mapeamento dos riscos operacionais existentes na organização, criando um dicionário de dados que estabelece as definições de cada evento de perda. Esta tarefa é extremamente difícil devido à natureza ampla e multidisciplinar do risco operacional. Portanto, é de suma importância a participação dos especialistas de cada área envolvida.

Para facilitar a identificação dos riscos operacionais e sua gestão, o Novo Acordo distingue os riscos entre categorias de causas, eventos e efeitos. Esta abordagem permite uma atuação distinta sobre cada um desses componentes (Guimarães, 2003) [46].

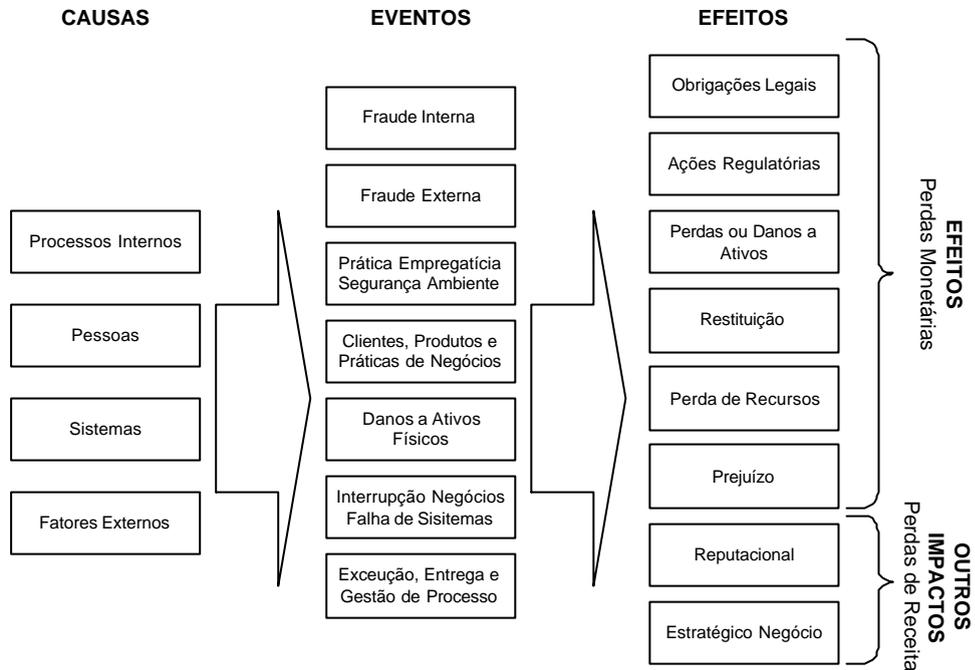


Figura 1 – Estrutura de Risco Operacional segundo o Novo Acordo da Basileia. Fonte: Guimarães (2003) [46].

A causa, ou tipo de evento de nível 0, é a ação ou conjunto de circunstâncias que levam a um evento. Ela pode ser tanto um fator interno (processos, pessoas e sistemas), quanto um fator externo (assaltos, ações da natureza etc), o que torna a gestão do risco operacional ainda mais complexa. A Figura 1 mostra as causas estipuladas pela Basileia para o risco operacional.

A identificação das causas é importante para o entendimento sobre os fatores que levam a materialização do risco operacional, fornecendo subsídios à criação de controles e testes³⁸. Os controles baseiam-se na experiência dos especialistas (gerência qualitativa), em modelos causais (gerência quantitativa) ou em ambos (utilizando técnicas bayesianas ou análise de cenários). As respostas destes controles atuam sobre os fatores causais com o objetivo de diminuir a probabilidade de ocorrência do risco operacional.

Os fatores causais são variáveis explicativas numéricas de dois tipos possíveis:

³⁸ Teste são procedimentos de averiguação do funcionamento dos controles.

- CEF (*Control Environment Factors*): *Proxies* da qualidade dos processos e controles internos do negócio. Explicam diretamente a quantidade ou o tamanho das perdas. Exemplo: tempo em que um sistema fica fora do ar.
- KRI (Key Risk Indicators): Indicadores das perdas ocorridas representando indícios da qualidade dos processos e controles internos do negócio. Não explicam diretamente a quantidade e nem o tamanho das perdas. Por exemplo: número de quedas na conexão entre os sistemas.

