

UNIVERSIDADE DO VALE DO RIO DOS SINOS — UNISINOS  
UNIDADE ACADÊMICA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO E  
SISTEMAS  
NÍVEL MESTRADO

DAVENILCIO LUIZ DE. SOUZA

AVALIAÇÃO DA EFETIVIDADE DE CARTAS DE CONTROLE  
MULTIVARIADAS NA DETECÇÃO DE SUSPEITAS DE FRAUDE FINANCEIRA

SÃO LEOPOLDO  
2017

Davenilcio Luiz de. Souza

AVALIAÇÃO DA EFETIVIDADE DE CARTAS DE CONTROLE  
MULTIVARIADAS NA DETECÇÃO DE SUSPEITAS DE FRAUDE FINANCEIRA

Dissertação apresentada como requisito  
parcial para a obtenção do grau de Mestre  
pelo Programa de Pós-Graduação em  
Engenharia de Produção e Sistemas da  
Universidade do Vale do Rio dos Sinos —  
UNISINOS

Orientador:  
Prof. Dr. André L. Korzenowski

São Leopoldo  
2017

S729a

Souza, Davenilcio Luiz de.

Avaliação da efetividade de cartas de controle multivariadas na detecção de suspeitas de fraude financeira / Davenilcio Luiz de Souza — 2017.

99 f.: il.; 30 cm.

Dissertação (mestrado) — Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Engenharia de Produção e Sistemas, São Leopoldo, 2017.

“Orientador: Prof. Dr. André L. Korzenowski.”

1. Suspeitas de fraude financeira. 2. Lavagem de dinheiro. 3. Carta de Controle Multivariada. 4. Lei de Benford. 5. Controle Estatístico de Processos. I. Título.

CDU 658.5

Dados Internacionais de Catalogação na publicação (CIP)

(Bibliotecário Responsável: Flávio Nunes — CRB 10/1298)

Esta dissertação é uma homenagem à Família, porto seguro em todos momentos.  
Nailde Lima de Almeida, Marcella Almeida de Souza e Roberta Almeida de Souza

*Se soubéssemos o que era aquilo que estávamos fazendo, não seria chamado de  
pesquisa.*

*Quanto mais me aprofundo na Ciência mais me aproximo de Deus.*

— ALBERT EINSTEIN

## AGRADECIMENTOS

A benção de estar vigiado por um Ser que criou o universo e as maravilhas que dele provêm. A família sempre presente, esposa Nailde Almeida de Souza e as filhas Marcella Almeida de Souza e Roberta Almeida de Souza que em todos os momentos foram incansáveis na trajetória da execução desta dissertação, apoiando e reconhecendo que o crescimento se concretiza pelo trabalho, abnegação e a compreensão das ausências necessárias.

Um período de novos saberes e vocês ficaram para sempre, uma relação construída ao longo de 2015 e 2016, que Deus abençoe a todos. Uma dádiva em cada pessoa. André Korzenowski mais que orientador, um amigo. Mestrandos, Ana Herrera, Pablo Fogaça, Rossano Noronha, Wagner Simões, parceiros de aula e da ambrosia. Dos amigos de classe e de estudos, Elisandro Vargas, Eduardo Batista, Lenon Pinheiro, Rafael Herzer, Christopher Pohlmann e demais Mestrandos de 2015, e os Galáticos, como esquecer?

Aos Professores, Claudia Viegas, Giancarlo Pereira, Guilherme Vaccaro, Leonardo D. Chiwiacowsky, Mirian Borchardt, Marcos Souza, ao Coordenador do PPGEPS, Professor Miguel Sellitto e aos demais professores do Programa.

Ao pessoal da secretaria que esteve sempre solícito nas demandas dos Mestrandos, em especial a Lilian Amorim.

Meus sinceros agradecimentos a todos, pelos momentos de troca e crescimento que construímos ao longo destes dois anos.

## RESUMO

Os crimes de lavagem de dinheiro têm provocado grandes perdas aos países e a seus sistemas financeiros, o volume de dados em transações digitais representa dificuldade para a detecção deste tipo de ilícito. As auditorias em dados financeiros mostram-se limitadas na identificação de fraudes, pois em grande parte, ainda são realizadas com dados coletados por amostragem e incapazes de identificar as situações de delito em tempo real. Este trabalho, visando auxiliar no atendimento a esta lacuna, tem por objetivo propor um método estatístico de monitoramento por Cartas de Controle multivariadas, com base na Lei de Benford, para a detecção de suspeitas de fraude em lançamentos financeiros, entre eles os devidos à lavagem de dinheiro. Foi definido um modelo conceitual com distribuição de probabilidades representando dados oriundos de lançamentos financeiros, e adotada a suposição de que aderem a distribuição da Lei de Benford. Posteriormente foi considerada a distribuição empírica, estimada a partir dos próprios dados e dois procedimentos foram testados para verificar as suspeitas de fraude por lavagem de dinheiro utilizando a avaliação dos primeiros dígitos significativos: A Carta de Controle multivariada  $\chi^2$  e a Carta de Controle multivariada  $T^2$  de Hotelling. Foram simulados dados com auxílio do *software R-Project* até a ocorrência do 50.000º sinal. Foram avaliados casos simulados e reais, com o fim de exemplificar a operação do método. A partir da simulação, as duas Cartas de Controle testadas foram avaliadas quanto ao *ARL*, isto é, o número médio de observações até sinalizar que a série passou a operar em um estado fora de controle, o que significa a suspeita de lançamentos fraudulentos. Após aplicação do método de análise retrospectiva, com base nas proporções dos primeiros dígitos de Benford em lançamentos financeiros da campanha para Prefeito em 2016, não foram evidenciadas suspeitas de fraude nos dados obtidos junto ao sítio do Tribunal Superior Eleitoral (TSE). Em um conjunto de dados de uma instituição financeira, foram observados sinais de divergência entre as frequências dos primeiros dígitos nos lançamentos e nos valores esperados, porém os pontos além dos limites de controle identificados encontram-se em um período próximo nas três análises realizadas, concentrando os dados de investigação para a auditoria financeira. A contribuição acadêmica deu-se pelo desenvolvimento de um modelo de aplicação de Cartas de Controle multivariadas e da Lei de Benford, com uma abordagem inovadora do controle estatístico de processos voltado à área financeira, utilizando recurso computacional acessível, de fácil processamento, confiável e preciso, que permite aprimoramento por novas abordagens acadêmicas. No que tange à contribuição à sociedade, se dá pelo uso do modelo por entidades que atuam com movimentações financeiras e pela comunidade, em dados de organizações civis e estatais divulgados nos canais de informação, de modo a proporcionar a prática cidadã pelo acesso à análise e a constatação da idoneidade dos fatos e dos dados.

**Palavras-chave:** Suspeitas de fraude financeira. Lavagem de dinheiro. Carta de Controle Multivariada. Lei de Benford. Controle Estatístico de Processos.

## ABSTRACT

Large losses are generated in the countries financial systems, by money laundering. The volume of financial data is big issue to identify digital crime and money laundering. Audits in financial data have limitations in detecting fraud, in large part it is still performed in a traditional way, data are collected by sampling and often unable to identify a real-time crime situation. This research is aiming to serve in addressing this gap, to propose an monitoring statistical method, from multivariate control chart based on Benford's law for detecting suspicious of fraud in financial data, including those due to money laundering. It was initially defined as a conceptual model in order to determine the type of probability distribution that represents data from financial launches. It was adopted an assumption that this type of data adheres to the Benford's Law distribution. Subsequently, an empirical distribution was obtained, estimated from the own data. Two procedures were tested to verify a suspected money laundering fraud through the significant first-digit assessment: The Multivariate  $\chi^2$  Control Chart and the Multivariate Hotelling's  $T^2$  Control Chart. Data were simulated using the R-Project software until the occurrence of the 50.000<sup>o</sup> signal. Finally, the simulation procedures were applied to real data in order to exemplify the method operationally. From the simulation, the two Control Charts tested were evaluated for *ARL*, that is, average number of observations until the signaling that the series started to operate in an out-of-control state, which it means suspicious of fraudulent launches. The application of the retrospective analysis method in the financial launchings of county's campaign from 2016 Elections in five capitals of Brazil, based on the expected proportions from the first digit given by Benford's Law, no suspicions fraud were evidenced in the data obtained from the site of Tribunal Superior Eleitoral (TSE). Considering the application in a set of data from a financial institution, signs of divergence between the frequencies of the first digits of the entries and the expected values were observed, but these points beyond the identified limits are close in all three analyzes. Indicating the period of the data which ones the audit will focus in a further investigation. Academic contribution is identified by developing a multivariate Control Chart together the Benford's law in an application model with an innovative approach to the statistical process control aimed at the financial area, using accessible, easy to process, reliable and accurate computational resources that allow improvement through new academic approaches. As regard to the contribution to society, it is given the opportunity of applying the model by financial entities and the community in the data of civil and state organizations, disclosed in the information channels in order to provide access to analysis and verification of the suitability of facts and data by citizen practice.

**Keywords:** Suspected financial fraud. Money laundering. Multivariate Control Chart. Benford's Law. Statistical Process Control.



## LISTA DE FIGURAS

Figura 1:	Área sob a curva da distribuição normal . . . . .	44
Figura 2:	Carta de Controle de Shewhart . . . . .	45
Figura 3:	Fluxograma de procedimentos da Simulação . . . . .	59
Figura 4:	Cartas de Controle multivariadas de Benford para avaliação de suspeita de fraude - Prefeitura da Capital A - Eleições 2016 . . . . .	72
Figura 5:	Cartas de Controle multivariadas de Benford para avaliação de suspeita de fraude - Prefeitura da Capital B - Eleições 2016 . . . . .	73
Figura 6:	Cartas de Controle multivariadas de Benford para avaliação de suspeita de fraude - Prefeitura da Capital C - Eleições 2016 . . . . .	74
Figura 7:	Cartas de Controle multivariadas de Benford para avaliação de suspeita de fraude - Prefeitura da Capital D - Eleições 2016 . . . . .	75
Figura 8:	Cartas de Controle multivariadas de Benford para avaliação de suspeita de fraude - Prefeitura da Capital E - Eleições 2016 . . . . .	76
Figura 9:	Cartas de Controle multivariadas de Benford para avaliação de suspeita de fraude - dados da Instituição Financeira . . . . .	79

## LISTA DE TABELAS

Tabela 1:	Artigos encontrados nas bases de dados CAPES . . . . .	17
Tabela 2:	Probabilidades dos primeiros dígitos de Benford . . . . .	37
Tabela 3:	Probabilidades da distribuição dos dois primeiros dígitos . . . . .	38
Tabela 4:	Avaliação da aplicação da Carta de Controle $\chi^2$ - Distribuição dos primeiros dígitos (dBenford) . . . . .	69
Tabela 5:	Avaliação da aplicação da Carta de Controle $T^2$ - Distribuição das médias da estatística (eBenford) . . . . .	69
Tabela 6:	ARLs e LSCs estimados . . . . .	70

## LISTA DE QUADROS

Quadro 1: Procedimento de execução do monitoramento de movimentações financeiras . . . . .	67
--	----

## SUMÁRIO

<b>1 INTRODUÇÃO</b> . . . . .	<b>13</b>
<b>1.1 Apresentação do Tema e Problema de Pesquisa</b> . . . . .	<b>13</b>
<b>1.2 Objetivos</b> . . . . .	<b>15</b>
<b>1.3 Justificativa</b> . . . . .	<b>16</b>
<b>1.4 Delimitações</b> . . . . .	<b>20</b>
<b>2 REFERÊNCIAL TEÓRICO</b> . . . . .	<b>21</b>
<b>2.1 Auditoria</b> . . . . .	<b>21</b>
2.1.1 Origem da Auditoria . . . . .	21
2.1.2 Da Auditoria Contábil à Auditoria de Sistemas de Informação . . . . .	24
<b>2.2 Lavagem de Dinheiro</b> . . . . .	<b>26</b>
2.2.1 Ação Internacional de Combate a Lavagem de Dinheiro . . . . .	26
2.2.2 Combate à Lavagem de Dinheiro no Brasil . . . . .	29
2.2.3 A Prevenção da Lavagem de Dinheiro . . . . .	31
<b>2.3 Auditoria em Lavagem de Dinheiro</b> . . . . .	<b>33</b>
<b>2.4 Lei de Benford</b> . . . . .	<b>36</b>
2.4.1 Conceito e Definição . . . . .	36
2.4.2 Aplicações da Lei de Benford . . . . .	38
<b>2.5 Cartas de Controle de Processos</b> . . . . .	<b>43</b>
2.5.1 Carta de Controle de monitoramento multivariada $\chi^2$ . . . . .	49
2.5.2 Carta de Controle $T^2$ de Hotelling . . . . .	50
<b>2.6 Síntese do referencial e abordagem proposta</b> . . . . .	<b>53</b>
<b>3 PROCEDIMENTOS METODOLÓGICOS</b> . . . . .	<b>56</b>
<b>4 APRESENTAÇÃO, VALIDAÇÃO E APLICAÇÃO DO MODELO DE MONITORAMENTO PROPOSTO</b> . . . . .	<b>65</b>
<b>4.1 Estrutura do modelo de monitoramento proposto para avaliação de suspeitas de fraude por lavagem de dinheiro</b> . . . . .	<b>65</b>
<b>4.2 Apresentação dos resultados</b> . . . . .	<b>66</b>
4.2.1 Resultados da validação do modelo por simulação . . . . .	66
4.2.2 Exemplos de aplicação do modelo em dados reais . . . . .	72
<b>5 CONSIDERAÇÕES FINAIS</b> . . . . .	<b>80</b>
<b>REFERÊNCIAS</b> . . . . .	<b>84</b>

<b>APÊNDICE A</b>	<b>CÓDIGOS NO <i>SOFTWARE R</i></b>	<b>95</b>
<b>A.1</b>	<b>Código geral: Arquivo main.R</b>	<b>95</b>
<b>A.2</b>	<b>Carta <math>\chi^2</math>: Arquivo quiquadrado.R</b>	<b>96</b>
<b>A.3</b>	<b>Carta <math>T^2</math> de Hotteling: Arquivo T2.R</b>	<b>98</b>

## 1 INTRODUÇÃO

### 1.1 Apresentação do Tema e Problema de Pesquisa

Fraude por lavagem de dinheiro é um ilícito previsto na legislação brasileira. A Lei nº 12.683, de 9 de Julho de 2012, deu nova redação à Lei nº 9.613, de 3 de Março de 1998 no trato penal desse tipo de ilícito. O Conselho de Controle de Atividades Financeiras (COAF) é o órgão criado no âmbito do Ministério da Fazenda para atuar em conformidade com os preceitos legais de prevenção e combate à lavagem de dinheiro e ao financiamento do terrorismo. Segundo o COAF, lavagem de dinheiro é um processo dinâmico que se caracteriza pelo afastamento de forma disfarçada de fundos da sua origem. A movimentação ocorre em três etapas: colocação, ocultação e integração. Na colocação, o dinheiro é colocado no sistema econômico geralmente de forma fracionada e em locais onde os controles legais são mais liberais. Na ocultação, o dinheiro é movimentado anonimamente para evitar o rastreamento contábil em paraísos fiscais ou através de contas fantasma. Por fim, na integração, o dinheiro é integrado ao sistema econômico legal de forma que evite suspeitas através de investimentos que facilitem os negócios das organizações criminosas. Uma vez que a rede esteja formada, as aparências encobrem as ações, tornando mais difícil identificar a origem do dinheiro ilegal. (COAF, 2015a).

Publicações relacionadas com o assunto abordam o tratamento de ocorrências de ilícitos em desvios, fraudes e em lavagem de dinheiro. No âmbito governamental, em se tratando de desvios e superfaturamento em órgãos públicos, há alta incidência de obras paradas, atrasadas ou inadequadamente licitadas no Brasil. Pacheco Filho (2004) afirma que o simples cumprimento do inciso I do artigo 7º, parágrafo 2º, da Lei de Licitações nº 8.666 de 21 de Junho de 1993, eliminaria grande parte das irregularidades constatadas todos os anos pelo Tribunal de Contas da União (TCU). O inciso I menciona que deve haver projeto básico aprovado por autoridade competente, critério técnico para projeto de execução de qualquer obra, descumprido sistematicamente por autoridades de todas esferas de governo. Pacheco Filho (2004) menciona ainda a Lei Complementar 101, promulgada em 04 de Maio de 2000, que trata da Responsabilidade Fiscal como o mecanismo legal e moralizador de gastos públicos. O autor refere a Lei como amparo do legislador e da sociedade para fazer cumprir a liberação de recursos do orçamento da

União somente para projetos concluídos e aprovados pelos órgãos reguladores, ministérios, entes controladores ou entidades do gênero.

A respeito de lavagem de dinheiro, Schneider (2010) afirma que além da economia oficial existe uma economia paralela, que inclui todos os tipos de atividades criminosas em conflito com a lei. Como exemplos, o autor cita o tráfico de seres humanos e o tráfico de drogas. À frente de ações criminosas, atividades à sombra da economia se manifestam, principalmente, em produtos e serviços como sobrepreço na economia oficial. Neste caso, a ilegalidade vem da evasão de impostos e da violação na regulamentação do mercado de trabalho. Schneider (2010) acrescenta que há estimativas de que, por meio de ações criminosas, em nível mundial, sejam lavados anualmente entre US\$ 2,0 a 2,5 trilhões, valor que representa em torno de 5 a 6% do produto bruto global no ano de 2006. Estimativas mostram que no setor financeiro mundial (fundos de investimento e bancos), foram lavados cerca de US\$ 0,5 a 1,0 trilhão somente no ano de 2004.

Os avanços tecnológicos propiciam maior sofisticação na prática de crimes de lavagem de dinheiro. Yuen (2014) afirma que as instituições financeiras devem intensificar e aprimorar o uso de dados não estruturados para melhor entender seus clientes e dar efetividade a suas iniciativas de controle. Soluções anti-lavagem de dinheiro devem ter capacidade de detectar atividades financeiras suspeitas por meio da identificação de indivíduos ou entidades que possam estar envolvidos diretamente nestas práticas. O conhecimento do perfil do cliente deve ser foco das instituições financeiras, elas são responsáveis por entender os seus clientes em diferentes áreas, incluindo fatos e circunstâncias, tipo de negócio, fontes de capital, os objetivos de transações específicas e natureza esperada das operações. A forma mais prática de cumprir os regulamentos anti-lavagem de dinheiro é identificar as transações comerciais normais dos clientes para que as atividades incomuns ou suspeitas sejam reconhecidas. Yuen (2014) cita o Canadá como exemplo de local do mundo onde houve a adoção dessa prática.

Reforçando o que cita Yuen (2014), Gao e Xu (2009) mencionam que, ações em nível mundial têm sido implementadas contra a lavagem de dinheiro. Na tentativa de coibir esse crime, cada vez mais sistemas de prevenção de lavagem de dinheiro têm sido implementados. O volume de transações nos bancos é grande e aumenta a cada dia, em razão disso os sistemas existentes não têm a capacidade de verificar todas as transações de uma forma abrangente e consistente. A presença de alto volume de dados em extensos relatórios não facilitam a detecção da lavagem de dinheiro em tempo real. As

soluções normativas tradicionais apresentam muitas desvantagens pelo elevado número de falsos positivos. (GAO; XU, 2009; YUEN, 2014). Embora estejam baseados em regras, e apresentarem alguma capacidade de reconhecimento de padrões, os sistemas não têm capacidade de aprendizagem com o processo. As funções de reconhecimento de padrões e a capacidade de processamento são limitadas. (GAO; XU, 2009).

No que tange à suspeição de crimes por lavagem de dinheiro, cabe salientar que toda sociedade é afetada por esse tipo de delito. Como justificativa de aplicação dos resultados da pesquisa, há como embasamento uma série de fatos relacionados com fraudes, lavagem de dinheiro e crimes de tráfico de drogas. Dentre os impactos e as consequências da não detecção de eventos que proporcionam a lavagem de dinheiro, há que se considerar o que concluem Amorim, Cardozo e Vicente (2012), o dinheiro ilícito causa diversos problemas de cunho social e econômico como a desvalorização da moeda nacional, pressão sobre o câmbio e a impunidade dos criminosos poderosos, o que gera descrédito na Justiça. Yuen (2014) complementa que existem custos sociais associados à lavagem de dinheiro, o impacto mais significativo é o aumento da corrupção e crimes na sociedade. Lavagem de dinheiro bem sucedida torna as atividades criminosas lucrativas permitindo que, contrabandistas, traficantes de drogas e outros criminosos continuem expandindo suas operações e consequentemente atraíam mais criminosos.

Neste contexto, devido aos danos causados pelos crimes de lavagem de dinheiro, uma pesquisa que envolve aspectos de monitoramento de padrões que permitem o aprendizado, ao mesmo tempo que são capazes de detectar suspeitas de atividades que alteram estes padrões, são relevantes para a sociedade. Procedimentos de monitoramento de processos na indústria que visam a melhoria contínua da qualidade, a exemplo do Controle Estatístico de Processos na aplicação de Cartas de Controle, apresentam características similares às necessárias no problema de identificação de suspeitas de fraude em lançamentos financeiros. Neste sentido, a questão de pesquisa pode ser assim enunciada: “Qual a efetividade do uso de Cartas de Controle multivariadas na detecção da presença de suspeitas de fraude financeira?”

## **1.2 Objetivos**

O objetivo geral desta pesquisa é avaliar a efetividade do uso de Cartas de Controle multivariadas na detecção de presença de suspeitas de fraude financeira. Para tanto, os



seguintes objetivos específicos devem ser alcançados:

- Verificar a efetividade de Cartas de Controle multivariadas na detecção da presença de movimentações financeiras com suspeição de fraude, com base na Lei de Benford;
- Propor uma Carta de Controle multivariada para detecção de suspeitas de fraude financeira; e
- Analisar a performance dos procedimentos citados nos objetivos específicos anteriores, no monitoramento, a partir da simulação de dados com a inclusão de movimentações fraudulentas.

### 1.3 Justificativa

A relevância do tema se dá por inserir uma proposta de inovação na aplicação de recursos estatísticos. Segundo Woodall e Montgomery (2014), recentemente várias atividades de investigação em monitoramento estatístico de processos têm sido realizadas. Em decorrência disso, a gama de aplicações expandiu-se para diversas áreas do conhecimento, tais como os trabalhos de Kulldorff (1997); Hjalmarsson et al. (1996); Duczmal, Kulldorff e Huang (2012) aplicando a estatística na área médica pela análise espaço-temporal de epidemias, Minamisava et al. (2009) em estudo espaço-temporal de ocorrências de mortes violentas em áreas recém urbanizadas no Brasil, Nigrini (2011, 2012) em seus estudos de aplicação da Lei de Benford em prevenção de fraudes na contabilidade forense, Silva, Korzenowski e Vaccaro (2014) na identificação de padrões estatísticos em suspeitas de fraude por lavagem de dinheiro.

Os processos de auditoria orientam-se, comumente, em procedimentos padronizados de verificações de documentação das movimentações contábeis e financeiras, com dados coletados por amostragem e muitas vezes incapazes de identificar, em tempo real, uma situação de delito. Segundo Coelho (2015), presidente do Instituto dos Auditores Independentes do Brasil (IBRACON), ao contrário do que o senso comum imagina, o auditor independente não tem a missão de buscar fraudes. Ele deve estar atento ao risco de sua ocorrência para fazer seu julgamento, testa os controles da empresa por amostragem, analisa sua estrutura de governança e solicita o diagnóstico de especialistas para apurar operações fora de sua área de conhecimento, dentre outros procedimentos definidos pelas normas que regem a profissão no Brasil e no exterior. Em sua definição

sobre fraude, Coelho (2015) declara que, se a fraude é um problema universal, então as medidas para coibi-la também devem ser universais. O desafio é construir uma cultura interna nas organizações que iniba a fraude.

Embora tenha havido evolução nas medidas preventivas para combate à lavagem de dinheiro, Yuen (2014) afirma que, indivíduos e empresas que praticam a lavagem de fundos tornaram-se especialistas em atacar os pontos fracos dos sistemas anti-lavagem de dinheiro.

O levantamento bibliográfico mostra a relevância do tema e as lacunas encontradas nas publicações que se relacionam com a lavagem de dinheiro. A Tabela 1 apresenta o volume de artigos e os totais encontrados no período de 2012 a 2016 por palavras-chave e suas combinações nas bases de dados pesquisadas.

**Tabela 1:** Artigos encontrados nas bases de dados CAPES

Palavras Chave	Artigos Encontrados / Ano					Total
	2012	2013	2014	2015	2016	
<i>Money Laundering + Financial Fraud</i>	46	29	18	22	1	116
<i>Money Laundering + Financial Crime</i>	66	48	44	37	4	199
<i>Money Laundering + Internal Controls</i>	11	9	7	9	24	60
<i>Money Laundering + Internal Audit</i>	3	82	90	85	15	275
<i>Money Laundering + Systems Information</i>	16	5	11	10	2	44
<i>Money Laundering + Systems Information Auditing</i>	1	33	46	39	8	127
<i>Money Laundering + Computer</i>	27	22	20	19	4	92
<i>Money Laundering + Computer Fraud</i>	8	5	4	5	16	38
<i>Money Laundering + Computer Crime</i>	8	7	6	4	13	38
<i>Money Laundering + Cybercrime</i>	8	8	5	6	3	30
<i>Anti-Money Laundering + Internal Audit</i>	38	32	32	37	4	143
<i>Anti-Money Laundering + Audit Systems</i>	35	22	29	31	5	122
<i>Fraud Prevention + Internal Controls</i>	12	3	2	14	26	57
<i>Fraud Prevention + Systems Information Auditing</i>	102	98	90	78	11	379
<i>Systems Information + Auditing</i>	168	170	189	123	23	673
<i>Systems Information Auditing + Internal controls</i>	8	5	11	8	56	88
<i>Systems Information Auditing + Computer Fraud</i>	111	99	73	64	8	355
<i>Benford's Law + Money Laundering</i>	1	0	0	2	1	4
<i>Benford's Law + Fraud</i>	8	14	14	15	3	54
<i>Benford's Law + Internal Controls</i>	12	9	8	10	3	42
<i>Benford's Law + Statistical Process Control</i>	21	30	37	26	2	116
<i>Benford's Law + Control Chart</i>	3	7	3	3	2	18
Total Geral	714	738	739	647	232	3.070

Fonte: Bases de dados das publicações.

No período dos cinco anos pesquisados, houve redução no Total Geral de artigos publicados na área de interesse do estudo. Apesar disso, encontrou-se número signi-

ficativo de artigos tratando das seguintes associações das palavras-chave, como segue: A palavra-chave *Money Laundering*, associada a *Financial Fraud*, *Financial Crime*, *Internal Audit*, *Systems Information Auditing* e *Computer*. A palavra-chave *Anti-Money Laundering*, associada a *Internal Audit* e *Audit Systems* apresenta número de publicações equilibrado entre 2012 e 2015. A palavra-chave *Fraud Prevention*, associada a *System Information Auditing*. As palavras-chave *Systems Information* e *Systems Information Auditing*, associadas a *Auditing* e *Computer Fraud*. Por fim, a palavra-chave *Benford's Law*, associada a *Statistical Process Control* e com *Control Chart*, apresenta-se relacionada ao tema proposto na aplicação da Lei de Benford, associada a Cartas de Controle.

Em se tratando da aplicação da Lei de Benford, relacionada com o monitoramento de processos, observa-se na Tabela 1 artigos relacionando *Benford's Law* com *Statistical Process Control* e com *Control Chart*.

A lei de Benford associada a recursos estatísticos tem sido usada para descobrir a manipulação e a fabricação de dados em diversas áreas.

Dentre os artigos encontrados nas bases de dados, há 18 publicações associando *Control Chart* a *Benford's Law* para a identificação de fraudes. O número de publicações enseja oportunidade de desenvolvimento de novas abordagens, o que corrobora com a relevância da presente pesquisa, que trata especificamente da aplicação da Carta de Controle multivariada na detecção de presença de suspeitas de fraude financeira. Dos 18 artigos envolvendo o uso de Cartas de Controle em conjunto com a Lei de Benford, comenta-se alguns exemplos das publicações.

Um estudo sobre fraude financeira foi realizado em empresas que manipulavam dados de teste de produtos, com o fim de obterem lucros ilícitos. Chung et al. (2016), identificaram a manipulação de testes realizados em cabos de energia para usinas nucleares e nos testes de suprimentos militares. O objetivo do estudo foi o de implementar um método para validar os sistemas de teste dos fornecedores, fazendo uso de Cartas de Controle e a Lei de Benford na detecção de dados falsos, utilizando o *software* popular MS Excel.

Um inquérito documenta prováveis declarações errôneas de dados oficiais da qualidade do ar em Pequim, no período de 2008 a 2013. Comparativos feitos com dados coletados pela embaixada dos USA em Pequim dão conta do erro. Stoerk (2016) demonstra em seu estudo testes estatísticos utilizando a Lei de Benford e a Carta de con-

trole  $\chi^2$ , demonstrando que os dados não aderem à Lei de Benford. Isso mostrou que as autoridades manipulavam os dados divulgados à comunidade, utilizando o que foi chamada de **corrupção estatística**.

A crescente descoberta de substâncias farmacêuticas destaca a falta de qualidade nos dados de desenvolvimento de produtos por parte dos laboratórios. Estudos prévios de Orita et al. (2013), identificaram a aderência à distribuição de Benford em dados de solubilidade e atividade de reagentes em componentes fármacos. Um novo estudo realizado em uma gama mais ampla de dados de sequenciamento em micro-partículas, aplicando Carta de Controle e a Lei de Benford, conduziu Orita et al. (2013) à constatação da má conduta científica ao identificarem fatores de erro pela manipulação de dados. Essas constatações conduziram ao desenvolvimento de um método de avaliação sistemática de monitoramento, utilizando um protocolo baseado na lei de Benford, fornecendo aos pesquisadores um método eficiente para a avaliação da confiabilidade da qualidade dos dados.

Como visto nas bases pesquisadas, a justificativa acadêmica se norteia pelo fato de que o estado da arte atual dos processos para a identificação de fraudes por lavagem de dinheiro se dá, em sua maioria, por modelos de amostragem. A aplicação do recurso de monitoramento por Carta de Controle, com base na Lei de Benford em específico, por Cartas de Controle multivariadas, apresenta poucas publicações. O modelo poderá ser implantado em organizações financeiras, órgãos de Estado e outras, incluindo as organizações privadas que tratam de movimentações financeiras, com o objetivo de assegurar a confiabilidade e transparência em seus processos e na divulgação de resultados à sociedade. Portanto, a contribuição da dissertação vem preencher uma lacuna na pesquisa em aplicação de ferramentas estatísticas no monitoramento de dados de lançamentos financeiros para avaliação de suspeitas de fraude por lavagem de dinheiro.

Enfatizando a relevância do tema, a Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) vem apoiando pesquisadores na criação de modelos estatísticos para detecção e prevenção de fraudes em operações financeiras. Alisson (2015) refere o levantamento da Serasa Experian que, a cada 15,4 segundos há registro de tentativa de fraude por identidade no Brasil, dados pessoais são usados para fazer negócios pela internet sob falsidade ideológica ou obter crédito com a intenção de não honrar o pagamento. O modelo estatístico foi criado para prever a ocorrência de fraudes em compras com cartões de crédito. (ALISSON, 2015).

Em resumo, quanto à justificativa acadêmica, percebe-se que no estado-da-arte há espaço para o desenvolvimento de meios de monitoramento para avaliar a suspeição de crime de fraude por lavagem de dinheiro. Por outro lado, no que tange aos aspectos aplicados da pesquisa, a disponibilidade de mecanismos que inibam os crimes de lavagem de dinheiro contribuirá para uma melhor vida em sociedade. Assim, de acordo com os argumentos apresentados, justifica-se o presente trabalho, tanto pelo aspecto acadêmico quanto pelo aspecto de aplicabilidade prática.

#### **1.4 Delimitações**

A pesquisa utilizará dados de movimentações financeiras envolvendo entidades que atuam no mercado legal. Na etapa de estudo de campo, buscou-se dados reais em agentes financeiros para o teste da eficácia dos meios de monitoramento, sendo que os dados obtidos foram utilizados para aplicar modelos de simulação no intuito de validar os procedimentos analisados nesta pesquisa. Dados das movimentações financeiras das campanhas para prefeito em 2016 também foram coletados nos sítios de internet dos Tribunais Eleitorais e os métodos foram aplicados para demonstrar a aplicabilidade das técnicas de avaliação em dados retrospectivos.

A dissertação se restringiu a desenvolver o método e validá-lo. Do mesmo modo, o algoritmo implementado é apresentado para que bancos e fundos de investimento também possam implementá-lo no monitoramento das movimentações financeiras de seus clientes, com o objetivo de detecção de movimentações fora do padrão histórico dos clientes, sem a necessidade da utilização de avaliações arbitrárias. Por fim, apenas um modelo de Carta de Controle multivariada foi avaliada em complemento a este trabalho, apresentado com base nos resultados encontrados na pesquisa, aliados a sugestões de outros procedimentos com o propósito de avaliar suspeição de fraudes em lavagem de dinheiro.

Com relação à tradução de *Control Chart*, encontra-se na literatura tanto Gráfico de Controle como Carta de Controle e para simplificar a leitura deste trabalho, a partir deste ponto terá a denominação de Carta de Controle.

## 2 REFERÊNCIAL TEÓRICO

### 2.1 Auditoria

#### 2.1.1 Origem da Auditoria

A Auditoria desenvolveu-se pela Ciência Matemática utilizada nos Sistemas de Escrituração Comercial. Iudícibus, Martins e Carvalho (2005) apontam seu uso a partir do final do século XIII e início do século XIV, sua evolução aos princípios da Contabilidade conhecida atualmente se deu naturalmente de Sistemas de Escrituração para Sistemas de Informação. Publicado em 1202, o livro *Liber Abaci*, de Leonardo Fibonacci, foi utilizado nas aritméticas comerciais do período medieval e renascentista. (PIRENNE, 1968). Em 1494, o Frei Luca Pacioli publicou o tratado de matemática *Summa de Arithmetica, Geometria, Proportioni et Proportionalita*, dedicado a exposição do sistema de escrituração por partidas dobradas (débito / crédito). (CHESNEAUX, 1995; IUDÍCIBUS; MARTINS; CARVALHO, 2005).

A evolução histórica da contabilidade como ciência, segundo Martins, Martins e Martins (2007) ocorreu em três fases:

**Contabilidade Proprietária:** usada para escriturar dados de propriedade particular dos homens de negócio;

**Contabilidade Financeira:** em razão da revolução industrial o financiamento de negócios necessitava avaliar o balanço patrimonial e apresentar demonstrações de resultados para o direcionamento dos investimentos; e

**Contabilidade Gerencial:** surgiu para a gestão interna dos meios de produção, com o objetivo de gerar informações para tomada de decisão futura na gestão de negócios.

Segundo Gomes, Araújo e Barboza (2009) a auditoria nasceu na Inglaterra. No século XIII o país dominava os mares, controlava o comércio mundial e foi o primeiro país a criar a taxaço do imposto de renda baseada no lucro das empresas. Cita Veyne (1990), que por toda a Idade Média e após o Renascimento surgiram associações profissionais nas diversas regiões da Europa, com a missão de executar as funções de auditoria, os Conselhos Londrinos em 1310; o *Collegio dei Raxonati* em 1581, na cidade

de Veneza; o Tribunal de Contas de Paris em 1640; e a *Accademia dei Ragionieri* em 1658, nas cidades de Milão e Bolonha. A auditoria consolidou-se na Inglaterra, a partir da Sociedade dos Contadores de Edimburgo. (CARDOZO, 2014). Fundada em 1853, deu origem ao Instituto dos Contadores Associados da Inglaterra e do País de Gales - *Institute of Chartered Accountants in England and Wales* (ICAEW), entidade de maior prestígio nas áreas da contabilidade e auditoria da Grã Bretanha. Depreende-se que a moderna auditoria nasceu na Inglaterra e foi introduzida nos Estados Unidos por Contadores Ingleses, enviados para auditarem firmas norte-americanas pertencentes à capitais britânicos. Cardozo (2014) salienta que, por influência inglesa, desde 1887 formou-se a associação profissional de Contadores dos Estados Unidos, que desde 1916 foi denominada Instituto Americano de Contadores Públicos Certificados - *American Institute of Certified Public Accountants* (AICPA). O AICPA atua na formação, avaliação do perfil técnico e impõe rigorosos padrões de avaliação e éticos dos profissionais de contabilidade e de auditoria. Seu renome mundial repercutiu na difusão de seus procedimentos técnicos nos países do mundo capitalista. (GOMES; ARAÚJO; BARBOZA, 2009; CARDOZO, 2014).

A exemplo dos Estados Unidos em relação aos investidores ingleses, a moderna auditoria chegou ao Brasil na década de 1940 com a entrada de companhias multinacionais. (VEYNE, 1990). Na década de 1960 os auditores brasileiros organizaram-se em associação de classe, e em 1971 foi legalmente reconhecida como Instituto dos Auditores Independentes do Brasil (IBRACON), pela Resolução nº 317, do Conselho Federal de Contabilidade e da Resolução nº 220, do Banco Central do Brasil, ambas de 1972. (GOMES; ARAÚJO; BARBOZA, 2009).

Crepaldi (2004) conceitua auditoria como um exame minucioso e sistemático das atividades desenvolvidas pela organização para averiguar se estão em conformidade com padrões definidos para atingir os objetivos organizacionais e atender a legislação. Com o objetivo principal de detecção de fraudes nos controles contábeis e de patrimônio, Díaz e Vera (2006) citam que as atividades de auditoria ampliaram-se ao final do século XIX pelos movimentos provocados pela revolução industrial e desenvolveu-se em dois tipos: i) Auditoria externa e ii) Auditoria interna. Cardozo (2014) define os dois tipos, a auditoria externa é realizada por auditores independentes da empresa que os contrata. O objetivo principal é o de emitir opinião sobre as demonstrações contábeis e fornecer subsídios para divulgação das demonstrações públicas dos resultados e de de-

sempenho do negócio, conforme a periodicidade legal. A auditoria interna é realizada por representantes diretos da administração da empresa e tem como objetivo principal controlar os movimentos internos da organização, com a finalidade de assessorar seus dirigentes na definição imediata de medidas e ações que influenciam nos resultados do negócio.

Segundo Porter, Burton e Ramírez (1980), o conceito de auditoria interna se firmou no período de 1940 a 1950, dada a importância que os controles internos assumiram nas grandes organizações, conceito reforçado pelos critérios da Lei *Sarbanes-Oxley* (SOX), chamada de ato SOX, lei criada por Paul Sarbanes e Michael Oxley, aprovada pelo congresso dos Estados Unidos, em 2002, com a finalidade de proteger os investidores, assegurar a precisão, confiabilidade, ética no cumprimento de regras de governança corporativa e nas publicações legais em conformidade com a legislação de valores mobiliários e outros fins. A Comissão de Valores Mobiliários - *Securities and Exchange Commission* (SEC) avaliza e administra o ato SOX, estabelecendo os padrões de conformidade, os requisitos e os prazos das publicações legais.

Dada a sua evolução, na década de 1950 a auditoria expandiu-se para outras áreas das organizações além da área contábil e financeira, tais como: Sistemas de Gestão da Qualidade, Processos de Negócio, Gestão do meio ambiente e Gestão da Segurança e Saúde. Segundo Díaz e Vera (2006), essas novas modalidades abriram caminho para a auditoria de ativos intangíveis em Sistemas de Informação e Gestão da Segurança da Informação.

A partir da década de 1960 a evolução dos equipamentos de informática através da redução do tamanho dos componentes eletrônicos, principalmente os circuitos integrados e processadores, permitiram grande avanço no desenvolvimento computacional. (CURY; CAPOBIANCO, 2014). Esse avanço possibilitou o processamento de grande massa de dados, o que, aliado à comunicação digital em tempo real permitiu fácil trânsito de informações pela rede mundial. O desenvolvimento global das comunicações e o alto volume de informação circulante favoreceram as transações entre as organizações internacionalmente. Carneiro (2009) cita a clara mudança na forma de operação das empresas que passaram de sociedades de indústria para sociedades de informação.