Conforme dito anteriormente, a finalidade do sistema definirá a necessidade ou não do modelo de dados apresentar campos referentes aos fatores causais. Neste caso, a vantagem de se trabalhar com metadados³⁹ é a possibilidade de se definir em tempo de execução quais os fatores farão parte do modelo causal.

Um evento de risco operacional pode ter várias causas associadas, por isso, além da categorização por causas, é necessária uma categorização por tipos de eventos. O Novo Acordo da Basileia define categorias de tipos de eventos em três níveis hierarquizados para classificação do risco operacional⁴⁰. A Figura 1 mostra os sete tipos de evento do nível 1. O nível 2 detalha um pouco mais o nível 1, sendo seus tipos também padronizados pelo Novo Acordo. O nível 3 apresenta exemplos de possíveis atividades referentes ao nível 2, podendo cada instituição acrescentar mais exemplos de acordo com suas realidades. Cada tipo de evento de nível 3 é materializado por intermédio das informações transacionais coletadas pela instituição. Essa hierarquização, além facilitar o processo de modelagem e de coleta de dados do risco operacional, auxilia no processo de análise do comportamento de cada uma dessas categorias, cujas distribuições de probabilidade são diferentes entre si. Portanto, basta coletar, tratar e armazenar os dados da transação correspondente a um evento de nível 3, para se obter informações nos demais níveis de tipos de eventos (Guimarães, 2003) [46].

³⁹ Literalmente significa dados sobre os dados. Na prática, modelar uma base de dados baseada em metadados é passar para os registros as informações que em um modelo normalizado estariam em campos. Isso flexibiliza o crescimento conceitual da base de dados em tempo de execução, sem necessitar de mudanças no modelo físico.

⁴⁰ Para maiores detalhes consultar anexo 7 em [12].

O efeito é a perda monetária ou não-monetária acarretada pelo evento. Perdas não-monetárias são perdas indiretas^{41,42}, e, portanto, complicadas de serem quantificadas. Conforme visto na definição de risco operacional, somente as perdas diretas⁴³ são consideradas pelo Novo Acordo.

Mesmo assim, nem sempre considerar somente perdas diretas torna mais fácil a tarefa da manutenção do banco de dados. Existem eventos cuja duração⁴⁴ é maior (dias, meses ou até anos). Nesses casos, deve-se estimar um valor para o efeito com a ajuda de um especialista. Este valor estimado é chamado de perda potencial (em unidades monetárias). Para auxiliar a tarefa de estimação, podem ser utilizados modelos causais baseados nos dados históricos. Neste caso, o modelo de dados deverá conter as variáveis explicativas daqueles modelos causais.

A classificação dos tipos de efeitos esclarece quais são os tipos de perdas⁴⁵ decorrentes de risco operacional que afetam os resultados da instituição. A identificação dos efeitos auxilia, principalmente no processo de criação de contingências financeiras como o estabelecimento provisões e a utilização de seguros.

Além da classificação por categorias de causas, eventos e efeitos, o Novo Acordo da Basileia define um mapeamento das linhas de negócio em dois níveis hierarquizados como forma de padronização⁴⁶. Isto ocorre, pois os perfis de risco das instituições financeiras variam com seus tipos de negócios.

Portanto, o Comitê da Basileia determina que a base de dados armazene as informações que possibilitem cálculos de capitais alocados para proteção contra o risco operacional separados por unidade de negócio, e por potencial atividade de perda. Neste contexto, o conceito de *Bottom Up*⁴⁷ é essencial para se obter o entendimento do ambiente de risco. A agregação para cada nível

⁴¹ São perdas que impactam indiretamente no resultado. Por exemplo: perdas acarretadas por depreciação da imagem.

⁴² Estas perdas podem ser consideradas em um modelo interno, desde que existam procedimentos para mensurá-las, independente do fato da não-exigência regulatória. Neste caso, as causas, os eventos e os efeitos destas perdas devem ser capturados.

⁴³ São perdas que impactam diretamente no resultado. Por exemplo: perdas por fraudes internas.

⁴⁴ Intervalo de tempo entre a concepção do evento até a sua conclusão, momento aonde se pode ter certeza sobre o valor monetário do impacto (efeito).

⁴⁵ Dado um conceito de concretização de risco voltado para a variabilidade, ganho pode ser considerado como um tipo de perda retificadora, uma vez que provem de um erro operacional.

⁴⁶ Para maiores detalhes consultar anexo 6 em [12].