### 2.1.2 Da Auditoria Contábil à Auditoria de Sistemas de Informação

Oliveira (2006) evidencia que os Sistemas de Informação (SI) existem desde a origem das relações humanas, independentemente da tecnologia. A tecnologia influenciou a forma como os dados são tratados. Segundo Cury e Capobianco (2014), os avanços tecnológicos inseriram velocidade no fluxo das informações devido à evolução nas técnicas de processamento. Oliveira (2006) destaca a importante raiz dos sistemas de informação no tratado matemático, publicado pelo Frei Luca Pacioli, em 1494, o autor afirma que, naturalmente, qualquer sistema informático está contido em um Sistema de Informação, por consequência, o surgimento da Tecnologia da Informação (TI) tornou os processos mais rápidos proporcionando contínuo crescimento na capacidade de armazenagem de grande volume de dados.

Os novos modelos de negócio são constituídos por milhões de transações diárias que necessitam captura e registro por processos com um nível de automação compatível, com intervenção humana praticamente nula. Registros contábeis ocorrem automaticamente no momento imediato das transações, sejam comerciais ou financeiras, em plataformas informáticas implantadas previamente nas organizações. A par dessa realidade, Martins (2013) enfatiza ser primordial a confiabilidade nos controles internos e na capacidade dos sistemas. A evolução das comunicações impôs a necessidade da gestão segura do fluxo de informações.

Como consequência do desenvolvimento da Tecnologia da Informação e nas Tecnologias da Informação e Comunicação (TIC) ampliadas pelo uso da internet, Cury e Capobianco (2014) assinalam que a facilidade no trânsito de dados viabilizou um crescimento sem precedentes das relações econômicas nos negócios em caráter global, no comércio através do *e-commerce* e nas transações entre as entidades financeiras. Estes avanços trouxeram ganhos, facilidades e vulnerabilidades ao mesmo tempo, em razão destas vulnerabilidades surgiu a necessidade de controles eficazes culminando na origem da Auditoria de Sistemas da Informação nas décadas de 1970 e 1980. Seus princípios básicos são a disponibilidade, integridade, confidencialidade, não repúdio e a autenticidade dos dados em estrita relação aos participantes, emissor e receptor da comunicação. DÍAZ; VERA (2006).

As instituições financeiras são as entidades que mais dependem de sua plataforma informática, as informações em seus sistemas são vitais à natureza do negócio. Neste

contexto, a auditoria ganha seu maior impulso na década de 1990, mais abrangente e sistemática. Porter, Burton e Ramírez (1980) destacam a ênfase nos controles internos e a auditoria contínua com caráter independente, associados à governança corporativa para gestão dos riscos organizacionais.

As transações realizadas pela internet exigem plataformas de alta performance em segurança da informação. Práticas e controles que oferecem robustez contra ataques e invasões foram desenvolvidos e aprimorados ao longo do tempo com a experiência adquirida, a segurança da informação e gestão de tecnologia da informação exigem qualificação e conhecimento sofisticados. Normas, padrões, práticas de prevenção a invasões e ataques aos sistemas foram desenvolvidos, dentre os quais, a norma ISO 27001 que trata dos requisitos para certificação de Sistemas de Gestão da Segurança da Informação.

A *Information Technology Infrastructure Library* (ITIL), do Reino Unido, é uma biblioteca que compila e divulga sistematicamente um conjunto de publicações das melhores práticas de governança e de Gestão da Tecnologia da Informação utilizadas mundialmente. Em termos de controles internos, o *Committee of Sponsoring Organizations of the Treadway Commission* (COSO), referência nos Estados Unidos (USA), é um modelo de processo para prover confiabilidade e conformidade com os regulamentos legais na gestão dos Sistemas de Informação. (GERTZ; GULDENTOPS; STROUS, 2013).

A Auditoria de Sistemas de Informação não é apenas uma extensão da auditoria tradicional. Segundo Oliveira (2006) devido à importância das plataformas informáticas nas organizações, há necessidade de controle e verificação constantes. A natureza desses sistemas é de serem invisíveis, acessíveis e modificáveis sem deixar vestígios. A pressão pela melhoria dos mecanismos de controle e de transparência demanda a modernização das técnicas de auditoria.

A Auditoria Contínua (AC), voltada para a análise de dados em formato eletrônico é um dos avanços. O recurso vem sendo adotado pela ampliação das transações sem papel e por imposições legais. Motta-Jr (2010) cita como exemplos o ato SOX dos USA, que trata da transparência na governança corporativa e a legislação brasileira em seu artigo 5º da Constituição Federal, onde determina que as informações relativas à execução orçamentária e financeira dos entes públicos sejam publicadas em tempo real em meio eletrônico.

## 2.2 Lavagem de Dinheiro

### 2.2.1 Ação Internacional de Combate a Lavagem de Dinheiro

Quando uma atividade criminosa gera lucros substanciais, o indivíduo ou o grupo envolvido procura encontrar uma maneira de controlar os fundos sem atrair a atenção para a atividade subjacente ou para as pessoas envolvidas. Ao fazerem isso, os criminosos disfarçam as fontes mudando a forma. Eles movem os fundos para um lugar onde eles são menos propensos a atrair a atenção. Venda ilegal de armas, contrabando e as atividades do crime organizado, incluindo as redes do tráfico de drogas e prostituição podem gerar enorme volume de recursos. Esquemas de peculato, abuso de informação privilegiada, suborno e fraude computacional também podem produzir grandes lucros e criar o incentivo para legitimar os ganhos ilícitos por meio da lavagem de dinheiro. (FATF, 2015).

Em resposta à crescente preocupação com a ameaça que a lavagem de dinheiro representa aos países, ao sistema bancário e às instituições financeiras em nível internacional, foi criado, em 1989, o *Financial Action Task Force* (FATF), um organismo inter-governamental liderado por presidentes, primeiro ministros e chanceleres, à época representantes dos sete países membro, grupo mais industrializado e desenvolvido economicamente do mundo G7. Na ocasião, o grupo de países foi formado por Canadá, França, Alemanha Ocidental, Itália, Japão, Reino Unido, Estados Unidos, mais a Comissão Européia. De atuação permanente desde sua formação, o FATF é um órgão com caráter de decisão política internacional, que funciona para gerar a vontade política necessária para promover reformas legislativas e regulamentares nos países membro. (FATF, 1989).

Em 1990, foram geradas quarenta Recomendações do FATF e em 1996 as Recomendações foram revistas pela primeira vez para refletirem a evolução das tendências e técnicas de lavagem de dinheiro e ampliar o seu âmbito para além da lavagem de dinheiro e do tráfico de drogas. Em outubro de 2001 o FATF ampliou o seu mandato para atuar contra o financiamento dos atos e das organizações terroristas, dando um passo importante na criação de Nove Recomendações Especiais que tratam sobre o financiamento do terrorismo. As Recomendações do FATF foram revistas pela segunda vez em 2003, juntamente com as Recomendações Especiais e foram aprovadas por mais

de 180 países. (FATF, 2012a). A proliferação de armas de destruição em massa também é uma preocupação de segurança em nível mundial. Em 2008 o mandato do FATF foi expandido para tratar desse assunto. Para combater esta ameaça, o FATF adotou a Recomendação Especial número 7, destinada a garantir a aplicação coerente e eficaz das sanções financeiras específicas, quando este é o chamado feito pelo Conselho de Segurança da Organização das Nações Unidas (ONU). (FATF, 2012b).

Em fevereiro de 2012 o FATF completou a revisão de suas normas e publicou suas quarenta Recomendações. Esta revisão destinou-se a reforçar as garantias globais e proteger a integridade do sistema financeiro internacional, fornecendo aos governos recursos robustos para tomar medidas contra a criminalidade financeira. As Recomendações foram ampliadas para lidar com novas ameaças, a exemplo do financiamento da proliferação de armas de destruição em massa. Para dar maior transparência às ações, e para que sejam mais duras contra a corrupção, nove Recomendações Especiais sobre o combate ao financiamento do terrorismo foram totalmente integradas com as medidas contra a lavagem de dinheiro. Isso resultou em um conjunto robusto e claro dos padrões globais, aprovados internacionalmente contra a lavagem de dinheiro e financiamento do terrorismo. As Recomendações do FATF sofrem atualizações no mês de Outubro de cada ano. (FATF, 2015).

O FATF está permanentemente em diálogo estreito e construtivo com o setor privado, a sociedade civil e outras partes interessadas. A revisão das Recomendações envolve extensa consulta e beneficia-se dos comentários e sugestões desses parceiros importantes para garantir a integridade do sistema financeiro. Em sua evolutiva atuação, produz orientações de melhores práticas, estas servem para orientar os países na aplicação das suas Recomendações. (FATF, 2015). Como auxiliares para efetividade de conformidade da aplicação das Recomendações, foram publicados documentos complementares: i) que trata de metodologia de aplicação, *For assessing Technical Effectiveness of AML / CFT Systems Compliance with the FATF Recommendations and the Effectiveness of AML / CFT Systems* (FATF, 2013a); e ii) que trata de procedimentos de avaliação mútua entre os membros, em sua quarta edição, *Procedures for the FATF Fourth Round of AML / CFT Mutual Evaluations*. (FATF, 2013b). Esses padrões dão transparência e permitem aos países agirem com sucesso contra o uso ilícito de seus sistemas financeiros.

Para dar a amplitude necessária na implementação de suas Recomendações, o FATF

depende de uma forte rede global, são organismos formados por grupos regionais de países chamados *FATF-Style Regional Bodies* (FSRBs), além de seus 36 países membro, atualmente, há oito FSRBs que têm papel importante na promoção da aplicação efetiva das Recomendações do FATF, no fornecimento de conhecimento e de cooperação na elaboração de suas políticas. Há mais de 180 jurisdições comprometidas com as Recomendações do FATF, pela rede de FSRBs e por associações globais ligadas ao FATF, estas, formadas por entidades relacionadas a órgãos responsáveis por movimentações financeiras, de controle e de combate a ilícitos financeiros e de combate a crimes contra a vida.

No ano de 2015, o FATF avaliou as medidas tomadas para combate do financiamento do terrorismo nas 198 jurisdições. Constam nos relatos da Conferência do G20 em novembro de 2015 que, enquanto a maioria dos países tiveram marcos legais abrangentes e significativos, muitos países ainda não possuem medidas básicas locais para combater o financiamento do terrorismo. A partir da Conferência do G20 o FATF intensificou esforços na aplicação de suas normas. Desde então, mais de 50 países iniciaram e concluíram alterações em suas legislações. (FATF, 2016a).

Em sua reunião plenária que ocorreu nos dias 17 a 19 de Fevereiro de 2016, em Paris, o FATF (2016b) tratou dos problemas existentes, as soluções alcançadas, as estratégias de ação, criação de padrões de controle e a efetiva troca de informações entre os seus membros no combate ao financiamento do terrorismo. O FATF congratulou-se com o progresso do Brasil sobre o congelamento de bens de terroristas, mas acentua que mais melhorias são necessárias para satisfazer plenamente as suas normas. A entidade manifestou preocupação com o continuado fracasso do Brasil na solução das deficiências identificadas na terceira avaliação mútua realizada em junho de 2010. O Brasil não criminalizou o financiamento do terrorismo desde 2004 por ocasião da segunda avaliação mútua. Em razão disso, o Brasil é chamado a cumprir o seu compromisso de adesão ao FATF adotando lei contra o financiamento do terrorismo que regule adequadamente as lacunas existentes de acordo com suas Recomendações, tendo como prazo a reunião plenária em fevereiro de 2016. (FATF, 2016b). Na reunião plenária de 22 a 24 de Junho de 2016, conforme FATF (2016c) a entidade congratula-se com o Brasil pela adoção da lei nº 13.260, de 16 de Março de 2016 que atende parcialmente a recomendação de acordo com FATF (2011, 2013a), sobre a criminalização do financiamento do terrorismo e salienta que o país deu um passo significativo no sentido de melhorar seu regime de

combate ao financiamento do terrorismo. Como consequência, o FATF decidiu não considerar os próximos passos no processo de acompanhamento, mas salienta que apesar deste passo positivo, ainda há uma série de deficiências do regime brasileiro no combate ao financiamento do terrorismo. Assim sendo, o FATF (2016c) recomenda que o Brasil deve eliminar estas deficiências, a fim de atingir um nível satisfatório de conformidade com os padrões internacionais.

### 2.2.2 Combate à Lavagem de Dinheiro no Brasil

Uma das principais características da criminalidade contemporânea é seu caráter transnacional. Com frequência, a atividade criminal envolve diversos países. Moro (2010) destaca que em um deles o crime é planejado, noutro executado e o produto da atividade criminal é enviado a um terceiro. No Brasil, a Lei nº 9.613, de 1998, em seu artigo 2º, II, faz explícita referência à possibilidade de que o crime antecedente ao da lavagem de dinheiro ocorra em outro país. Há entendimento que a prova indiciária é fundamental no processo por crime de lavagem de dinheiro, inclusive quanto à prova de que o objeto da lavagem é produto de um crime antecedente. (MORO, 2008). Portanto, para o recebimento da denúncia basta a **prova indiciária**, ou seja, ainda não categórica do crime antecedente e a bem da verdade do próprio crime de lavagem, como é a regra geral para recebimento da denúncia em qualquer processo criminal. Já para a condenação será necessária prova categórica do crime de lavagem, o que inclui prova convincente de que o objeto desse delito é produto de crime antecedente. (MORO, 2010).

O caráter transnacional do crime de lavagem de dinheiro e a necessidade da cooperação internacional motivou o Brasil a firmar sua participação como membro do FATF no ano 2000. A partir de então, consta nos registros da entidade em seu relatório anual (FATF, 2000), bem como passou a ser membro do grupo regional dos países da América Latina – *Financial Action Task Force of Latin América* (GAFILAT). Dos compromissos assumidos dentre as atribuições do Ministério da Justiça do Brasil, está a responsabilidade de planejar e executar políticas públicas que propiciem o combate à lavagem de dinheiro e à corrupção no país. Em 2003, ocorreu ação conjunta (Governo, Judiciário e Ministério Público) para desenvolverem estratégias de ação permanentes com a criação de tribunais especializados em matéria de lavagem de dinheiro e combate à corrupção.

Dentre as ações executadas, conforme Enccla (2004), destacam-se: a coordenação do processo de recuperação de ativos enviados para o exterior, pelo Departamento de Recuperação de Ativos e Cooperação Jurídica Internacional da Secretaria Nacional de Justiça (DRCI) / (SNJ); criação da gestão da Estratégia Nacional de Combate à Corrupção e à Lavagem de Dinheiro (ENCCLA); a coordenação da Rede de Laboratórios de Tecnologia contra à Lavagem de Dinheiro (LAB-LD); o Programa Nacional de Capacitação e Treinamento para o Combate à Corrupção e à Lavagem de Dinheiro (PNLD) e a criação do Conselho de Controle de Atividades Financeiras (COAF).

O Conselho de Controle de Atividades Financeiras (COAF) é o órgão criado no âmbito do Ministério da Fazenda, instituído pela Lei nº 9.613, de 1998, que atua eminentemente na prevenção, combate à lavagem de dinheiro e ao financiamento do terrorismo. As competências do COAF estão definidas nos artigos 14º e 15º da referida lei e tem a principal função de receber, examinar e identificar as ocorrências suspeitas de atividades ilícitas. (COAF, 2015a). São ainda funções do COAF comunicar as autoridades competentes para a instauração de procedimentos cabíveis nas situações em que o Conselho concluir pela existência de fundados indícios de crimes, de lavagem, ocultação de bens, direitos e valores, disciplinar e aplicar penas administrativas. (COAF, 2015b). Como entidade específica, o COAF organiza os regramentos de todas as entidades nacionais relacionadas com a contabilidade, auditorias e demais registros de movimentações de valores. O COAF tem como principal pressuposto que as entidades, físicas ou jurídicas, têm a responsabilidade de realizar suas operações com os devidos registros, seguindo os preceitos legais e em sendo ao contrário, estarão sujeitos aos rigores da lei que infringirem. COAF (2015b).

O COAF tem ainda a competência de regular os setores econômicos para os quais não haja órgão regulador ou fiscalizador próprio. Nesses casos, cabe ao mesmo definir as pessoas abrangidas, os meios e os critérios para envio de comunicações, bem como a expedição das instruções para a identificação de clientes e manutenção de registros de transações, além da aplicação de sanções previstas no artigo 12º da lei de sua criação. (COAF, 2015a). Como parte da responsabilidade legal, devem ser comunicadas ao COAF todas as operações e propostas realizadas pelas Entidades Reguladas. As Comunicações de Operações Automáticas (COA) são comunicações efetuadas pelos Setores Obrigados nos termos do artigo 11º da Lei nº 9.613, de 1998. Essas comunicações são realizadas sem análise de mérito, em razão de valores ou situações previamente defini-

das nas normas emitidas pelos órgãos reguladores. As Pessoas Obrigadas são pessoas físicas ou jurídicas listadas em (COAF, 2015c), e detalhadas em (COAF, 2015d) referente ao artigo 9º da Lei nº 9.613 de 1998 e são aquelas para as quais a lei impõe obrigações para a prevenção e combate ao crime de lavagem de dinheiro e de financiamento do terrorismo. As obrigações estabelecidas pelos artigos 10 e 11 da Lei referem-se ao dever de identificar clientes, manter registros e comunicar operações financeiras ao órgão responsável, observadas as orientações contidas nos normativos específicos emitidos pelos órgãos reguladores das respectivas pessoas físicas e jurídicas sujeitas a essa Lei. Por outro lado, as Comunicações de Operações Suspeitas (COS) são comunicações efetuadas ao COAF pelos Setores Obrigados, levando-se em conta as partes envolvidas, valores, modo de realização, meio e forma de pagamento, além daquelas que, por falta de fundamento econômico ou legal possam configurar sérios indícios da ocorrência dos crimes previstos na Lei nº 9.613, de 1998, ou com eles relacionar-se. (COAF, 2015d). As comunicações de Operações financeiras devem ser encaminhadas ao COAF, no prazo de 24 horas a contar da conclusão da operação ou da proposta de operação, conforme o inciso II do artigo 11º da Lei nº 9.613, de 1998. (COAF, 2015d).

Em razão da manutenção da conformidade e da transparência das suas atividades, as Pessoas Obrigadas das entidades reguladas pelo Ministério da Fazenda, através do COAF, devem periodicamente realizar a Comunicação de Não Ocorrência ou Declaração Negativa de Ocorrência. Essa comunicação é o ato pelo qual a Pessoa Obrigada deverá comunicar ao órgão regulador ou fiscalizador da sua atividade a não ocorrência de propostas, transações ou operações passíveis de serem comunicadas ao COAF na periodicidade e forma definidas pelos mesmos. Esse prazo está definido pelo COAF. Em sua maioria, as Comunicações de Não Ocorrência tem prazo anual, tendo como data limite os dias 10, 15, e 31 de Janeiro de cada ano calendário. A exceção se dá a entidades reguladas pela Superintendência de Seguros Privados (SUSEP), uma vez que essas devem fazer a Comunicação de Não Ocorrência mensalmente até o dia 20 do mês subsequente. (COAF, 2016).

### 2.2.3 A Prevenção da Lavagem de Dinheiro

O tratado internacional, ocorrido em 1930 na Conferência de Haia, deu origem ao Banco Internacional de Mediação - *Bank for International Settlements* (BIS), com sede



na Suíça, seus membros são os bancos centrais e as autoridades monetárias. O BIS tem a missão de ser o regulador dos bancos centrais na manutenção da estabilidade monetária e financeira e promover a cooperação internacional entre seus membros. A Comissão de Valores Mobiliários - *Securities and Exchange Commission* (SEC) foi fundada em 1929 antes da queda do mercado acionário dos Estados Unidos. Segundo Bertoni e Carvalho (2014), em 1960 a SEC passou a realizar um movimento de orientação à contratação de responsáveis por conformidade (*compliance officers*) para a criação de procedimentos internos de controle, treinamento de pessoas, o monitoramento e supervisão de atividades suspeitas. A criação do BIS foi o marco inicial das atividades de auditoria e controles internos das instituições na prevenção de fraudes e a lavagem de dinheiro, atividades essas reforçadas com a aprovação da lei Sarbanes-Oxley (SOX), dos USA em 2002, tornando as organizações que atuam no mercado de capitais, entidades obrigadas à prática da governança corporativa e de controles internos robustos, além de primarem pela transparência em seus relatórios nas publicações legais. (BERTONI; CARVALHO, 2014).