⁴⁷ Este é um dos principais conceitos por de trás do modelo *Mark-to-Future* da empresa Algorithmics, conforme será visto no capítulo 4. Informações mais detalhadas sobre este modelo podem ser vista em (Reynolds & Syer, 2002) [77].

acima, tanto nos tipos de evento quanto nas linhas de negócio, trará informações gerenciais para toda a empresa.

Estas informações gerenciais são potencializadas quando os dados de perdas se relacionam com os planos de ação para mitigação dos riscos, os planos de contingência, os produtos, os processos, os controles internos, as localidades geográficas, as unidades organizacionais e seus respectivos custos. Assim torna-se possível realizar análises de custo e benefício, e medir a eficiência e a eficácia das tomadas de decisão.

Estes foram os principais pontos abordados na modelagem de dados. Entretanto, as informações que serão vista adiante, justificam alguns ajustes no modelo de dados final.

3.2.

Coleta de Dados

A coleta de dados é um processo constitutivo que traduz perigo em risco para tomada de decisão. A fidelidade desta tradução depende da qualidade e da quantidade de dados.

A qualidade é importante uma vez que não adianta ter um bom modelo ajustado aos dados errados. Por isso, o Comitê de Supervisão Bancária da Basileia exige que a base de dados, a partir da qual será calculada a alocação de capital, seja constantemente avaliada quanto à integridade e confiabilidade de seus dados. Esta preocupação está associada, principalmente à credibilidade do modelo a ser implementado pelos bancos (Guimarães, 2003) [46].

Já a quantidade possibilita um tratamento estatístico adequado⁴⁸.

Logo, é vital que os dados sejam coletados corretamente não só pelos acidentes e incidentes ocorridos, mas também por auditorias internas e avaliações da gerencia de risco operacional.

Atualmente, existem 4 formas alternativas de se coletar as perdas decorrentes de risco operacional:

- Registrar manualmente as ocorrências de risco operacional à medida que os eventos de perda ocorrem. Para isso é necessário um sistema de cadastro de perdas, tendo usuários capacitados a identificar e registrar os mais variados tipos;

⁴⁸ Cuidado: uma informação estatisticamente significativa pode não ser relevante ou útil para o modelo.

- Coletar os dados de fontes gerenciais, pois determinados tipos de perdas já dispõem de acompanhamento gerencial (exemplo: fraudes com cartão de crédito). Muitas dessas fontes já estão informatizadas, o que possibilita a extração das perdas por intermédio de regras automatizadas;
- Buscar nos registros contábeis as perdas decorrentes de eventos de risco operacional – exemplo: multas (Guimarães, 2003) [46];
- Auto-avaliar com auxílio de questionários e de técnicas *Delphi*⁴⁹, quando não se têm os dados históricos (Coleman, 2000; Cruz, 2002) [20,26].

Forma de Coleta	Vantagens	Desvantagens
Manual	É possível ter um maior detalhamento da perda, como a identificação precisa das causas, efeitos e ações mitigadoras.	Requer muito recurso humano e treinamento para atingir a cobertura total das perdas.
Gerencial	O esforço humano é menor, e é possível adaptar a informação para a taxonomia de riscos operacionais.	A informação é menos confiável do que a contábil. O volume de dados é muito alto. Os formatos dos dados são diferentes. Os locais de armazenamento são distribuídos.
Contábil	O esforço humano é mínimo e a confiabilidade deste tipo de informação é muito maior do que as outras.	Ainda não existem contas específicas para este tipo de perda e é difícil identificá-las dentro da taxonomia de riscos operacionais.
Auto-avaliação	Não há a necessidade de dados históricos.	Requer informações de vários especialistas das mais diversas áreas de negócio, devido à característica multidisciplinar e ampla do risco operacional.

Tabela 3 – Procedimentos de coleta de dados de perdas operacionais. Fonte adaptada: Guimarães (2003) [46]

Estes tipos de coleta não são mutuamente excludentes. Logo, há a necessidade de apoio tecnológico para garantir a unificação dos dados, mas de forma a possibilitar a execução de consultas em várias dimensões, tal como na utilização de *data warehouse* (Middlemiss, 2003) [65].

⁴⁹ Para maiores informações sobre auto-avaliação consulte o apêndice B do capítulo 2 de (Cruz, 2002) [26].

Todavia, existem grandes problemas na coleta de eventos de risco operacional dentro das instituições financeiras. Dentre esses vários obstáculos destacam-se:

- Eventos de risco operacional têm sua natureza descentralizada, ou seja, este tipo de risco pode se materializar em qualquer unidade de negócio, podendo estar relacionado a qualquer processo, produto ou serviço (Guimarães, 2003) [46].
- Há uma carência cultural no processo de coleta deste tipo de perda, principalmente pelo fato de ser uma informação sensível, e ainda não haver um claro entendimento dos benefícios, como o ganho de eficiência que este rastreamento de perdas pode trazer (Guimarães, 2003) [46].
- Falta de padrão na identificação⁵⁰, e na classificação dos dados de perdas, embora o Comitê da Basileia tenha minimizado sobremaneira esse problema. A assunção dos padrões do Novo Acordo da Basileia tem ajudado na melhoria da consistência da coleta de dados nas organizações e entre as organizações bancárias. Isto possibilita que os bancos comparem suas experiências de perdas através das linhas de negócio, e os supervisores as experiências de perdas entre os bancos (de Fontnouvelle et al., 2003) [30]. Além disso, permite também que os vendedores de base de dados externas possam levantar os fatos estilizados que auxiliam na elaboração de modelos para a exposição ao risco. O conceito de base de dados externa será visto adiante.
- A má classificação das perdas gera resultados suspeitos na análise quantitativa. Os principais exemplos de erros na classificação das perdas são: a sobreposição de categorias para um mesmo evento, e o vínculo equivocado entre o evento e seu respectivo impacto monetário⁵¹. Entretanto, o Comitê da Basileia (2001a; 2004) [6,12] reconhece a dificuldade de se classificar com clareza os eventos de

⁵⁰ O Novo Acordo da Basileia não define com clareza o que é um evento de perda operacional. Certamente, este evento deve se distinguir de um custo operacional normal, mas o que separa tais conceitos? Além disso, as perdas operacionais devem se diferenciar também das perdas já levadas em consideração pelos riscos de mercado e de crédito (Nyström & Skoglund, 2002) [68].

⁵¹ Exemplos numéricos para estes dois tipos de erros de classificação são apresentados no item 5.1 de Álvarez (2001) [1].

perda, e já conta com certa ambigüidade. Portanto, independente do padrão adotado pelos bancos, o esquema de classificação das perdas deve ser lógico, consistente, com categorias mutuamente exclusivas, e coerência entre as definições de perda e seus respectivos impactos monetários (Álvarez, 2001) [1].

- Muitas das vezes o que se percebe é o efeito ocasionado por um ou mais eventos de perda, inclusive com uma defasagem de tempo considerável. Nestes casos, pode ser extremamente complicado levantar todos os dados relacionados a este(s) evento(s), sendo necessário o preenchimento dos dados faltantes segundo algum critério pré-estabelecido.
- Conforme mencionado anteriormente, uma duração de evento mais longa acarreta a necessidade da estimação da perda potencial. Isso dificulta a coleta, pois, além dos dados do evento em si, dados para o modelo causal de previsão monetária da perda potencial devem ser capturados.
- Normalmente no processo de coleta de dados, existe um limite mínimo para o impacto financeiro do evento de perda coletado chamado limiar. Isto acontece, pois é muito custoso coletar e armazenar toda e qualquer perda ocorrida. Além disso, nos estudos dos modelos estatísticos do método das distribuições de perdas, a preocupação maior recai sobre os eventos de perda mais severos, isto é, perdas cujo impacto financeiros são maiores. Será visto mais adiante que para melhorar o gerenciamento do risco operacional devem-se considerar também as perdas pouco severas, porém muito freqüentes chamadas de *Near-Miss*. Isso aumenta e muito a quantidade de dados a ser coletada. O capítulo 4 mostrará a influência deste limiar no modelo estatístico.
- Particularmente para trabalhos acadêmicos na área de risco operacional, o principal problema na aquisição dos dados é a confidencialidade. A divulgação dos dados de perdas operacionais pelos bancos ainda é considerada restrita ao público, devido à importância estratégica desses dados. Isso desestimula os trabalhos acadêmicos na área, e explica o fato da maioria dos estudos realizados terem sido feitos no âmbito das instituições financeiras por suas áreas de pesquisa e desenvolvimento. Para poder publicar os resultados práticos de tais pesquisas, geralmente

os dados são mascarados, sem, contudo, perder a proporcionalidade entre eles. Uma solução bastante utilizada é tornar os dados adimensionais, dividindo-os por um indicador de exposição ao risco, como por exemplo, a receita bruta da área de negócio em questão. Para garantir a confidencialidade, o valor deste indicador não deve ser apresentado. Esta solução além de resolver o problema de sigilo, torna os resultados e conclusões passíveis de comparação entre instituições financeiras afins. Com a disciplina de mercado, descrita no Terceiro Pilar do Novo Acordo da Basileia, espera-se que esta barreira seja transposta.