O desenvolvimento tecnológico possibilitou novos e sofisticados meios de fraudar e burlar os meios legais de negociação com amplitude internacional. Daniali (2014) trata da prevenção da lavagem de dinheiro por meio eletrônico (*e-money launderig*), assinala a dependência na robustez dos controles internos e das ações de governança corporativa, aliados à auditorias sistemáticas nos Sistemas de Informação para tornar os processos seguros. Pinho (2015) relata a situação da economia angolana que afasta-se da dependência exclusiva do petróleo. Seu sistema bancário caracteriza-se pela intensa movimentação de capitais gerados nos setores formal e informal e há sérios fatores de risco sobre a origem dos fundos e dos fluxos financeiros, que em alguns casos provêm de atividades criminosas. Essa realidade fez os bancos revisarem rapidamente seus sistemas de controle, estabelecendo metodologias e procedimentos de controle interno de acordo com as orientações emitidas pelo organismo nacional de supervisão, Banco Nacional de Angola (BNA) e de outros organismos internacionais, a exemplo do FATF.

No Brasil, com a adoção da Lei nº 9.613 de 1998 que criou o COAF e a publicação dos normativos, passou-se a exigir a implementação de controles internos e conformidade (*compliance*) dos setores sujeitos (obrigados) à lei de lavagem de dinheiro. Amorim, Cardozo e Vicente (2012) constataram que houve uma evolução no número de Relatórios de Inteligência Financeira elaborados e crescimento do número de inves-

tigados, indiciados e condenados e seus valores bloqueados na prevenção e combate à lavagem de dinheiro. Ao comparar os números anteriores à aplicação da lei, Amorim, Cardozo e Vicente (2012) identificaram que as comunicações de atividades suspeitas por parte dos setores obrigados, na maioria dos anos, cresceu 300% e em alguns anos cresceu até 500%, sinalizando excelente resultado na aplicação dos padrões de conformidade à lei. Dada a importância da competência dos responsáveis por conformidade (*compliance officers*) para avaliação dos riscos de uma organização influenciada por fatores internos e externos, Isa et al. (2015) enfatizam que essa função tão melhor será desempenhada de acordo com a robustez dos sistemas de controle internos e a aplicação de tecnologia na identificação dos riscos tão melhor quanto conhecer o perfil de seus clientes e a origem de suas transações.

Casos de repercussão internacional, no período de 1995 a 2015 foram enumerados por Suspiro (2015), envolvendo entidades financeiras, empresas e governos em escândalos financeiros, fraudes, tráfico de influência e lavagem de dinheiro, a exemplo do BARINGS, ENRON, Worldcon, Parmalat, Lehman Brothers, Siemens, Petrobrás, Toshiba, Volkswagen, dentre outros. Esses fatos reforçam a necessidade de maior rigidez nas funções de auditoria, controle interno e de governança corporativa para atender a conformidade com os requisitos legais de combate à corrupção, financiamento do terrorismo e a lavagem de dinheiro em caráter mundial.

### **2.3 Auditoria em Lavagem de Dinheiro**

Santos, Diniz e Corrar (2005) desenvolveram um estudo de caso em vinte municípios do Estado da Paraíba para verificar a eficácia da Lei de Benford. Utilizaram uma amostra de uma população de aproximadamente 104 mil notas de empenho. A fundamentação do estudo foi através do método dedutivo e aplicação prática da Lei de Benford, abstraindo o 1º dígito e medindo o nível de significância por meio de testes de aderência  $Z$  e  $\chi^2$ , introduzidos por Nigrini (2000) para determinar a concordância com valores esperados. Constataram que os testes adotados foram capazes de demonstrar a aderência dos dados à Lei de Benford.

Os autores definem a Lei de Benford como robusta para auxiliar o Auditor a detectar erros contábeis, simples erros de digitação e fraudes contábeis no contexto de uma auditoria contábil tradicional ou de uma auditoria digital. Destacaram ainda que é possível

inferir que a Lei de Benford é segura e eficaz no planejamento de amostras em razão de que os erros são sensivelmente minimizados e conhecendo o comportamento natural do sistema decimal, extratos que não se moldam à Lei de Benford são passíveis de uma maior concentração de testes de auditoria como forma de validação dos demonstrativos contábeis. Constataram também que esta é uma distribuição capaz de determinar o comportamento padrão das despesas praticadas pelos gestores públicos, traçando o perfil no qual a composição do primeiro dígito deve se comportar. Outra importante constatação dos autores é que a análise feita pela Lei de Benford foi capaz de explicar fatos que derivam de uma análise qualitativa. Os dados que apresentaram problemas de aderência à Lei de Benford são de municípios que realmente apresentaram problemas de gestão de conformidade com dados do Tribunal de Contas do Estado estudado. As conclusões finais permitiram identificar que a Lei de Benford é eficaz e precisa. Santos, Diniz e Corrar (2005) evidenciaram o indício de fraude por superfaturamento e fracionamento de despesas para burlar o limite estabelecido pela Lei Federal nº 8.666 /1993, que disciplina regras de aquisições mediante licitações.

Johnson e Weggenmann (2013) realizaram um estudo nos Relatórios financeiros dos cinquenta estados nos Estados Unidos. Por força legal, as demonstrações financeiras de fundos governamentais americanos utilizam o regime de competência modificado da contabilidade e devem ser auditados por auditores independentes, conforme os critérios do Conselho Governamental de Contabilidade - *Governmental Accounting Standards Board* (GASB). (GASB, 2016). As auditorias garantem a confiabilidade necessária para outros governos, órgãos legislativos, agências federais, agências de classificação de risco, os meios de comunicação e a sociedade americana. O estudo foi baseado na localização geográfica e a representatividade das populações, versa sobre os indicadores das demonstrações financeiras de Governo em Receitas do Resultado Primário de Governo, Saldo Geral de Fundos e o Fundo Total de Equilíbrio Governamental, com o objetivo de verificar se essas contas estão em conformidade com a Lei de Benford. Essas contas são as mais importantes referências da saúde financeira dos estados e oferecem amplo espectro de informações. Para desenvolver seu trabalho os auditores dependem de técnicas sofisticadas. Johnson e Weggenmann (2013) citam a aplicação de recursos como MS Excel para organizar os dados coletados de três anos consecutivos, que resultaram em 150 pontos para a análise de cada conta, a variação de Drake e Nigrini (2000) para a Média do Desvio Absoluto - *Mean Absolute Deviation* (MAD) e o Idioma

de Comandos de Auditoria - *Audit Command Language* (ACL), *software* que tem como função filtrar os dados duplicados e realizar outras análises digitais. Esse *software* inclui o código de análise da Lei de Benford, que permite ao auditor iniciar a identificação de anomalias nos dados. Durtschi, Hillison e Pacini (2004) afirmam que a aplicação da Lei de Benford é mais eficaz em grandes conjuntos de dados transacionais que representam mais de uma distribuição. Segundo Hill (1995), com média maior que a mediana e assimetria positiva, evitando assim os erros tipo I, falsos positivos e tipo II, falsos negativos. (MOOD; GRAYBILL; BOES, 1974).

O estudo de Johnson e Weggenmann (2013) demonstra que a Lei de Benford é eficaz em detectar desvios em pequenos conjuntos de dados, desde que acompanhados de testes de aderência adequados como o teste  $\chi^2$  e o teste  $Z$ , evitando erros e perda de tempo investigando dados não representativos. Como conclusão, os autores evidenciaram que a aplicação da Lei de Benford nas auditorias em dados financeiros governamentais é adequada e confiável para demonstrar o grau de credibilidade e transparência da gestão de recursos aos cidadãos.

Segundo Cunha e Bugarin (2014), as auditorias em obras públicas têm gerado oportunidade de aprofundamento de verificação da aplicação da verba pública por parte do governo brasileiro, recentemente mais discutidas estão as relacionadas com as obras realizadas nos estádios para a Copa do Mundo de 2014. A exemplo de Nigrini (2011, 2012), autor que introduziu a aplicação da Lei de Benford na prática da contabilidade forense nos Estados Unidos, Cunha e Bugarin (2014) realizaram um estudo nos dados da auditoria realizada em 2014 pelo Tribunal de Contas da União (TCU) sobre os recursos aplicados na reforma do Estádio do Maracanã. Os autores utilizaram a Lei de Benford associada à curva ABC aplicada na auditoria do TCU para a análise do sobrepreço nos valores praticados na obra. Além da análise dos Primeiros Dígitos de Benford, foi aplicada a análise dos dois primeiros dígitos e o teste da soma desenvolvido por Nigrini (2011, 2012), capaz de confrontar os resultados apurados, aumentando a capacidade de verificação da conformidade dos dados. O Teste da Soma, por sua vez, é um teste avançado, onde ao simular uma distribuição de Benford verifica-se que as somas dos números de cada grupo de primeiros dígitos 10, 11, ..., 99 resultam valores aproximadamente iguais, ou seja,  $1/90$ . Todavia, em Nigrini (2011, 2012) houve a constatação de que dados reais raramente se conformam a esse padrão, isto porque apresentam valores muito altos e muitas repetições de números de valor médio. Quando o teste da soma

é confrontado com o teste dos dois primeiros dígitos, também confronta-se o volume de dados contidos em cada grupo com a soma de seus valores. No caso de orçamento de obras públicas, identifica-se a materialidade e a relevância de cada grupo de dados, isso permite selecionar os dígitos que merecem uma análise mais crítica por parte do auditor. Em seus achados, Cunha e Bugarin (2014) recomendam a aplicabilidade da Lei de Benford em auditorias públicas. Como conclusão, identificaram valores com frequências excessivas pela análise dos dois primeiros dígitos de Nigrini (2011, 2012), caracterizando o sobrepreço equivalente a 71,54% dos valores apurados na auditoria do TCU.

## **2.4 Lei de Benford**

### **2.4.1 Conceito e Definição**

A Lei de Benford é uma distribuição estatística aplicada em diversas áreas do conhecimento, tais como estudos em vários ramos das ciências naturais realizados por Sambridge, Tkalčić e Jackson (2010), em pesquisas e análises de dados geofísicos para a identificação e determinação de sinais sísmicos na detecção de terremotos feitas por Díaz, Gallart e Ruiz (2014); Toledo, Riquelme e Campos (2015) na discussão sobre a análise desenvolvida por Pericchi e Torres (2011) em dados de processos eleitorais. A utilização da Lei de Benford tem sido ampliada por vários autores na identificação de suspeitas de fraude em dados contábeis e de movimentações financeiras. Nigrini (2011, 2012) a aplica na contabilidade forense em auditorias para detecção de fraudes e na análise e investigação de suspeitas de lavagem de dinheiro; Durtschi, Hillison e Pacini (2004) a utilizam na detecção de fraude em dados contábeis; Bhattacharya, Xu e Kumar (2011) a aplicam como sistema de apoio à análise e decisão sobre dados de relatórios financeiros; Clippe e Ausloos (2012) a utilizam como recurso estatístico na investigação e análise de dados de contribuições financeiras realizadas para entidades religiosas.

Frank A. Benford, que viveu de 1883 a 1948, foi físico pesquisador da GE e formulou a conhecida lei dos primeiros dígitos. A lei é uma descoberta realizada por Simon Newcomb, em 1881, ao observar que as primeiras páginas de um livro de logaritmos eram mais desgastadas do que as últimas. Em 1938, Frank A. Benford criou a lei que está baseada na distribuição estatística dos primeiros dígitos significativos de dados

oriundos de fenômenos naturalmente distribuídos. A distribuição de Benford (1938) revela que para conjuntos de dados, sejam contábeis, fenômenos naturais (natureza diversa), dados reais não viciados (sem manipulação), seguem a distribuição formulada por Benford. A Lei mostra que dentre os números significativos, isto é, 1, 2, 3, 4, 5, 6, 7, 8 e 9, verifica-se que o primeiro dígito tem maior probabilidade de ocorrer. Este pensamento não é trivial, uma vez que o senso comum deduziria que a distribuição dos primeiros dígitos significativos deveriam apresentar igual frequência de ocorrência.

Hill (1995, 1998) contribuiu para o tema apresentando a formulação matemática da Lei de Benford, além de demonstrar a origem de diversos tipos de dados que a seguem, como exemplo a extensão de rios, população, calor específico, constantes matemáticas e físicas. Em resumo, a Lei foi descoberta por Simon Newcomb, formulada por Frank A. Benford e a equação foi apresentada por Theodore P. Hill. A fim de simplificar a leitura deste trabalho, a distribuição de probabilidade baseada na Lei de Benford será denominada a partir desse ponto, Lei de Benford.

A formulação da Lei de Benford segue o modelo matemático apresentado na Equação (2.1)

$$P(d) = \log[1 + (1/d)], \quad (2.1)$$

onde  $P$  é a probabilidade de ocorrência do primeiro dígito significativo e  $d = 1, 2, \dots, 9$  é o valor do dígito significativo avaliado.

A Tabela 2 mostra a distribuição de probabilidades de cada dígito significativo segundo a Lei de Benford.

**Tabela 2:** Probabilidades dos primeiros dígitos de Benford

Dígito ( $d$ )	1	2	3	4	5	6	7	8	9	Total
$P(d)$	30,1%	17,6%	12,5%	9,7%	7,9%	6,7%	5,8%	5,1%	4,6%	100%

Fonte: Adaptado de Benford (1938)

Uma extensão da Lei de Benford é a distribuição dos dois primeiros dígitos aplicada por (NIGRINI, 2011, 2012) e está de acordo com o que segue: a frequência esperada da ocorrência de um dígito  $D_2 = d_2$  como segundo dígito em um número, dado que o primeiro dígito é  $D_1 = d_1$ , segundo a Lei de Benford é dada pelo modelo matemático apresentado na Equação (2.2).

$$P(D_1, D_2 = d_1, d_2) = \log[1 + (1/(d_1, d_2))], \quad (2.2)$$

onde  $P(D_1, D_2)$  é a probabilidade de ocorrência dos dois primeiros dígitos e

$$d_1, d_2 \in \{10, 11, \dots, 99\}.$$

A Tabela 3 mostra a distribuição de probabilidades dos dois primeiros dígitos significativos segundo esta abordagem.

**Tabela 3:** Probabilidades da distribuição dos dois primeiros dígitos

Dígitos ( $d_1, d_2$ )	0	1	2	3	4	5	6	7	8	9
1	4.14%	3.78%	3.48%	3.22%	3.00%	2.80%	2.63%	2.48%	2.35%	2.23%
2	2.12%	2.02%	1.93%	1.85%	1.77%	1.70%	1.64%	1.58%	1.52%	1.47%
3	1.42%	1.38%	1.34%	1.30%	1.26%	1.22%	1.19%	1.16%	1.13%	1.10%
4	1.07%	1.05%	1.02%	1.00%	0.98%	0.95%	0.93%	0.91%	0.90%	0.88%
5	0.86%	0.84%	0.83%	0.81%	0.80%	0.78%	0.77%	0.76%	0.74%	0.73%
6	0.72%	0.71%	0.69%	0.68%	0.67%	0.66%	0.65%	0.64%	0.63%	0.62%
7	0.62%	0.61%	0.60%	0.59%	0.58%	0.58%	0.57%	0.56%	0.55%	0.55%
8	0.54%	0.53%	0.53%	0.52%	0.51%	0.51%	0.50%	0.50%	0.49%	0.49%
9	0.48%	0.47%	0.47%	0.46%	0.46%	0.45%	0.45%	0.45%	0.44%	0.44%

Fonte: Elaborada pelo autor adaptado de Benford (1938)

Exemplificando a amplitude diversa de aplicações da Lei de Benford, seguem-se exemplos de estudos realizados.

#### 2.4.2 Aplicações da Lei de Benford

Com base nas experiências de Hill (1998), Nigrini e Mittermaier (1997), que comprovam a robustez da aplicação da Lei de Benford em diversas áreas do conhecimento, principalmente na área contábil e financeira desde a década de 1980, Rocha e Cymrot (2009) realizaram um estudo de confiabilidade do controle de resíduos de processos industriais que afetam o meio ambiente através da análise da aderência dos dados à Lei de Benford. Os conjuntos de dados analisados são de processos de fundição, correspondem à medida do comprimento de um tipo de filtro para metais líquidos e o teor de alumínio em porcentagem retirados do banco de dados do Controle Estatístico de Processos

(CEP) da organização. Para os testes de hipóteses foi definido  $H_0$  como sucesso e  $H_1$  como fracasso da aderência das amostras à Lei de Benford. Para testar a aderência à distribuição, foi aplicado o teste  $\chi^2$  com nível de significância de 1% e  $k - 1 = 8$  graus de liberdade.

Foram testadas amostras de teor de alumínio também para o terceiro e quarto dígitos, nesse caso, a aderência do terceiro dígito comprovou a aleatoriedade dos dados e a não existência de interferências no processo, o quarto dígito apresentou-se não aderente à Lei de Benford devido ao uso de arredondamento de valores, também comprovando a eficácia da Lei de Benford na detecção de possível manipulação nos dados. (ROCHA; CYMROT, 2009). A análise da amostra do comprimento do filtro necessitou uma abordagem diferente por se tratar dos limites inferior e superior de uma especificação, uma vez que possui número diferente de dígitos significativos. Como os dados tem limitação imposta pela especificação de produto, foi necessário o teste da aderência à Lei de Benford nos dígitos livres de limites externos impostos. Corroborando os dizeres de Hales et al. (2008): a Lei de Benford é simples, pouco onerosa, precisa e segura para a análise estatística da confiabilidade de dados para a tomada de decisão em processos industriais, como visto no estudo de Rocha e Cymrot (2009) envolvendo dados de sustentabilidade e preservação do meio ambiente.

Ao realizar um estudo para identificar a possível manipulação de dados reais de controle diário de refugo de um dos produtos em uma empresa multinacional de autopeças, Favarettto (2007) relata que o interesse em analisar esses dados se deu em razão de que parte da remuneração dos operadores é inversamente proporcional à quantidade de refugo. Foram utilizados 34.200 registros do sistema ERP da organização, dados obtidos por observação direta dos próprios operadores na identificação das causas de refugo. Ao analisar a aderência dos dados à Lei de Benford, Favarettto (2007) afirma que a mesma mostrou-se adequada para análises primárias da qualidade de dados e que a adulteração de uma pequena parte poderia ser detectada. O resultado permitiu deduzir que a Lei de Benford é uma distribuição útil para análise da qualidade de grande volume de dados industriais e, portanto, recomendada para aplicação em estudos futuros.

Com o intuito de verificar se a Lei de Benford é obedecida em dados de origem biológica, Cáceres et al. (2008) realizaram um estudo a partir de diferentes fontes de dados: comprimento do gene de bactérias, incidência de sarampo antes da vacinação e de incidência epidêmica em valores absolutos de gravações de Magneto-encefalogramas



(MEG) humanos. Os pesquisadores constataram que a frequência dos primeiros dígitos da Lei de Benford é obedecida em um caso de dados epidemiológicos, em um foco de disseminação putativa (incidência do sarampo na pré-vacinação em Preston, Inglaterra). No entanto, foram observadas amostras peculiares para a distribuição do comprimento do gene em microorganismos, magneto-encefalogramas (MEG) e em dados epidemiológicos agrupados de grandes regiões geográficas. As frequências dos primeiros dígitos foram calculadas e confirmadas em testes de aderência. Em média, os dados coletados em um cenário de eventos que se propagam podem explicar pequenas distorções observadas na frequência dos primeiros dígitos da Lei de Benford. Como conclusões do estudo, Cáceres et al. (2008) referem que isto ajudou a entender o comportamento dos conjuntos de dados das epidemias, mostrando a presença de ondas que se propagam e revelam a peculiaridade dos dados reais de origem biológica ao aderirem à Lei de Benford.

Os exemplos citados demonstram a utilidade da Lei de Benford na avaliação da aleatoriedade, da influência intencional ou na alteração de resultados observacionais. Em todos estudos, as conclusões confirmam o uso científico da Lei de Benford de forma segura e precisa. Rocha e Cymrot (2009) e Favarettto (2007) confirmaram em suas análises no meio industrial, i) a confiabilidade dos dados para tomada de decisão ii) a capacidade de análise em grande volume de dados e de identificação de manipulação dos mesmos. Já Cáceres et al. (2008) identificaram a aderência à Lei de Benford em dados biomédicos, portanto, a lei é aplicável também em abordagens na área da saúde.

Pesquisas realizadas comprovam a utilidade da aplicação da Lei de Benford também na identificação de fraudes. Lagioia et al. (2011) fazem uso da Lei de Benford para análise de dados quantitativos em empresas fiscalizadas na prestação de contas de Impostos Sobre Serviços (ISS), sujeitas a penas de acordo com a Lei nº 8.137, de 1990. O uso da Lei de Benford sinalizou a aderência à distribuição, apontou desvios e fraudes na identificação da emissão de notas fiscais frias. A pesquisa tem sua relevância pela comprovação da aderência à Lei de Benford, além da identificação de desvios. Quando do reteste do modelo com dados corretos (expurgo dos dados manipulados), os resultados mostraram aderência à probabilidade esperada. Com a aderência ao modelo da Lei de Benford, tornou-se pioneiramente possível antever a sinalização de fraudes e desvios de dinheiro, antes somente baseada em confronto com os dados dos tomadores de serviço por ocasião de intensa análise documental. Lagioia et al. (2011) concluíram que

o estudo mostra a conveniência da aplicação da Lei de Benford nas auditorias municipais em tributos, pela simplicidade, capacidade de tratar grande volume de dados e principalmente pelo baixo custo de aplicação.

Forster (2010) estudou a aplicação da Lei de Benford em atividade do Terceiro Setor, mais especificamente em 159 instituições sem fins lucrativos do Distrito Federal, nas contas Caixa, Bancos e Despesas dos anos de 2002 e 2003. Para confirmar a aderência dos dados à Lei de Benford, procedeu dois testes de regressão,  $Z$  e  $\chi^2$ . Como resultado dos testes, identificou o teste  $Z$  em consonância com a Lei de Benford com exceção do primeiro dígito 2 na conta de Receita e dígitos 5, 6 e 9 para a conta de Despesa, ambas do ano de 2002. Com relação ao teste  $\chi^2$ , a única conta que não apresentou aderência nos testes foi a de Despesa do ano de 2002. O autor ressalta que o fato de parte dos dados não aderirem à Lei de Benford não significa, a princípio, erro ou fraude, as variações podem se relacionar a dados corretos que não estejam em consonância com a Lei de Benford. Ao considerar o alcance da Lei de Benford, verificando o comportamento das frequências dos segundos dígitos, os mesmos mostraram resultados homogêneos ao redor de 1/10 para cada um. Conclusão para esse conjunto de dados, a Lei de Benford é válida para os primeiros dígitos e não para os segundos, com base nisso o autor destaca que a maior parte dos dados aderiu à Lei de Benford com baixa presença de erros ou fraudes nas instituições estudadas. (FORSTER, 2010).

Para confirmar a aplicabilidade da Lei de Benford em processos eleitorais, foi realizado um estudo nas eleições presidenciais dos Estados Unidos de 2004, para governador de Porto Rico, em 1996, 2000 e 2004 e o referendo presidencial da Venezuela, em 2004. Pericchi e Torres (2011) constataram a clareza que a distribuição permite ao analisar os dados, utilizando as variações ao realizarem testes com o primeiro e o segundo dígitos de Benford, confirmando a aderência pelos resultados obtidos. As análises foram organizadas para identificar a aderência à distribuição de Benford para o primeiro e para o segundo dígitos por meio do teste  $\chi^2$  dado pelo tamanho da amostra e o  $p$  - valor, dado pela probabilidade de que o  $\chi^2$  calculado seja maior ou igual ao  $\chi^2$  observado, para a hipótese nula verdadeira com número mínimo de 2.250 eleitores, o que representa o melhor ajuste para um mínimo de 999 simulações. De acordo com o método de calibração do  $p$ -valor de Sellke, Bayarri e Berger (2001), se a hipótese nula é aceita os dados seguem a distribuição, se ao contrário a hipótese nula seja rejeitada os dados não seguem a distribuição. As análises permitiram identificar a melhor aderência dos

dados à Lei de Benford no escrutínio dos Estados Unidos tanto para o primeiro quanto para o segundo dígito, sem necessidade do uso de restrição de tamanho da amostra. Os dados de Porto Rico não apresentaram aderência à distribuição nos testes do primeiro dígito, bem como na aplicação da restrição de tamanho da amostra, no entanto, ocorreu aderência nos testes do segundo dígito. Já os dados da Venezuela foram rejeitados em todos os testes, com ou sem restrição, demonstrando a não aderência à distribuição, ou seja, o comportamento dos dados são tendenciosos e necessitam revisão minuciosa nos resultados. As conclusões que Pericchi e Torres (2011) chegaram a partir de repetidos processamentos para assegurar a confiabilidade dos resultados, são de que a qualidade dos dados não está relacionada ao método de escrutínio, manual ou eletrônico, mas sim com a forma de tratamento do evento (eleições). O grau de exigência do método de ajuste e a não limitação de estudo apenas ao primeiro dígito, como a proposta pelos autores em testar a distribuição no segundo dígito, tem se firmado como recurso preciso e eficaz em eleições majoritárias. (MEBANE JR, 2006).

Em respeito ao uso de análises digitais em auditorias, Silva, Korzenowski e Vaccaro (2014) desenvolveram a aplicação de um método científico baseado em testes de hipóteses através da verificação da sua aderência à Lei de Benfor, com o objetivo de aprimorar os controles internos das organizações na verificação de suspeitas de fraude por Lavagem de Dinheiro. Durtschi, Hillison e Pacini (2004) afirmam que, aplicada de forma adequada, a Lei de Benford é útil na identificação de suspeitas de fraude em dados contábeis e financeiros, com a vantagem dos recursos de análise digitais baseadas na Lei de Benford estarem incorporadas em *softwares* populares, a exemplo do *software R*, disponibilizado em Team (2016). Durtschi, Hillison e Pacini (2004) salientam que o Instituto Americano de Contadores Públicos - *American Institute of Certified Public Accountants* (AICPA), em suas publicações, Declarações de Padrões de Auditoria - *Statements on Auditing Standards* (SAS) nºs 56 e 96, recomendam aos auditores a previsão de aplicação de testes analíticos robustos na fase de planejamento da auditoria e a utilização de especialistas na execução dos testes para identificarem transações não usuais e tendências que caracterizem suspeita de fraude em dados de demonstrações financeiras.

A Norma Brasileira de Contabilidade (NBC), NBC T 11.11 – Amostragem, aprovada pela Resolução do Conselho Federal de Contabilidade - CFC nº 1.012 de 21/1/2005, ao determinar a extensão de um teste de auditoria ou método de seleção de itens a testar, define que o auditor pode empregar técnicas de amostragem em suas abordagens.

Ribeiro e Filho (2007) propõem a não limitação de recursos estatísticos designados por uma norma, a exemplo da Lei *Sarbanes-Oxley* (SOX), que tem como princípio a governança corporativa e robustez dos controles internos e define que a lei exige e confere ao auditor a autoridade para decidir sobre a expertise estatística mais adequada com a realidade. Francis (2004) afirma que a correta forma de medir o risco inicia-se pela avaliação do sistema de controles internos da organização, sabidamente, o conjunto de medidas de proteção dos ativos contra erros e fraudes.

Drake e Nigrini (2000) e Nigrini (2011, 2012) tratam de demonstrar os fundamentos de uso e critérios de avaliação da aderência dos dados à Lei de Benford. Os padrões relativos à distribuição de Benford mostram-se valiosos para analistas de dados e especificamente para auditores na identificação de omissões, erros, distorções, fraudes e irregularidades, além de recomendações de funções analíticas relativas ao contexto de análise de dados forenses.

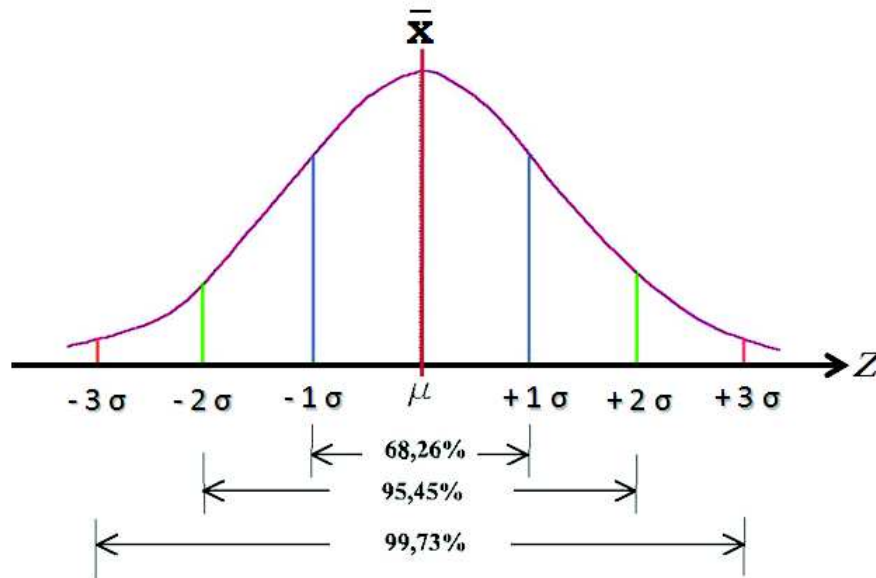
## 2.5 Cartas de Controle de Processos

A inferência estatística define que os eventos de processos, sejam naturais ou produtivos, seguem uma distribuição normal. Montgomery (2016) destaca que no Controle Estatístico da Qualidade a curva normal é provavelmente a mais importante distribuição, tanto na teoria quanto na prática da estatística. A importância da distribuição normal foi extensamente utilizada por Shewhart (1931) pela aplicação da mesma no Controle Estatístico de Processos. A distribuição normal é peculiar pela sua aparência, sua forma é de uma curva simétrica em forma de sino, unimodal. Sua larga aplicação gerou uma notação especial  $x \sim N(\mu, \sigma^2)$ , para dizermos que o evento  $x$  é normalmente distribuído  $N$ , tem média  $\mu$ , variância  $\sigma^2$  e desvio-padrão  $\sigma$ .

Montgomery (2016) descreve uma das mais importantes medidas de dispersão, o desvio-padrão  $\sigma$ . O mesmo tem uma interpretação simples, ilustrada na Figura 1. Note-se que 68,26% dos valores populacionais caem entre os limites centrados no valor da média populacional, mais ou menos um desvio-padrão ( $\mu \pm 1\sigma$ ); 95,46% dos valores populacionais caem entre os limites centrados no valor da média populacional, mais ou menos dois desvios-padrão ( $\mu \pm 2\sigma$ ); e 99,73% dos valores populacionais caem entre os limites centrados no valor da média populacional, mais ou menos três desvios-padrão ( $\mu \pm 3\sigma$ ). Concluindo, o desvio-padrão mede a distância na escala horizontal, asso-

ciada aos limites de abrangência percentuais representados pela área abaixo da curva normal, nas proporções de 68,26%, 95,46% e de 99,73%.

**Figura 1:** Área sob a curva da distribuição normal



Fonte: Montgomery (2016)

Outro importante recurso na aplicação da distribuição normal é a padronização da distribuição, onde pode-se calcular a função acumulada da **distribuição normal padrão** (média = 0, desvio-padrão = 1). Dessa forma, simplifica-se o cálculo, encontrando valores padronizados  $z$  em proporções de desvios-padrão, chamada de **padronização**.

Convertendo uma variável aleatória da distribuição normal  $N$ , com média  $\mu$  e variância  $\sigma^2$ ,  $N(\mu, \sigma^2)$ , em uma variável aleatória da distribuição normal  $N$ , com média  $\mu$  igual a zero e um desvio-padrão  $\sigma$  igual a 1,  $N(0; 1)$  e a forma simplificada de cálculo dos valores padronizados  $z$ , segundo Montgomery (2016) é dada pela Equação 2.3.

$$z = \frac{x - \mu}{\sigma} \quad (2.3)$$

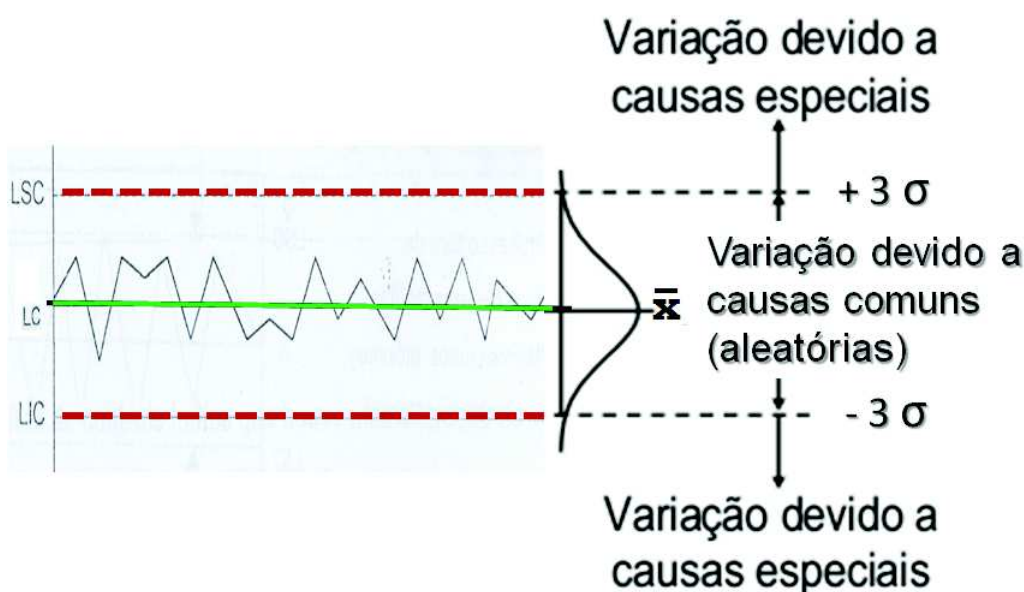
Shewhart (1931) introduziu os conceitos e as bases da aplicação de Cartas de Controle de processos com o objetivo de monitorar padrões de manufatura de produtos. Desde então, a estatística aplicada difundiu-se do meio industrial para outros ramos de atividade, áreas que ainda não haviam experimentado passaram a protagonizar a aplica-

ção de ferramentas estatísticas.

Como exemplo de novos ramos de atividade utilizando Cartas de Controle temos a área da saúde, onde Fry et al. (2012) usam no controle de evidências em procedimentos cirúrgicos e Nordstrom et al. (2012) utilizam em diagnósticos oncológicos. Na área Geoquímica, Andriotti (2007) utiliza para definir escopo de projetos em processos de geoprospecção. Em atividade relacionada ao solo, Milan e Fernandes (2002) aplicam para o ganho de produtividade na agropecuária.

Todos processos sofrem variações conforme sua natureza Shewhart (1931) identificou duas causas de variação, as causas **comuns** ou **aleatórias**, e as **atribuíveis** ou **especiais**. O detalhamento das Cartas de Controle e seus fatores influentes são mostrados na Figura 2.

**Figura 2:** Carta de Controle de Shewhart



Fonte: Adaptado de Montgomery (2016)

A típica Carta de Controle é uma representação gráfica de uma característica de qualidade ou de um evento, medidos ou calculados a partir de uma amostra *versus* o número da amostra ou do tempo decorrido. É composta de uma **linha central**, que representa o valor médio da característica da qualidade que corresponde ao estado **sob controle estatístico**, isto é, apenas as causas aleatórias estão presentes. Duas outras

linhas horizontais, chamadas de **limite superior de controle** (*LSC*) e **limite inferior de controle** (*LIC*), são também mostrados no gráfico e representam o intervalo de  $\pm 3\sigma$ , em relação à média.

As variações **comuns** ou **aleatórias** são o efeito cumulativo de muitas causas pequenas, essencialmente inevitáveis (naturais). No sistema de Controle Estatístico da Qualidade essa variabilidade natural é chamada de **sistema estável de causas aleatórias**. Um processo que opera apenas com as causas aleatórias de variação está **sob controle estatístico**, portanto, as causas aleatórias fazem parte dos processos.

As variações **atribuíveis** ou **especiais** são oriundas de interferências que ocorrem de forma excepcional no processo. Tal variabilidade é muito grande quando comparada às naturais. Estas fontes de variabilidade têm sua origem por interferência externa ao processo (não naturais), em geral são máquinas ajustadas inadequadamente, erros na operação, matéria prima defeituosa, excessiva interferência no processo e/ou dados manipulados. Diz-se que um processo que opera com presença de causas atribuíveis está **fora de controle estatístico**. (SHEWHART, 1931) e (MONTGOMERY, 2016).

A Carta de Controle é uma das técnicas para monitoramento do processo em tempo real, são utilizadas para estimar parâmetros de um processo e determinar sua capacidade. Ela pode também fornecer informação útil à melhoria do processo. Montgomery (2016) lembra que o objetivo do Controle Estatístico do Processo é a eliminação da **variabilidade no processo**, o que nem sempre é possível, mas com certeza a Carta de Controle é uma ferramenta eficaz na redução da variabilidade nos processos.

A aplicação das Cartas de Controle de Processos seguem critérios importantes para assegurar a sua adequada implementação. Montgomery (2016) orienta para a correta implementação de Cartas de Controle. Deve-se seguir os passos que envolvem o que se chama de **Fase I** e **Fase II**, com dois objetivos distintos. Na fase I, um conjunto de dados do processo é coletado e analisado de uma vez, em uma análise chamada de **retrospectiva**. Com base nessa análise, determina-se os **limites de controle de teste** para avaliar se o processo estava sob controle estatístico durante o período em que os dados foram coletados e para verificar se limites de controle confiáveis podem ser estabelecidos para o monitoramento de produção futura. Na fase I, as Cartas de Controle ajudam a levar o processo para um estado de **sob controle estatístico**. A fase II se inicia após a obtenção de um conjunto de dados chamado **limpo** (somente causas aleatórias agem sob o mesmo). Os dados são coletados sob condições estáveis e representativas do desem-

penho de um processo **sob controle estatístico**. Portanto, na fase II a Carta de Controle é usada para **monitorar** o processo compara-se a estatística amostral para cada amostra sucessiva, à medida que ela é extraída do processo com os seus limites de controle. Na fase I, os limites de controle são calculados a partir de  $m = 20$  ou mais subgrupos de dados. Os limites de controle são calculados com base nos  $m$  subgrupos e as estatísticas geradas são plotadas na Carta de Controle. Pontos que se situam fora dos limites de controle devem ser investigados a procura de causas atribuíveis, uma vez essas causas tenham sido identificadas as mesmas podem ser tratadas e eliminadas. (SHEWHART, 1931) e (MONTGOMERY, 2016).

Com relação ao número de características monitoradas por Cartas de Controle, define-se **Carta de Controle univariada** para processos onde apenas uma variável é monitorada e **Carta de Controle multivariada** para processos onde duas ou mais variáveis são monitoradas.

Quanto ao agrupamento das amostras para o cálculo das estatísticas, as mesmas devem ser tomadas na sequência temporal da ocorrência dos eventos geradores. Os cálculos das estatísticas podem ser feitos de duas formas i) dentro da **própria amostra**, onde as estatísticas são obtidas dentro da mesma amostra coletada e ii) entre **intervalos móveis**, ou seja, as estatísticas são obtidas por agrupamento de determinado número de amostras, onde elimina-se a amostra mais antiga do grupo de dados, incorpora-se uma amostra atual e assim sucessivamente no processo de monitoramento. O uso de intervalos móveis aplica-se em Cartas de Controle de Shewhart, quando há o interesse em detectar pequenas mudanças no processo, tanto em monitoramento de Cartas de Controle univariadas, bem como em Cartas de Controle multivariadas, a exemplo das Cartas de Controle *CUSUM*, *EWMA* e *MEWMA*, desenvolvidas por Roberts (1959), Crowder (1989) e Lucas e Saccucci (1990).