Além das dificuldades no processo de coleta de dados de perdas operacionais, os bancos enfrentam problemas na utilização da base de dados histórica:

- Baixa quantidade e qualidade de dados de perdas operacionais pré-existentes nos bancos, independente dos processos de coletas atuais estarem aderentes às exigências da Basileia. Isso se deve a incipiência dessa disciplina de risco no mercado financeiro, levando os bancos a terem de garimpar os dados mais antigos para algumas linhas de negócios e/ou tipos de eventos.
- Especialistas questionam o uso da base de dados histórica como único sinal da organização empresarial, uma vez que os controles internos estão sempre evoluindo à luz da experiência sobre as perdas. Logo, as medidas das distribuições de probabilidade podem ser influenciadas pelo gerenciamento, sendo mais comportamentais que somente técnicas (Power, 2003) [75].
- Conforme será visto no capítulo 4, a distribuição de severidade das perdas operacionais geralmente possui caudas pesadas, isto é, eventos de baixa frequência e alto impacto. Porém, a situação atual nos bancos é que os dados antigos representando estes eventos não foram registrados por razões óbvias. Isso torna a base de dados histórica viesada em direção a perdas de baixa severidade, dificultando assim a estimação adequada dos parâmetros que definem a distribuição da população, e subestimando o capital alocado para o risco operacional (Frachot & Roncalli, 2002) [43].

Para resolver estes problemas, os bancos utilizam bases de dados externas (pública ou consórcio), e bases de dados geradas por análises de

cenários (artificial), complementando assim os dados históricos. Estas bases de dados são classificadas de acordo com a fonte dos dados da seguinte maneira:

- Pública⁵²: baseada em dados de perdas publicamente divulgadas. A maioria dessas perdas é muito severa, uma vez que é muito difícil esconder impactos extremamente altos da mídia (Baud et al., 2002) [14].
- Consórcio⁵³: baseada em dados de perdas coletados por um consórcio de bancos. Funciona por um acordo de confidencialidade entre um conjunto de bancos que se comprometem em alimentar uma base de dados com suas perdas internas. Em contrapartida, os bancos consorciados podem utilizar estes dados para complementar seus dados internos (Baud et al., 2002a) [14].
- Artificial: baseada em dados de perdas simuladas a partir da análise de cenários gerados artificialmente através de auto-avaliação. Quando esta base de dados artificial é gerada por intermédio de modelos causais que explicam razoavelmente bem a situação dos dados de perdas internas, estes dados são ditos efetivos para captura de riscos operacionais potenciais (Hidetoshi, 2002) [51].

Entretanto, existem algumas questões a se considerar, quando se pretende mesclar dados internos com dados externos e/ou dados simulados:

- Como tornar compatíveis para mescla perdas externas e internas com diferenças de severidade muito acentuadas? Provavelmente, esta avaliação dependerá de diferentes fatores tais como: número de empregados, capital econômico, tamanho dos ativos, peso de cada negócio etc. Porém, é muito difícil apresentar um método de escalonamento razoável. Uma idéia simples⁵⁴ é tornar as perdas internas e externas adimensionais, dividindo-as por seus respectivos

⁵² Como exemplo de produtos de base de dados externos públicos destacam-se: *OpRisk Analytics' OpRisk Global Data* do SAS Institute, Inc., e *OpVantage First* do Fitch Risk Management. Ambos garantem que coletam somente dados de perda disponíveis em fontes públicas, e que excedam US \$1 milhão. Além disso, as perdas seguem a classificação de linhas de negócio e eventos previstos pela Basileia, sendo que cada perda tem uma explicação descritiva. Outro exemplo: a base de dados de perdas operacionais CORE (Compilation of Operational Risk Events) que foi elaborada pela empresa ORI (Operational Risk, Inc.) para os gerentes de risco entenderem e avaliarem riscos operacionais que ocorrem externamente.

⁵³ Exemplo de produtos de base de dados externos de consórcio: Gold of BBA (*British Bankers' Association*).

⁵⁴ Outras soluções mais complexas são indicadas no item 3.4 de Chapelle et al. (2004) [18].

indicadores de exposição ao risco – exemplo: receita bruta (Hidetoshi, 2002) [51].

- Mais ainda: será que as bases de dados externas representam bons indicadores em relação aos controles internos dos bancos (Cruz, 2002) [26]? Com certeza, bases de dados de bancos que estão sediados em diferentes países, com culturas e leis específicas, níveis de desenvolvimento econômico e político diferenciado, linhas de negócio distintas e etc não serão bons *proxies* umas das outras. Portanto, também é muito importante avaliar qualitativamente o grau de similaridade entre os bancos que são fontes de dados destas bases.
- Quando dados simulados em cenários de estresse são adicionados aos dados internos, a transparência da auditoria interna e da avaliação destes cenários é indispensável. Mais ainda, deve-se prestar atenção ao intervalo de tempo entre a auditoria e a avaliação, e a objetividade da criação do cenário de perda (Hidetoshi, 2002) [51].
- Como auditar dados simulados baseados em análise de cenários elaborados por especialistas? Neste caso, o máximo que a supervisão poderá exigir é uma documentação consistente que comprove os vínculos dos dados com os cenários. Caso estes cenários sejam gerados por intermédio de modelos causais, a supervisão poderá ir mais além, executando avaliações quantitativas.
- Em relação aos dados externos públicos, a raridade de alguns eventos pode estar associada à não divulgação de perdas destes tipos, ou a divulgação das perdas sem os respectivos impactos, ou ainda a erros de classificação quando do processo de coleta destes dados.
- Já que a maioria das perdas da base de dados pública é muito severa, espere-se um limiar maior para este tipo de base que para a base de dados de consórcio (Baud et al., 2002a) [14]. Esta informação é extremamente útil para decisão do método de mescla dos dados internos e externos (Chapelle et al., 2004) [18].
- O grau de incerteza em relação ao limiar da base de dados de consórcio é muito grande, uma vez que cada consorciado pode não seguir a risca o limiar pré-estabelecido no acordo. Neste caso, o limiar da base de dados do consórcio deve ser considerado como uma

variável aleatória ou um vetor aleatório⁵⁵ (Baud et al., 2002a; Frachot & Roncalli, 2002) [14,43].

- A simples junção das bases de dados internas e externas, sem o tratamento estatístico adequado⁵⁶, acarreta na estimação de parâmetros viesados para a distribuição subjacente⁵⁷, e na superestimação do capital alocado para o risco operacional⁵⁸. Isto se deve aos diferentes limiares das bases de dados, aliados aos seus graus de incerteza (Baud et al., 2002a; de Fontnouvelle et al., 2003) [14,30].

Uma vez tomadas às devidas ações em relação às questões supracitadas, a complementação das bases de dados internas por bases de dados externas e/ou simuladas apresenta algumas vantagens⁵⁹:

- Reduz o capital alocado para o risco operacional, caso o tratamento estatístico adequado seja feito (de Fontnouvelle et al., 2003) [30].
- Aumenta a quantidade de eventos extremos, quer seja pelas experiências de outros bancos (dados externos), quer seja pela captura de perdas potenciais em cenários de estresse simulados por especialistas (dados artificiais), tornando a estimação da cauda da distribuição de severidade mais precisa (Baud et al., 2002a; Frachot et al., 2001; Frachot & Roncalli, 2002) [14,42,43].
- Possibilita o “envelhecimento” do portfólio de eventos de perdas operacionais. Isto significa que a estimação da distribuição de perdas um ano à frente, conforme estabelecido no Novo Acordo para o cálculo do VaR, não precisa levar em consideração somente os eventos *ex-post*⁶⁰. Assim, por intermédio de avaliações de especialistas nos controles internos *ex-ante*⁶¹, os dados artificiais são gerados, complementando os dados internos.

⁵⁵ Detalhes técnicos serão vistos no capítulo 4.

⁵⁶ Este tratamento estatístico será abordado com detalhes no capítulo 4.