No planejamento do monitoramento por Cartas de Controle deve-se especificar tanto o **tamanho da amostra**, bem como a **frequência de amostragem**. A extração de grandes amostras muito frequentes favorecem a capacidade de detecção de mudanças no processo, no entanto, isso pode apresentar limitação econômica. A prática recente tende a favorecer as amostras pequenas, mais frequentes, principalmente na produção em alta escala ou em eventos, onde podem ocorrer um maior número de causas atribuíveis. Os sistemas de monitoramento automáticos, por meio de tecnologia e dos recursos computacionais facilitaram e tornaram mais eficazes as coletas de amostras em tempo real.



A probabilidade de ocorrer um ponto fora de controle determina a maneira de decisão sobre o método e a frequência de amostragem. A forma de medir a performance das Cartas de Controle é avaliada pelo Número Médio de Amostras Coletadas até a detecção de sinais, *Average Range Length* (ARL). Se as observações do processo não são correlacionadas, então, na aplicação de qualquer Carta de Controle de Shewhart, segundo Montgomery (2016), o ARL pode ser calculado conforme a Equação (2.4).

$$ARL = \frac{1}{p}, \quad (2.4)$$

onde  $p$  é a probabilidade de ocorrência de um ponto fora dos limites de controle.

Considerando a área sob a curva normal, temos que 99,73% dos valores populacionais caem entre os limites definidos pela média, mais ou menos três desvios-padrão ( $\mu + 3\sigma$ ), ou seja, em um processo onde somente causas comuns ou aleatórias estão presentes. Para o cálculo de  $p$  temos,  $1 - 0,9973$ , então  $p$  é igual a 0,0027.

Tomando a Equação (2.4) para o cálculo do ARL, o Número Médio de Amostras Coletadas até a detecção de sinais de que o processo está fora de controle é dado na Equação (2.5) a seguir.

$$ARL = \frac{1}{0.0027} = 370. \quad (2.5)$$

Portanto, em um processo onde somente causas comuns ou aleatórias estão presentes, o Número Médio de Amostras Coletadas até a detecção de um sinal de que o processo está fora de controle é igual a 370. Concluindo, teremos como probabilidade 370 pontos plotados na Carta de Controle até que um sinal aleatório se apresente fora de controle.

Caso estes sinais não sejam consequências de causas atribuíveis no processo, serão denominados de alarmes falsos, mas se são consequências de alguma causa atribuível no processo e a Carta de Controle não detectar teremos um erro de detecção. Geralmente a detecção de sinais é realizada por simulação e são calculadas estatísticas, a partir das quais é feita a avaliação da capacidade da Carta de Controle em detectar corretamente quando o processo saiu ou não de controle.

### 2.5.1 Carta de Controle de monitoramento multivariada $\chi^2$

Métodos multivariados são necessários sempre que se quer monitorar diversas variáveis de qualidade com o propósito de avaliar as relações entre elas. Um trabalho recente demonstrou que a utilização de métodos de seleção de variáveis pode melhorar o desempenho da Carta de Controle. Woodall e Montgomery (2014) recomendam a Carta de Controle  $\chi^2$  para atender esse propósito. Portanto, o modelo de Carta de Controle  $\chi^2$  é adequado para atender o monitoramento da aderência dos dados à Lei de Benford como proposta desse estudo. Umit e Cigdem (2001) executaram um estudo onde levantam as perspectivas históricas do controle de processos multivariados. Os autores mencionam diversas Cartas de Controle multivariado, entre elas a Carta de Controle  $\chi^2$ . Segundo os autores, o monitoramento e a detecção de pontos fora de controle depende da adequada definição dos limites de controle. O limite superior de controle (LSC) para a Carta de Controle  $\chi^2$  é dado pela Equação (2.6).

$$LSC = \chi_{\alpha, 2g.l.}^2 \quad (2.6)$$

A estatística  $\chi^2$ , por sua vez, é calculada para cada uma das amostras. Então, pode-se definir o vetor da estatística  $\chi^2$  conforme a Equação (2.7).

$$\chi_i^2 = n(x_j - \mu_j)\Sigma^{-1}(x_j - \mu_j); i = 1, \dots, t; j = 1, \dots, n, \quad (2.7)$$

onde  $n$  é o tamanho da amostra ou subgrupo,  $x_j$  é o valor da observação da variável  $j$  coletada no processo,  $\mu_j$  é a média da população para a variável  $j$ ,  $\Sigma$  é a matriz de covariâncias e  $i$  indica a amostra ou subgrupo coletados no tempo  $t$ . No caso específico do presente trabalho, tanto devido à magnitude dos valores observados - valores no intervalo (1;9), com a restrição de que a soma das proporções dos primeiros dígitos é igual a um, as estimativas de variâncias e covariâncias produziram matrizes  $\Sigma$  que não eram inversíveis ou que o resultado da estatística assumiram valores muito grandes, impossibilitando a plotagem para o monitoramento. Com isso, definiu-se a matriz  $\Sigma$  como uma matriz diagonal, sendo a diagonal composta pelas variâncias estimadas das proporções de cada dígito. Esta é uma suposição forte, pois assume a independência entre as proporções individuais do valor de cada primeiro dígito, o que de fato não é observado, devido à restrição imposta, onde a soma das proporções dos valores de todos

primeiros dígitos é igual a um.

Uma vez obtido um modelo de distribuição, baseado em dados discretos, supõe-se que seja possível monitorar um processo através de uma Carta de Controle  $\chi^2$ , baseando-se em uma estatística como a apresentada na Equação (3.2). Deste modo, dados discretos como os oriundos da análise do primeiro dígito significativo de lançamentos financeiros podem ser monitorados verificando continuamente a sua aderência a um modelo probabilístico, como o definido na Lei de Benford. Esta estratégia pode então ser adotada na execução de auditorias para identificação de alterações significativas no comportamento dos dados, levando à suspeita, por exemplo, da prática de crimes como o de lavagem de dinheiro.

### 2.5.2 Carta de Controle $T^2$ de Hotelling

No controle estatístico da qualidade univariada usa-se a distribuição normal para descrever o comportamento de uma característica de qualidade contínua. Segundo Montgomery (2016), a descrição da função densidade da distribuição normal é dada pela Equação (2.8).

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty. \quad (2.8)$$

onde  $f(x)$  é a função densidade,  $x$  é a variável,  $\mu$  é a média da distribuição normal e  $\sigma^2$  é a variância. Essa quantidade mede o quadrado da distância padronizada de  $x$  à média  $\mu$ , onde o termo "padronizada" significa que a distância se expressa em unidades de desvio-padrão  $\sigma$ .

Para a geração dos dados de controle (Fase I) utiliza-se a distribuição  $T^2$  de Hotelling. Mason, Tracy e Young (1997) destacam que a distribuição  $T^2$  de Hotelling é uma estatística comum usada em Cartas de Controle multivariadas para observações individuais. Ela é definida como a distância generalizada de uma amostra dimensional  $p$  do ponto  $X = (x_1, x_2, \dots, x_p)$  para sua média amostral. O vetor de médias e a matriz de covariância amostral são estimadores não viesados das quantidades populacionais correspondentes, isto é, visto da forma  $E(\bar{X}) = \mu$  e  $E(S) = \Sigma$ .

Segundo Mason, Tracy e Young (1997), a estatística de teste plotada na Carta de Controle  $\chi^2$  para cada amostra é conhecida como Distância de Mahalanobis – Equ-

ção (2.9)

$$\chi_0^2 = n(\bar{x} - \mu) \Sigma^{-1} (\bar{x} - \mu) \quad (2.9)$$

onde  $\chi^2$  (qui-quadrado) é a Distância de Mahalanobis,  $n$  é o tamanho da amostra ou subgrupos,  $\bar{x}$  é a média amostral e  $\mu$  é a média populacional.

Segundo Mason, Tracy e Young (1997), quando  $\mu$  (média populacional) é desconhecida recorre-se à aplicação de  $\bar{X}$  (média amostral) no cálculo de cada variável da amostra. Os termos  $(\bar{X}_j)$  são elementos do vetor  $\bar{X}$  e a variância  $S$  de dimensão  $p \times p$  das matrizes de covariância amostral se apresenta na matriz (2.10).

$$S = \begin{bmatrix} \bar{S}_1^2 & \bar{S}_{12} & \bar{S}_{13} & \dots & \bar{S}_{1p} \\ & \bar{S}_2^2 & \bar{S}_{23} & \dots & \bar{S}_{2p} \\ & & \bar{S}_3^2 & \dots & \bar{S}_{3p} \\ & & & \ddots & \vdots \\ & & & & \bar{S}_p^2 \end{bmatrix}. \quad (2.10)$$

Suponha que  $S$ , da Equação (2.10), seja utilizada para estimar o  $\Sigma$  e que o vetor  $\bar{X}$  seja o valor sob controle do vetor médio do processo. Substituindo  $\mu$  por  $\bar{X}$  e  $\Sigma$  por  $S$  na Equação (2.9), de acordo com Umit e Cigdem (2001), a estatística de teste passa a ser como apresentado pela Equação  $T^2$  de Hotelling (2.11)

$$T^2 = n(\bar{X} - \bar{\bar{X}}) S^{-1} (\bar{X} - \bar{\bar{X}}), \quad (2.11)$$

onde  $\bar{\bar{X}}$  e  $S^{-1}$  são os estimadores comuns do vetor médio e da matriz de covariância obtida de uma série histórica de dados e  $n$  é o tamanho da amostra ou subgrupos.

Como se vê em na Equação (2.11), o procedimento chamado de Carta de Controle  $T^2$  de Hotelling apresenta-se como um gráfico direcionalmente invariante, ou seja, sua habilidade em detectar uma mudança média no vetor depende apenas da magnitude escalar da mudança e não de sua direção. Umit e Cigdem (2001) e Alt e Grimshaw (2013) observam a importância da seleção dos limites de controle na aplicação do controle estatístico multivariado pela estatística  $T^2$  de Hotelling. Há duas fases distintas na aplicação da Carta de Controle. A Fase I é o uso da Carta de Controle para o estabelecimento do controle, isto é, testa se o processo está ou não sob controle estatístico, quando os  $m$

subgrupos são extraídos e as estatísticas amostrais  $\bar{X}$  e  $S$  são calculadas. O objetivo na Fase I é obter um conjunto de observações sob controle, de modo que se possa estabelecer o limite de controle para a Fase II quando então se fará o monitoramento do processo. A análise na Fase I é também chamada de análise retrospectiva, trata de dados históricos do estudo. O limite de controle da Carta de Controle  $T^2$  de Hotelling na Fase I é descrito por Montgomery (2016) e seu cálculo é feito conforme a Equação (2.12).

$$LSC = \chi_{\alpha,p}^2, \quad (2.12)$$

onde  $\chi^2$  é o quantil  $\alpha$  da distribuição qui-quadrado com  $p$  graus de liberdade.

Quando  $\mu$  e  $\Sigma$  são estimados a partir de um grande número de amostras preliminares costuma-se usar a Equação (2.12) como limite superior de controle em ambas as Fases I e II. (MONTGOMERY, 2016). A análise retrospectiva das amostras preliminares, o teste do controle estatístico e a determinação do limite de controle também ocorrem no ambiente da Carta de Controle univariada. Para a Carta de Controle  $\bar{X}$ , a prática mostra que se forem utilizadas  $m \geq 20$  amostras preliminares os valores dos limites de controle da Fase I e da Fase II praticamente coincidirão, portanto, não havendo necessidade de distinção entre os mesmos. Por outro lado, Jensen et al. (2006) alertam que na Fase I tamanhos amostrais maiores são necessários para a definição do desempenho do processo pelo Número Médio de Amostras Coletadas até um Sinal de que o Processo não sofre Alteração - *Average Run Length* (ARL). Isso permitirá a definição probabilística correta do período amostral esperado para ocorrência dos alarmes devido a causas comuns. Lowry e Montgomery (1995) e Montgomery (2016) enfatizam a necessidade de muitas amostras preliminares para oferecer a precisão necessária aos limites de controle a utilizar na Fase II. Os valores de  $m$  recomendados, segundo Lowry e Montgomery (1995), devem ser superiores a 20 amostras preliminares e na prática, frequentemente são superiores a 50 amostras. Observando que provavelmente esses tamanhos de amostras recomendados sejam pequenos, Jensen et al. (2006) sugerem tamanhos de amostras de pelo menos 200 para conferir robustez na estimação do limite de controle para Fase II. Neste estudo em particular foi utilizado um tamanho de série  $m = 500$  para estimativa de parâmetros.

## 2.6 Síntese do referencial e abordagem proposta

O referencial teórico transcorre sobre elementos históricos e a evolução da auditoria até a auditoria de sistemas de informação e a sua relação com as Tecnologias de Informação, avançando rumo ao tema da lavagem de dinheiro, assunto internacionalmente recorrente desde a década de 1980 e carregando elementos que unem estas temáticas, isto é, abordando as práticas de auditoria em lavagem de dinheiro. O texto transita por apresentar o desenvolvimento dos sistemas de informação baseados na comunicação digital pelo aprimoramento dos recursos da tecnologia da informação através dos avanços da computação e da miniaturização de componentes eletrônicos, consolidando o *e-commerce* e as transações financeiras em caráter internacional através da internet.

Os controles desenvolvidos através dos processos de auditoria e seus aprimoramentos envolveram a auditoria contábil e financeira, migrando para outras áreas das organizações e, especificamente, a auditoria de Sistemas de Informação. Um paralelo pode ser feito em relação às auditorias de sistemas da qualidade, onde audita-se os elementos documentais e registros do sistema de informação. Porém, um sistema auditado e aprovado não garante a qualidade dos produtos. Do mesmo modo, a auditoria de sistemas de informação, de modo genérico, e entre eles a auditoria contábil de sistemas de informação garante que o sistema é confiável, baseado nas evidências apresentadas. Note que, assim como um sistema de gestão da qualidade não garante a qualidade do produto, uma auditoria no sistema de informação, a princípio, não deve ser capaz de avaliar a fidedignidade dos documentos, mas sim se estes estão de acordo com as diretrizes do sistema auditado.

Os processos de auditoria desenvolvidos, bem como os sistemas de controle, a exemplo da segurança da informação, tendem a assegurar que a transferência de dados ocorra dentro dos princípios estabelecidos de disponibilidade, integridade, confidencialidade, não repúdio e a autenticidade, ou seja, comunicação entre o emissor A para o receptor B, não ocorrendo o julgamento sobre o dado transferido com relação à sua origem e idoneidade. Assim, a abordagem proposta visa a auxiliar os controles das instituições financeiras na atividade de combate à lavagem de dinheiro através de ferramentas de monitoramento de processos, complementando as atividades de auditoria dos sistemas de informação.

Estudos na área de lavagem de dinheiro apontaram a aderência de dados de mo-

vimentação financeira sem indícios de suspeitas de fraude à distribuição dos primeiros dígitos, conhecida como Lei de Benford. Assim, a ação de disponibilizar um recurso capaz de sinalizar movimentações financeiras suspeitas, utilizando a estatística com suporte da aplicação da Lei de Benford em uma Carta de Controle multivariada complementa a atividade de auditoria ao apontar fontes de evidências suspeitas de serem fraudulentas.

O Controle Estatístico do Processo desenvolvido na década de 1920 para a manufatura de produtos foi viabilizado através de Cartas de Controle (*Control Charts*) criadas por Shewhart (1931) e teve seu espectro de aplicação ampliado para além do meio industrial. Na identificação de tendências em novas pesquisas, Woodall e Montgomery (2014) citam três principais categorias onde há inovações na aplicação do Controle Estatístico do Processo: monitoramento da saúde, vigilância da saúde pública e vigilância de síndromes. Frisén (2008) trata em seu livro sobre métodos de vigilância estatística aplicada na análise financeira. A performance das Cartas de Controle é usualmente avaliada na literatura através de medidas de tempo, Média de Tempo Padrão - *Time Average Standard* (TAS) ou de Número Médio de Amostras coletadas até a detecção de sinais - *Average Range Length* (ARL). Estes sinais, caso não sejam consequência de alterações no processo são denominados de alarmes falsos, mas se os sinais são consequências de alguma alteração no processo e a Carta de Controle não detectar considera-se um erro de detecção. Geralmente estas avaliações são feitas por simulação e estatísticas são calculadas a partir das quais se avalia a capacidade da Carta de Controle em detectar corretamente quando o processo saiu ou não de controle. O ARL, ou número médio de amostras coletadas até um sinal quando o processo não sofre alteração é denominado de  $ARL_0$ , assim como quando o processo sofre uma alteração o número médio de amostras coletadas necessárias até que a Carta de Controle sinalize é denominado de  $ARL_1$  e depende de tamanho do desvio ao qual o processo foi submetido.

Em resumo, a literatura sugere a necessidade de um recurso de vigilância estatística a ser aplicada na análise financeira. Em paralelo, pode-se estabelecer que divergências em relação à Lei de Benford são indicativos de suspeita de dados fraudulentos. Para unir estas duas condições, pode-se utilizar uma Carta de Controle de processo multivariada como recurso de suporte à auditoria de sistemas de informação financeira para o monitoramento de lançamentos financeiros.

No recurso proposto nesta pesquisa o operador avaliará a frequência de movimen-

tações financeiras de seu cliente para identificar o padrão de suas movimentações. A partir de então, implantará o recurso de acordo com a frequência esperada dessas movimentações. Por exemplo, um estudo das movimentações financeiras de uma empresa de grande porte, em função do volume de lançamentos financeiros, muito provavelmente deverá ser monitorado de forma distinta aos lançamentos financeiros de uma pessoa física. Isto porque a estimativa de parâmetros necessita de um certo volume de dados. Do mesmo modo, os limites de controle devem ser adequados à velocidade de detecção desejada e à taxa de alarmes falsos decorrentes da velocidade de detecção implementada. Por exemplo, um  $ARL_0 = 20$  prevê a ocorrência de um alarme falso em uma frequência esperada de 20 eventos, que podem ser semanais, mensais, anuais ou períodos definidos pelo operador em relação à frequência das movimentações financeiras a monitorar. Note que em relação a um período semanal, 20 semanas será o tempo decorrido entre eventos onde é esperado um alarme falso, ou seja, espera-se que transcorram 5 meses até um provável sinal falso. Os períodos serão escolhidos pelo operador, em princípio, conforme o padrão das movimentações financeiras e de acordo como o ARL determina, pela intenção de investigação de um alarme em maior ou menor período de tempo, ou então, regulado pela disponibilidade do operador em investigar um sinal de suspeita nas movimentações financeiras. Os capítulos 3 e 4 detalham o desenvolvimento da pesquisa, o procedimento proposto e os resultados da validação do mesmo.



### 3 PROCEDIMENTOS METODOLÓGICOS

Acerca da pesquisa de modelagem como a realizada no presente estudo, Hevner et al. (2004) definem que esta caracteriza-se por manipular diretamente as variáveis relacionadas ao objeto de estudo através de técnicas matemáticas e tecnologia para a solução do modelo matemático. Para Miguel e Fleury (2012), a orientação do pesquisador pode ser racional, dedutiva axiomática, baseada em reconstrução artificial da realidade para visualizar as conclusões e decisões sugeridas pelo modelo, essa abordagem é importante nos casos de pesquisa axiomática que desenvolvem métodos novos e mais efetivos. Dentre os tipos de pesquisa com o viés de modelagem está a simulação computacional, Miguel e Fleury (2012) destacam que os avanços teóricos em matemática e estatística aliados às novas tecnologias em ciências da computação permitem a manipulação das variáveis, proporcionando o estudo da relação entre as causas e os efeitos de determinado fenômeno. Segundo Prodanov e Freitas (2013) e Miguel e Fleury (2012), através da criação de situações de controle evita-se interferências de variáveis intervenientes. Interfere-se diretamente nos dados, manipulando a variável independente e observa-se o que ocorre na variável dependente. Frequentemente as avaliações das conclusões inferidas mostram necessidade de revisão da modelagem, o que implica na repetição do ciclo tantas vezes quanto necessárias à adequação do modelo matemático.