⁵⁷ A distribuição subjacente é a distribuição de probabilidade que melhor se ajusta aos dados da amostra. Neste caso, a hipótese é que a amostra representa as características da população. Logo, infere-se que esta distribuição é a da população, e, por conseguinte, os parâmetros desta distribuição são não-viesados. Quando as bases de dados estão truncadas, e sem nenhum tratamento, a hipótese anterior é falsa. Portanto, os parâmetros da distribuição são viesados.

⁵⁸ Isto será visto no capítulo 4. Mais detalhes em Baud et al. (2002a) [14].

⁵⁹ Uma explicação detalhada destas vantagens será vista no capítulo 4.

⁶⁰ Eventos que já ocorreram.

⁶¹ Controles responsáveis por mitigar eventos que ainda não ocorreram.

- Garante que a distribuição da frequência dos eventos de perdas operacionais esteja vinculada à competência gerencial da instituição, e não a sorte, ou ao azar, da ocorrência destes eventos⁶² (Frachot & Roncalli, 2002) [43].

Por isso, o Novo Acordo da Basileia impõe os seguintes padrões quantitativos para os bancos que adotarem a AMA, no que se refere aos tópicos dados internos, dados externos e análise de cenário⁶³:

- Utilizar uma base de dados de perdas com um histórico de cinco anos, sendo que um mínimo de três anos será aceito no início da vigência do Novo Acordo.
- Considerar também em seu modelo de perdas a utilização de dados externos em conjunto com *Expert Opinions*, especialmente nos casos em que há razões para acreditar que o banco está exposto a perdas de baixa frequência e alta severidade.

Uma outra solução para o problema de baixa quantidade de dados é a simulação de dados baseados em modelos ajustados, onde o *bootstrap*⁶⁴ é o principal exemplo de processo de geração destes dados, bastando os dados internos para implementar esta solução.

Embora uma atenção maior seja depositada nas perdas pouco frequentes e muito severas, pois são estas que modelam a cauda da distribuição subjacente, as perdas muito frequentes e pouco severas, denominadas *Near-Miss*, apresentam um papel estratégico para o gerenciamento de risco operacional⁶⁵. Este tipo de dado já é coletado nas indústrias aeronáuticas, automobilísticas e químicas, aonde impactos mais severos podem até causar mortes de pessoas. As *Near-Misses* indicam que os controles internos não estão funcionando bem, e que a qualquer hora o impacto das perdas aumentará, podendo chegar até a “morte” de pessoas jurídicas, no caso do mercado financeiro.

Logo, o limiar está intimamente relacionado a *Near-Miss*, pois é ele quem define o limite inferior da severidade da perda coletada.

⁶² Isto será visto em detalhes na seção 4.3.3.1.2

⁶³ Estes padrões são encontrados em [12] nos parágrafos 670-675.

⁶⁴ Ajusta-se uma distribuição aos dados. Obtem-se dados simulados baseados nesta distribuição. Os dados originais são complementados por estes dados gerados. A partir daí o processo se repete. Mais informações, basta consultar o apêndice A3 de Coleman (2003b) [24].

⁶⁵ Para obter mais informações sobre o gerenciamento de risco operacional utilizando *Near-Miss* basta consultar Muermann & Oktem (2002) [67].

O processo de coleta de dados torna-se mais caro e difícil à medida que o limiar diminui. Por outro lado, um limiar menor garante mais dados coletados, aumentando a acurácia da medida de risco, e a quantidade de *Near-Misses* coletadas. Portanto, cabe uma avaliação de custo e benefício para se determinar o limiar ideal para cada instituição (Frachot et al., 2001) [42].

Em um primeiro momento, as *Near-Misses* não são desejadas por aumentarem o capital requerido. Inicialmente, os advogados e os especialistas em RH não as desincentivam, porém o departamento de risco operacional tenta escondê-las. O real incentivo em coletá-las está em um melhor gerenciamento e controle do risco operacional (Coleman, 2003b) [24].

3.3.

Tratamento e Armazenamento

Os dados coletados podem necessitar de tratamento ao longo do tempo para retificação e compatibilização das informações de perdas operacionais. Por exemplo:

- Tornar a severidade adimensional: mudança de escala dividindo os impactos das perdas por um indicador de exposição de cada linha de negócio (por exemplo: receita bruta). Motivos:
 - Não é possível tornar pública a imagem das perdas do banco devido a confidencialidade dos dados.
 - É necessário para entender a severidade do evento de perda, e tornar possível uma comparação do impacto entre linhas de negócios diferentes em um banco. Isso vale também para o indicador de exposição do banco como um todo, permitindo a comparação entre bancos.
- Inflação e correção monetária: trazer a valor presente os dados monetários. Isso garante que as medidas e controles baseados nos dados armazenados estão atualizados em relação ao poder de compra real da moeda utilizada pela instituição financeira.
- Ajustes nos valores das perdas potenciais: com o passar do tempo, as perdas potenciais têm seus valores estabelecidos. Logo, além do ajuste necessário para traduzir o real impacto da perda, é possível avaliar a estimativa das perdas potenciais, visando a melhoria contínua dos modelos de previsão.

- Ajustes nos valores de perdas recuperadas: a gerência de risco operacional busca minimizar o impacto financeiro dos eventos de perda ocorridos, tentando recuperação do valor monetário da perda no todo ou em parte. Portanto, quando uma perda é recuperada o devido ajuste deve se feito.
- Ajustes nos valores de perdas restituídas: sempre que houver restituição de uma perda, esta deve ser registrada, e o dado original ajustado.
- Dados faltantes (*Censored*): muitas das vezes dispõe-se do número de observações de um evento de perda, e o valor total do impacto financeiro dessas perdas, mas não se conhece o valor de cada observação. Neste caso, pode ser necessária a elaboração de um processo de geração desses eventos de perda. Caso não seja possível, é extremamente importante indicar na base estes dados agregados, pois existe um tratamento estatístico adequado a ser feito para a estimação da distribuição da severidade⁶⁶.

Vale ressaltar que todas as transformações devem ser registradas, formando assim um histórico das mudanças sofridas pelos dados. Isso é importante para auditoria, avaliação das estimativas, estudo de sensibilidade, avaliação da taxa de recuperação das perdas etc.

3.4.

Análise dos Dados

Análises exploratórias nos dados de perdas podem ser realizadas, após a coleta, o tratamento e o armazenamento, pelos seguintes motivos:

- Auditoria interna: garantia da qualidade dos dados, evitando sanções regulamentares e aumentando o grau de confiabilidade dos modelos internos.
- Supervisão regulatória (auditoria externa): conforme estabelece o Segundo Pilar do Novo Acordo da Basileia. Os supervisores devem elaborar métodos e procedimentos para atestarem a qualidade dos dados, uma vez que estes são os insumos básicos para as medidas de risco operacional.

⁶⁶ Detalhes técnicos deste método serão vistos no capítulo 4.

- Indicação preliminar dos modelos estatísticos: comparando os fatos estilizados levantados pela comunidade de risco operacional e as informações da análise exploratória dos dados de perdas, verificam-se os candidatos a modelos estatísticos de medida e controle.

Estas são algumas das ferramentas utilizadas para esta análise:

- Inspeção dos dados por especialistas: visa depurar a base, confrontando as informações dos dados com os fatos estilizados baseados na experiência dos especialistas. A maior dificuldade desse tipo de tarefa é a característica multidisciplinar do risco operacional, e a avaliação qualitativa empregada.
- Modelos de validação dos dados: métodos e procedimentos quantitativos utilizados para identificar contradições, informações faltantes e não-conformidades que definam a má qualidade dos dados. Apesar da vantagem do uso de processos automatizados, estes modelos normalmente não são gerais, tendo de ser criados caso a caso.
- Estatísticas descritivas: cálculo de estatísticas descritivas sobre as ocorrências, as severidades, os CEF's e os KRI's das perdas operacionais por níveis de eventos, níveis de linhas de negócio, produtos, localizações geográficas, unidades organizacionais etc. Estas estatísticas indicam as características básicas das distribuições subjacentes dos dados analisados. Por exemplo, para a maioria dos eventos de perda operacional a distribuição empírica tem caudas pesadas, isto é, uma grande quantidade de perdas com valores baixos, e uma pequena quantidade de perdas com valores altos. Logo, um excesso de curtose, que é uma estatística descritiva básica, rapidamente indicaria este fato.

Portanto, a análise dos dados subsidia o conhecimento do comportamento da amostra, extraindo indicações sobre que modelos estatísticos e distribuições subjacentes utilizar para a medida e controle do risco operacional. O próximo capítulo é dedicado ao modelo de distribuição de perdas – o modelo de medida do risco operacional estudado neste trabalho, e o posterior, a modelagem causal responsável pelo controle, peça fundamental para o gerenciamento do risco operacional.