De acordo com Marconi e Lakatos (2010) esta pesquisa é suportada por uma abordagem quantitativa com objetivo exploratório de natureza aplicada. Miguel e Fleury (2012) definem que a pesquisa axiomática quantitativa produz conhecimento sobre o comportamento de certas variáveis do modelo, baseada em premissas do comportamento de outras variáveis desse mesmo modelo. Quanto à natureza dos métodos de trabalho, segundo Miguel e Fleury (2012) esta é uma pesquisa de modelagem e simulação de Monte Carlo de dados discretos, ou seja, representam sistemas em que as mudanças ocorrem em pontos discretos do tempo. As seguintes etapas compõem o método de trabalho e cada etapa do método de trabalho é detalhado a seguir.

- Levantamento bibliográfico;
- Modelagem e simulação;
  - Definição do modelo conceitual;
  - Implementação do modelo em *software*;
  - Validação do modelo; e

- Simulação do modelo;
- Aplicação de métodos e análise dos resultados; e
- Aplicação em caso real.

Uma vez definido o tema e objetivos, buscou-se o referencial teórico de suporte à pesquisa. Nesta fase, foram levantadas informações sobre lavagem de dinheiro, sistemas de auditoria e ferramentas de monitoramento e de controle estatístico de processos. Estes temas foram organizados em dois grupos, a saber: i) temas de suporte teórico às discussões executadas nesta pesquisa e ii) temas de suporte metodológico, tais como ferramentas e formulações.

O levantamento do referencial teórico orientou-se na busca de artigos que tratam de disciplinas relacionadas com as áreas de Negócios e Economia, Engenharia e Estatística, nas bases de dados: *EBSCO*, *Emerald Insigth*, *Taylor & Francis*, *Scopus (Elsevier)*, *Google Scholar*, *Science Direct*, *Home Springer* e *ProQuest*, na identificação de artigos científicos relacionados aos temas associados às palavras-chave. As palavras-chave principais utilizadas foram: *Money Laundering*, *Anti-Money Laundering*, *Fraud Prevention*, *Systems Information*, *Systems Information Auditing* e *Benford's Law*, combinadas com *Financial Fraud*, *Financial Crime*, *Internal Controls*, *Internal Audit*, *Systems Information*, *Systems Information auditing*, *Computer*, *Computer Fraud*, *Computer Crime*, *Cybercrime*, *Audit Systems*, *Auditing*, *Money Laundering*, *Fraud*, *Statistical Process Control* e *Control Chart*, incluindo versões equivalentes em língua portuguesa, no período de 2012 a 2016.

Considerando a modelagem e simulação dos dados, Miguel e Fleury (2012) destacam que em engenharia de produção a gestão de sistemas de produção bens e serviços tem a pesquisa quantitativa baseada em modelos matemáticos e na simulação. A utilização de modelos permite compreender e identificar problemas e suas oportunidades no ambiente em estudo, com o objetivo de dar tratamento adequado e de forma sistemática. Processos operacionais reais podem se apresentar como sistemas complexos e difíceis de serem modelados cientificamente no que se relaciona com a reprodução e desempenho. Bertrand e Fransoo (2002) afirmam que em análises de problemas idealizados pode-se gerar conhecimento valioso relativo aos problemas reais. Um modelo caracteriza-se por elementos essenciais, concretos ou abstratos, devem ser suficientemente detalhados para representar um sistema real, em contrapartida, deve ser suficientemente simples, a fim de facilitar seu tratamento por métodos de análise e de resolução

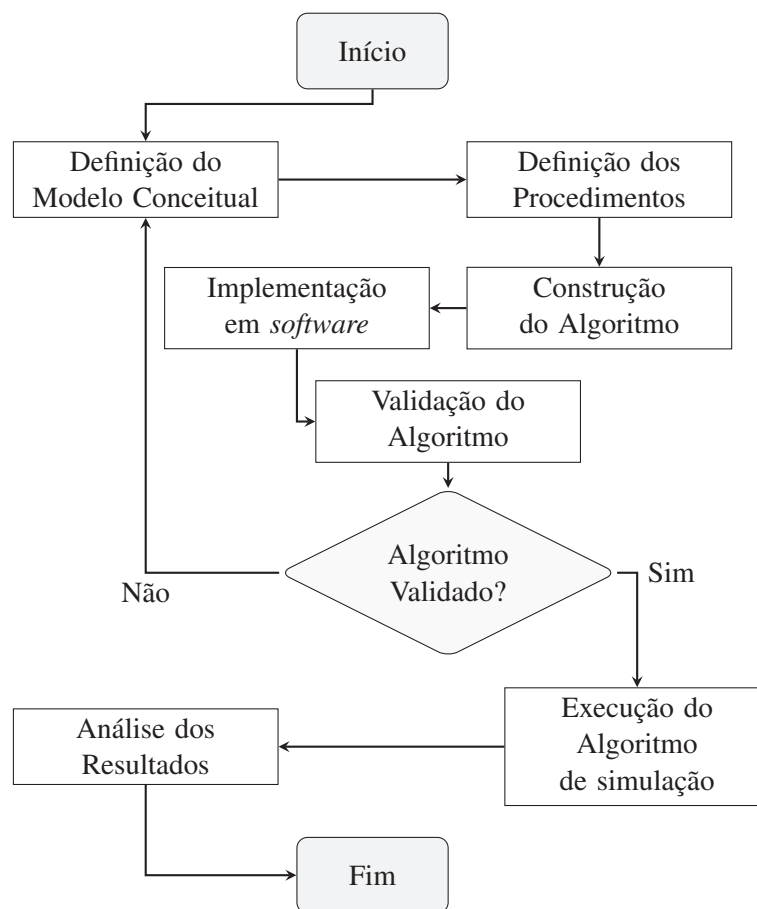
conhecidos. (MIGUEL; FLEURY, 2012). Os avanços teóricos da simulação estocástica vistos em Kelton e Law (2000) e da matemática, aliados à evolução da tecnologia de informação em *software* e *hardware* permitiram o aumento da complexidade da formulação de modelos, consolidando a pesquisa operacional como recurso de apoio ao processo de decisão.

Para atingir o objetivo de monitorar suspeitas de indicativo de fraude por lavagem de dinheiro baseado em Cartas de Controle de processos, foi definido inicialmente um modelo conceitual com base no tipo de distribuição de probabilidades que representam os dados oriundos de lançamentos financeiros. Suposição de dados que se comportam conforme a distribuição da Lei de Benford e também a análise de uma função de distribuição cumulativa empírica, utilizando estimadores a partir dos dados financeiros reais coletados. A avaliação de desempenho do modelo de Carta de Controle escolhido para esta aplicação segue o fluxograma apresentando na Figura 3.

A implementação do modelo em *software* consistiu em duas etapas: a primeira, com a geração de dados aleatórios de acordo com as características identificadas na definição do modelo conceitual na segunda, a aplicação da formulação de monitoramento, baseada nas Cartas de Controle definidas.

De posse das sequências de dados, pode-se avaliar a qualidade destas informações e verificar sua aderência à Lei de Benford, validando desta forma o modelo de geração dos dados. Tabelas de distribuição de frequências dos primeiros dígitos significativos foram construídas para cada série de dados. Os resultados foram comparados, a fim de estabelecer as variações de frequências observadas em cada dígito. Estas informações foram posteriormente utilizadas para inserir ruído na geração de dados simulados. Foi também avaliada a taxa de ocorrência de alarmes falsos, a fim de validar os limites de controle determinados para o monitoramento. Essa sequência baseia-se no modelo apresentado por Korzenowski et al. (2017).

Ainda na fase de simulação, foram geradas perturbações adicionais a fim de caracterizar alterações na distribuição dos dígitos significativos. Esta perturbação foi necessária para avaliar a capacidade de detecção das ferramentas de monitoramento adaptadas para este estudo. Os ruídos foram inseridos com base em uma distribuição Normal multivariada, com vetor de médias zero e matriz de covariâncias similar à observada, conforme apurado nas distribuições de frequências realizadas na análise inicial. Estas simulações foram utilizadas na aplicação dos métodos e permitiram a análise dos resultados.



**Figura 3:** Fluxograma de procedimentos da Simulação

O procedimento de inserção de ruídos para avaliação da performance das ferramentas de controle foi incorporado ao algoritmo de simulação. O procedimento de simulação foi elaborado conforme o Algoritmo 1.

Para a viabilidade de construção do algoritmo, adotou-se simplificações no modelo de geração dos dados em função das características do processo de simulação. Foram desenvolvidas rotinas computacionais no formato de funções para a distribuição de Benford no *software* R. O código destas rotinas estão disponíveis no Apêndice A e elas foram utilizadas para a geração de dados aleatórios com este comportamento. O Algoritmo 1 representa a implementação na Linguagem R-Project (TEAM, 2016) utilizada na geração e análise de dados empregados nesta dissertação.

Para garantir a reprodutibilidade dos resultados, foi fixada a semente aleatória do

---

**Algoritmo 1:** Para análise dos resultados obtidos por simulação
 

---

**Result:** Matriz com os resultados do ARL por desvio  $\delta$  adicionado à média de cada série gerada

```

1 begin
2   Fixar a semente aleatória
3   Definir  $m$ ;                               Número de alarmes
4   Definir vetor de desvios  $\delta$ ;
5   Define  $w$ ;                                   Tamanho da série na Fase I
6   Define o vetor  $ARL$ ;   Vetor para armazenar o  $ARL$  para os  $m$  alarmes por
   cada  $\delta$ 
7   for  $i = \delta$  do
8     Define o vetor  $RL$ ; Vetor para armazenar os tamanhos das corridas dos  $m$ 
     alarmes
9     for  $i = 1:m$  do
10      Gera  $w$  vetores com dados da distribuição de Benford;
11       $k = w$ ;
12      Define o vetor  $Z$  das estatísticas de monitoramento; Calcula a
      estatística de monitoramento  $\chi^2$  ou  $T^2$  e salva no vetor  $Z$ ;
13      Define  $UCL$ ;
14       $Z[end] = 0$ ;                               Última posição do vetor  $Z$  recebe zero
15      while  $Z[end] \leq UCL$  do
16         $k = k + 1$ ;
17        Gera um subgrupo baseado na distribuição de Benford,
        adicionado de um desvio  $\delta$  à média da série;
18         $Z[end + 1] =$  Estatística de monitoramento do subgrupo no tempo
         $k$ ;
19      end
20       $RL[end + 1] = k - w$ ;                        $RL$  até o sinal
21    end
22     $ARL[end + 1] = mean(RL)$ ;
23  end
24  Return Matriz do  $ARL$  para cada valor de  $\delta$  adicionado à média;
25  Os valores do desvio  $\delta$  são dados por proporções do desvio-padrão  $\sigma$ 
  de cada série, calculados na ordem de 0, 0; 0, 5; 1, 0; 1, 5; 2, 0; 3, 0 e 5, 0,
  adicionados à média de cada série gerada.
26 end

```

---

gerador de números aleatórios ( $set.seed = 42$ ). O número de alarmes para o cálculo do  $ARL$  foi definido em  $m = 50.000$  e o vetor de desvios da distribuição esperada  $\delta \in$

$\{0, 0; 0, 5; 1, 0; 1, 5; 2, 0; 3, 0; 5, 0\}$  em proporções do desvio-padrão  $\sigma$  foi assim definido em função dos desvios usuais utilizados na literatura da área. Por fim, o tamanho da série na Fase I para estimar a variabilidade em cada dígito, bem como dos valores de proporções médias dos primeiros dígitos, foram definidos em  $w = 500$ . Note-se que no algoritmo apresentado no Anexo A.1 a geração de dados na Fase I utilizou a semente aleatória (*set.seed* = 2016) e para a avaliação do ARL ela foi novamente fixada em 42. O restante do algoritmo é autoexplicativo: um novo conjunto de dados é gerado com ou sem desvio, a estatística de monitoramento é determinada e confrontada com os limites de controle, linha 15. Caso não ocorra um alarme, um contador é atualizado ( $k$ ) e o processo reinicia. Quando um alarme é identificado, o contador  $k$  é salvo na posição final do vetor  $RL$ . Ao final do processo, o vetor  $RL$  terá  $m = 50.000$  corridas e então, o valor médio deste vetor poderá ser determinado obtendo o  $ARL$ . Este procedimento repete-se para todos os desvios  $\delta$  definidos, retornando uma matriz com os resultados.

Uma vez gerados dados de controle, dados perturbados e as técnicas estatísticas de monitoramento terem sido implementadas, seus resultados puderam ser avaliados. Foi utilizado o teste  $\chi^2$  para a verificação de aderência.

O teste  $\chi^2$  de Karl Pearson é um teste de hipóteses que serve para avaliar a dispersão entre duas variáveis nominais qualitativas. (PLACKETT, 1983). Umit e Cigdem (2001) apresentam a estatística  $\chi^2$  como meio de avaliar a aderência de dados à distribuições de probabilidade. Os autores afirmam que se os verdadeiros parâmetros de uma distribuição de probabilidades são conhecidos, a distribuição  $\chi^2$  deve ser utilizada para avaliar a aderência de um conjunto de dados a esta distribuição. Por exemplo, presumindo que  $x_1$  e  $x_2$  são as variáveis de um processo com distribuição normal e  $\mu_1$  e  $\mu_2$  são as médias populacionais dessas variáveis, tem-se que  $\bar{X}_1$  e  $\bar{X}_2$  são as médias das amostras, supondo conhecer os desvios padrão ( $\sigma_1$  e  $\sigma_2$ ) e a covariância ( $\sigma_{12}$ ). Assim, a estatística  $\chi^2$  apresentada pelos autores é reproduzida na Equação (3.1).

$$\chi^2 = \frac{n}{\sigma_1^2 \cdot \sigma_2^2 - \sigma_{12}^2} \left[ \sigma_2^2(x_1 - \mu_1)^2 + \sigma_1^2(x_2 - \mu_2)^2 - 2\sigma_{12}(x_1 - \mu_1)(x_2 - \mu_2) \right], \quad (3.1)$$

Neste caso, a estatística  $\chi^2$  segue uma distribuição  $\chi^2$  com 2 graus de liberdade. Silva, Korzenowski e Vaccaro (2014) definem em seu estudo de suspeita de fraude por lavagem de dinheiro a importância do teste de aderência dos dados à distribuição de

Benford. Mencionam que há os erros sistemáticos e erros de aleatoriedade. Os erros sistemáticos estão relacionados à incerteza de medição, erro de digitação, de leitura ou de coleta de dados. Os erros aleatórios são comuns a qualquer evento natural, regidos pela probabilidade inerente ao fenômeno em estudo. Segundo os autores, esses erros podem levar a decisões errôneas. Silva, Korzenowski e Vaccaro (2014) apresentam ainda o teste da estatística  $Q$ , proposto por Karl Pearson em Plackett (1983) usado para avaliação de dados contábeis baseado na estatística  $\chi^2$ . A formulação da estatística aplicada pelos autores é reproduzida na Equação (3.2).

$$Q_k = \sum_{j=1}^{k+1} \frac{(N_j - np_j)^2}{np_j}, \quad (3.2)$$

onde  $k$  é o número de graus de liberdade da distribuição  $\chi^2$ ,  $N_j$  é o número de observações do valor  $j$ ,  $n$  é o tamanho da amostra,  $p_j$  é a proporção de valores  $j$  esperados segundo a distribuição,  $np_j$  portanto é o número de observações esperadas em  $j$ , supondo  $p$  e  $j = 1, 2, 3, \dots, 9$ , correspondente aos dígitos significativos de Benford.

O poder de confiança define a capacidade da análise de aderência a uma distribuição, para um efeito fixo, com nível de significância fixo em 5%, percebe-se que somente é possível aumentar o poder de um teste se aumentarmos o tamanho da amostra. Do mesmo modo, se é desejado identificar diferenças menores, isto é, deseja-se identificar tamanhos de efeito menores como significativos, amostras maiores são necessárias.

Para determinar o tamanho da amostra para a execução de um teste qui-quadrado sobre uma tabela de contingência com 8 graus de liberdade (a Lei de Benford refere-se a 9 dígitos significativos) deve-se definir o tamanho do efeito que deseja-se identificar. *Softwares* podem ajudar na tarefa de determinar o tamanho do efeito, mas, segundo os autores, cabe aos especialistas na área definir o limite a ser detectado. (SILVA; KORZENOWSKI; VACCARO, 2014).

O método de monitoramento escolhido foi a Carta de Controle  $\chi^2$ , quando assumiu-se como verdadeira a distribuição de Benford (parâmetros conhecidos). Quando optou-se pela avaliação de dados anteriores com parâmetros estimados na Fase I (parâmetros desconhecidos) utilizou-se a Carta de Controle  $T^2$  de Hotelling. A formulação, bem como as devidas adaptações necessárias à implantação no contexto explorado neste trabalho são detalhadas no Capítulo 4. Para avaliação dos procedimentos de monitoramento *off-line* foram computadas as taxas de erros do tipo I e II, considerando os níveis

de significância usualmente adotados em monitoramento de processos produtivos vistos em Montgomery (2016), bem como o poder estatístico esperado/observado. Foram avaliados outros níveis para o erro do tipo I, em função do tipo de aplicação para a qual a ferramenta poderá ser utilizada, conforme sugerido por Silva, Korzenowski e Vaccaro (2014), alterando assim a observância de alarmes falsos de acordo com o volume de dados observados e definindo limites de controle para estas situações específicas.

Os resultados são apresentados em tabelas descritivas, onde as taxas de erro são convertidas em ARLs. Além dos valores médios são apresentados os desvios padrão, o tempo dispendido na simulação e análise dos dados pelo algoritmo computacional desenvolvido para esta pesquisa. Por simulação também foram definidos limites de controle a serem utilizados em situações onde deseja-se um  $ARL_0$  específico.

Após validados os procedimentos, foram coletados dados reais de movimentação financeira de prestação de contas do pleito municipal de 2016 junto aos Tribunais Regionais Eleitorais, por intermédio dos dados do Tribunal Superior Eleitoral do Brasil. TSE (2016a). Um segundo exemplo de dados coletados trata-se de uma planilha de movimentação de contas resultante de auditoria em uma instituição financeira. A Carta de Controle para identificação de suspeitas de fraude por lavagem de dinheiro foi aplicada nestes dados e os resultados são exibidos em 4.2.2, no intuito de elucidar o funcionamento do método.

No primeiro exemplo, foram utilizados dados da prestação de contas da campanha a prefeituras no Brasil, em 2016. Os dados selecionados referem-se à divulgação da prestação de contas dos candidatos a prefeito das capitais dos estados que representam os cinco maiores colégios eleitorais do país. Segundo dados do Tribunal Superior Eleitoral TSE (2016b), a relação do primeiro ao quinto lugar são os estados de São Paulo, Minas Gerais, Rio de Janeiro, Bahia e Rio Grande do Sul, suas capitais em mesma sequência são São Paulo, Belo Horizonte, Rio de Janeiro, Salvador e Porto Alegre. Os dados das prestações de contas foram coletados na página do TSE (2016a), onde se tem acesso às regiões do país onde seleciona-se o estado, na sequência a capital do estado, em seguida os candidatos e, por fim, se tem acesso às contas de cada candidato, após, cada tipo de conta e demais dados estão acessíveis a abertura e a visualização. Dentre todas informações constantes, foram selecionadas as contas dos candidatos que ficaram em primeiro e em segundo lugar, independentemente da ocorrência de candidato vencedor no primeiro ou no segundo turno. Para o processamento foram acessados os arquivos



em Excel.csv com as movimentações das contas das receitas e das despesas. Os dados das duas planilhas de cada candidato foram consolidadas em apenas uma única planilha, foram agrupados os valores movimentados de receitas e despesas em uma única coluna, em outras duas colunas constam a data da movimentação e o número da semana do ano. Estas planilhas foram nomeadas para cada candidato em formato Excel.csv. O processamento se dá a partir da leitura do algoritmo no *software R*.

Estabelecida a distribuição de Benford é feita a leitura da base de dados no diretório onde as planilhas foram armazenadas. O intervalo de tempo é definido pela identificação prévia do número da semana nas linhas dos lançamentos financeiros, em planilha individual de cada candidato. Em seguida, define-se o período inicial e final em semanas para o intervalo selecionado. São extraídos os primeiros dígitos significativos dos dados dos lançamentos financeiros por cada semana. São calculadas as proporções dos primeiros dígitos. Na sequência, calcula-se o qui-quadrado das proporções dos primeiros dígitos para cada intervalo. Uma vez os dados padronizados ocorre o armazenamento dos valores qui-quadrado de cada semana.

Na sequência, toma-se o score do ARL padronizado da distribuição simulada para limite de controle de acordo com o número de ocorrências (semanas) de cada conta, e são tabulados os pontos na Carta de Controle. Esse processo se repete nos dados armazenados para todos candidatos, em que se tratando de dois para cada uma das cinco capitais totalizam dez.

No segundo exemplo, tomou-se dados de uma instituição financeira em uma planilha em Excel.csv, onde constam dados de cinco meses em 12.766 linhas de lançamentos financeiros. O processamento se dá a partir da leitura do algoritmo no *software R*. Estabelecida a distribuição de Benford é feita a leitura na base de dados, realizando a seleção das linhas válidas dos lançamentos. A seguir é feita a identificação das datas de início e término do período em análise, atribuição do tamanho do intervalo móvel  $i$  de análise, criação de uma matriz para acumular as proporções dos dígitos significativos, cálculo das proporções dos dígitos significativos e a definição dos pontos para plotar na Carta de Controle. É calculado o limite de controle com 0,9973 de probabilidade, então é plotada a Carta de Controle com as estatísticas nos intervalos móveis definidos para análise. Para definição da distribuição de Benford e o limite de controle foi utilizada a função de distribuição acumulada empírica, *Empirical cumulative distribution function* (ecdf), segundo Maechler (2008).

## 4 APRESENTAÇÃO, VALIDAÇÃO E APLICAÇÃO DO MODELO DE MONITORAMENTO PROPOSTO

### 4.1 Estrutura do modelo de monitoramento proposto para avaliação de suspeitas de fraude por lavagem de dinheiro

Como visto na seção 2.6, a abordagem propõe uso da Estatística, aplicando a Lei de Benford em uma Carta de Controle de processo multivariada como recurso de vigilância estatística para auxiliar as instituições financeiras no processo de monitoramento de movimentações financeiras de seus clientes. O processo se dá com o monitoramento das movimentações financeiras com a finalidade de determinar seu padrão e identificar sinais que, analisados, fornecerão informações sobre a manutenção ou não desse padrão para tomada de decisão sobre possíveis suspeitas. Os passos do processo de monitoramento proposto são descritos a seguir:

1. O histórico da conta do cliente fornece dados para o cálculo das estatísticas descritivas das suas movimentações financeiras. Esta etapa fornece as médias, desvio-padrão, variância e covariância das frequências dos primeiros dígitos observados no conjunto de dados históricos;
2. A determinação do limite de controle se dá pela padronização dos dados a partir das estatísticas descritivas obtidas na etapa anterior. Esta etapa fornece o limite de controle de acordo com a Carta de Controle de monitoramento dos dados adotada  $\chi^2$  ou  $T^2$ ;
3. Monitora-se movimentações financeiras em um intervalo de tempo definido e calcula-se as estatísticas descritivas do conjunto de dados nesse intervalo de tempo. Nesta etapa, obtém-se o vetor de distribuição dos primeiros dígitos que servirá para comparação com o padrão histórico das movimentações do cliente;
4. A estatística de monitoramento se utiliza do vetor de distribuição dos primeiros dígitos da etapa 3 para o cálculo da estatística de monitoramento multivariada. Nesta etapa, de acordo com a Carta de Controle de monitoramento  $\chi^2$  ou  $T^2$ , obtém-se a estatística de monitoramento multivariada  $\chi^2$  ou  $T^2$  para verificação de aderência ao padrão dos dados;
5. A decisão de suspeição sobre as movimentações financeiras se verifica pela análise da aderência dos dados encontrados com o padrão histórico. A estatística

multivariada e o limite de controle, definem se os dados monitorados no período aderem ao padrão histórico (o limite de controle pode estar pré-definido pela frequência de sinais esperada em um valor do  $ARL_0$  padronizado, correspondente ao período de monitoramento). Se os dados estão de acordo com o padrão histórico atribui-se valor = 0 (zero), não há soma de um alarme e o monitoramento prossegue, ou seja, o processo está sob controle, não há pontos fora de controle na Carta de Controle, não apresentando indícios de mudança no padrão das movimentações financeiras que requeiram análise mais detalhada nos dados. Por outro lado, se os dados não estão de acordo com o padrão histórico atribui-se valor = 1 (um), soma-se um alarme, isto sinaliza que ocorreu um ponto fora de controle, portanto, os dados da Carta de Controle indicam que o processo está fora de controle. Isso requer uma análise nas movimentações financeiras do cliente para verificar se ocorreu um alarme falso ou, em caso contrário, o ponto fora de controle sinaliza a uma mudança real no padrão e configura suspeita de fraude, exigindo uma análise mais detalhada nas movimentações financeiras.

Sintetizando as etapas e passos seguidos no processo de monitoramento proposto, o Quadro 1 descreve as atividades, entradas, procedimentos e saídas esperadas em cada etapa.

## 4.2 Apresentação dos resultados

### 4.2.1 Resultados da validação do modelo por simulação

Para realizar as Simulações foram utilizados como recurso o *software* livre *R* disponibilizado por Team (2016) e de *hardware* dois computadores portáteis. A configuração dos computadores utilizados é listada a seguir.

- *Notebook* Dell - Modelo Inspiron 1545; Processador Pentium(R) Dual-Core CPU 2.00 GHz; Memória (RAM) 4,00 GB; S.O. Windows 32 Bits.
- *Notebook* Dell - Modelo Inspiron N4050; Processador Intel(R) Core (TM) i3-2350 M CPU 2.30 GHz; Memória (RAM) 4,00 GB; S.O. Windows 64 Bits.

Para a validação do procedimento de monitoramento adotou-se duas estratégias de análise. A primeira considerou a distribuição dos primeiros dígitos da movimentação

**Quadro 1:** Procedimento de execução do monitoramento de movimentações financeiras

Etapa	Atividade	Entrada	Procedimento	Saida
1	Analisar o histórico das movimentações da conta do cliente	Dados de movimentações financeiras	Cálculo de estatísticas descritivas das movimentações financeiras	Medidas de tendência central e de variabilidade (frequências dos primeiros dígitos observados)
2	Determinar o limite de controle	Estatísticas descritivas obtidas na Etapa 1	Cálculo do limite de controle com base na Carta de Controle de monitoramento adotada	Limite de controle
3	Obter a estatística descritiva com base em dados das movimentações financeiras do cliente num intervalo de tempo definido	Dados de movimentações financeiras	Cálculo de estatísticas descritivas das movimentações financeiras	Vetor de distribuição dos primeiros dígitos
4	Calcular a estatística de monitoramento	Vetor de distribuição dos primeiros dígitos obtido na Etapa 3	Cálculo da estatística de monitoramento com base na Carta de Controle de monitoramento adotada	Estatística de monitoramento multivariada
5	Verificar a aderência dos dados das movimentações financeiras do cliente no intervalo de tempo definido ao histórico das movimentações	Estatística de monitoramento multivariada e o limite de controle obtido na Etapa 2	Comparar a estatística com o limite de controle para a tomada de decisão quanto à aderência dos dados	Dados encontrados: De acordo com o padrão histórico = 0. Caso contrário = 1

Fonte: Elaborado pelo Autor.

financeira como sendo equivalente à distribuição de Benford. Neste caso, a estatística do teste é computada comparando o vetor de primeiros dígitos observados com o valor de referência dado pela distribuição de Benford, configurando um ambiente de monitoramento com parâmetros conhecidos, o que caracteriza a Carta de Controle  $\chi^2$ . A segunda, geração de um conjunto de dados históricos pela simulação de uma matriz de dados segundo a distribuição de Benford com  $m = 500$  linhas (amostras) e  $p = 9$  colunas que representam as frequências dos primeiros dígitos significativos.

Com base nesta matriz foram calculadas as estimativas da média e da variabilidade para a proporção de cada dígito significativo. Ao final, este vetor de proporções médias foi normalizado para atender à restrição de que a soma das  $p_i$  proporções de cada dígito ( $i \in \{1, 2, 3, \dots, 9\}$ ) é igual a um. Este ambiente de monitoramento com parâmetros desconhecidos e estimados a partir de um conjunto de dados histórico caracteriza a Carta de Controle  $T^2$  de Hotelling. Assim, inicialmente serão apresentados os resultados de ARL para a Carta de Controle  $\chi^2$  e na sequência para a Carta de Controle  $T^2$  de Hotelling.

Para a confirmação da aderência dos dados simulados à distribuição, foram gerados lançamentos para obtenção de primeiros dígitos significativos. Ao número de lançamentos *NumLanc* foi atribuído um comportamento normalmente distribuído, com média 200 e desvio-padrão igual a 15 e, posteriormente, foram gerados *NumLanc* dados a partir da distribuição de Benford. A proporção de cada dígito significativo foi então obtida e utilizada para o cálculo da estatística de monitoramento  $\chi^2$ . Após, esta estatística de monitoramento foi confrontada com o limite de controle definido para a identificação do *status* de controle estatístico do processo. Este procedimento repetiu-se até que fosse obtido um alarme, quando então foi registrado o número de observações  $k$  até a ocorrência deste alarme. O processo repetiu-se até que fossem observados 50.000 alarmes. Os desvios  $\delta$  correspondem às proporções de 0; 0,5; 1,0; 1,5; 2,0; 3,0 e 5,0 desvios-padrão  $\sigma$ , adicionados à média em cada série de 50.000 alarmes gerados.

Os resultados da análise do monitoramento para a Carta de Controle  $\chi^2$  são apresentados na Tabela 4.

Com base na matriz de dados histórica gerada, obteve-se a estatística  $T^2$  de Hotelling, com distribuição normal, a partir da qual incrementou-se desvios  $\delta$  adicionados à média na ordem de 0; 0,5; 1,0; 1,5; 2,0; 3,0 e 5,0 desvios-padrão  $\sigma$  para o cálculo do ARL correspondente a cada desvio  $\delta$ . Os resultados dos dados simulados são apresen-

**Tabela 4:** Avaliação da aplicação da Carta de Controle  $\chi^2$  - Distribuição dos primeiros dígitos (dBenford)

$\delta$	<i>ARL</i>	$\sigma$ RL	<i>k</i>	Alarmes	Tempo (s)
0,0	141,00	140,2	7.049.946	50000	309.370,8
0,5	76,44	76,3	3.821.896	50000	99.975,6
1,0	33,64	33,2	1.681.955	50000	22.512,5
1,5	14,91	14,3	745.676	50000	4.336,0
2,0	7,21	6,7	360.263	50000	981,7
3,0	2,41	1,8	120.444	50000	172,4
5,0	1,06	0,2	52.893	50000	61,1

Fonte: Dados da simulação

tados na Tabela 5. Para cada desvio  $\delta$  o ARL, o desvio-padrão do RL, o número de vetores médios  $v$  gerados por meio da simulação até a obtenção de 50.000 alarmes e o tempo dispendido pela simulação são apresentados.

**Tabela 5:** Avaliação da aplicação da Carta de Controle  $T^2$  - Distribuição das médias da estatística (eBenford)

$\delta$	<i>ARL</i>	$\sigma$ RL	$v$	Alarmes	Tempo (s)
0,0	141,13	140,4	7.056.499	50.000	164.458,3
0,5	77,59	77,4	3.879.606	50.000	53.757,0
1,0	34,32	33,9	1.715.858	50.000	13.262,3
1,5	15,34	14,8	767.134	50.000	2.797,7
2,0	7,32	6,7	366.140	50.000	580,3
3,0	2,43	1,9	121.457	50.000	91,3
5,0	1,06	0,3	53.128	50.000	32,4

Fonte: Dados da simulação

Os RLs são as medidas de performance para o monitoramento do número de estatísticas até um alarme falso. Destaca-se que os valores de ARL gerados diminuem à medida que o desvio  $\delta$  é incrementado. Os incrementos no desvio  $\delta$  tem a finalidade de aproximar o limite de controle dos vetores médios. A redução dos ARLs representa a possibilidade de ocorrer um alarme em um período de tempo menor, ou seja, um número de estatísticas menor ocorrerá para um alarme se manifestar. A diferença significativa entre os dados gerados, apresentados nas duas tabelas, está nos tempos de processamento que são maiores para a estatística  $\chi^2$  e menores para a estatística

$T^2$ . Comparando-se os dados gerados para as duas Carta de Controle nas Tabelas 4 e 5, verifica-se que os valores encontrados de ARL para cada desvio  $\delta$  adicionados à média das séries, bem como os desvios padrão  $\sigma$  do RL, resultaram em valores muito próximos. Isso determina que as duas estatísticas concordam entre si, confirmando a aplicabilidade do método desenvolvido na simulação.

A partir desta constatação foram geradas novas simulações com desvio  $\delta = 0$  (zero) adicionado aos vetores médios para determinação do limite de controle padronizado para cada  $ARL_0$  específico, assumindo-se os valores 10; 20; 30; 40; 50; 75; 100 e 141 como referência. Os  $ARL_0$  e os LSC padrão servem para utilização no processamento de dados reais. Uma vez que os dados gerados também serão padronizados, o período de expectativa de um alarme falso pelo valor do ARL é correspondente ao número de estatísticas geradas. Os valores encontrados para esses ARLs são apresentados na Tabela 6.

**Tabela 6:** ARLs e LSCs estimados

$ARL_0$	LSC
141	23,57
100	22,45
75	21,51
50	20,17
40	19,43
30	18,48
20	17,07
10	14,62

Fonte: Dados da simulação

Os limites de controle  $ARL_0$  foram obtidos por simulação, uma vez que as probabilidades associadas aos limites observados não concordavam com as taxas de alarmes falsos teóricas. Esta divergência possivelmente deu-se pelas suposições acerca da matriz de covariância dos dados coletados transformados em primeiros dígitos significativos. Foram testados procedimentos de inversão generalizada para a matriz de covariância, mas os resultados não foram satisfatórios devido a problemas de convergência. Uma vez que cada linha da matriz de dados de proporções de primeiros dígitos, tem a restrição da soma de seus valores em ser igual a um. Qualquer uma das colunas pode ser representada como uma combinação linear das demais, como se vê na Equação (4.1),

ou seja, os valores obtidos pelas proporções são dependentes entre si.

$$P_j = 1 - \sum_{i \neq j} P_i \quad (4.1)$$

onde:

$$i, j \in \{1, 2, \dots, 9\}.$$

A abordagem do segundo dígito teve repercussão pela aplicação na área forense. Nigrini (2012) define que o segundo dígito significativo tende a ter uma distribuição uniforme, mas isso é pouco interessante sob o ponto de vista da análise desenvolvida neste estudo. O interesse ocorre por tratar dos 2 dígitos significativos em conjunto.

Uma alternativa atrativa foi a de tratar o primeiro dígito significativo e tomar a proporção de vezes que ele aparece e distribuí-la no segundo dígito significativo, ou seja, a partir do primeiro dígito significativo gerar os segundos dígitos, 10, 11, 12, ..., 19, simplesmente dividindo esta proporção por 10. Da mesma forma para o dígito 2, gerar a sequência 20, 21, 22, ..., 29, sucessivamente até o dígito 9, gerando a sequência 90, 91, 92, ..., 99. Dessa forma, obtém-se a distribuição dos dois primeiros dígitos de forma conjunta. Mas ocorre um problema em razão disso, o vetor multivariado a ser analisado passa a ter não apenas 9 posições do primeiro ou 10 posições do segundo dígito, mas tem 90 posições com uma proporção muito pequena para cada uma dessas posições e uma variabilidade observada nos testes de simulação menores ainda. Isto faz com que a estratégia proposta de análise multivariada através dos procedimentos das Cartas de Controle  $\chi^2$  e  $T^2$  se tornem inviáveis, uma vez que é necessário inverter a matriz de covariância para efetuar o cálculo, a inversão da matriz resulta em um número de magnitude muito grande, o que faz com que não seja viável a adoção de um limite de controle pelas vias tradicionais. Uma pesquisa dedicada e mais aprofundada se faz necessária para o desenvolvimento de uma abordagem robusta para a análise conjunta dos dois dígitos significativos. Portanto, foi realizada a análise do segundo dígito significativo, mas identificou-se que as abordagens desenvolvidas e as ferramentas utilizadas não apresentaram viabilidade, optou-se então por não dar andamento na abordagem apenas do segundo dígito significativo neste estudo.

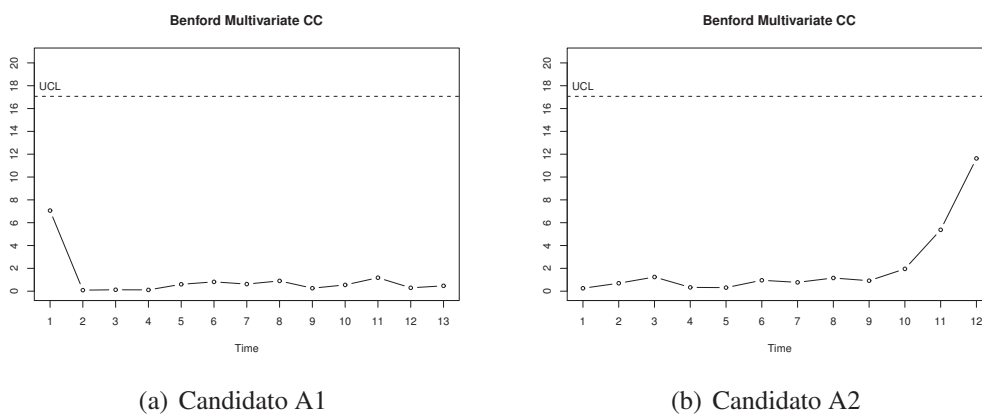


## 4.2.2 Exemplos de aplicação do modelo em dados reais

### 4.2.2.1 Eleições municipais 2016

No primeiro exemplo, são vistas as Cartas de Controle geradas pela análise dos dados da prestação de contas da campanha a prefeituras no Brasil em 2016. Optou-se por analisar as contas dos candidatos que ficaram em primeiro e em segundo lugar, independentemente da ocorrência de candidato vencedor no primeiro ou no segundo turno e os dados analisados constam do agrupamento de receitas e despesas de cada candidato em uma única planilha em Excel (salvo em formato .csv).

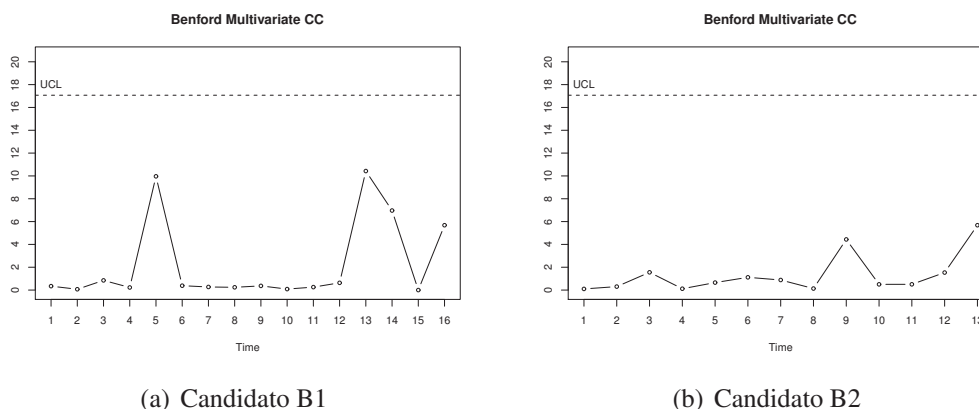
Na sequência, aparecem as cinco capitais analisadas, sendo que em cada figura aparecem as Cartas de Controle dos dois candidatos – Figuras 4 a 8. A figura da esquerda (a) refere-se ao candidato que ficou em segundo lugar e a figura da direita (b) refere-se ao candidato que ficou em primeiro lugar no pleito. Para efeito da análise da aderência dos dados, tomou-se o limite de controle padronizado  $LSC = 17,07$  referente ao  $ARL_0 = 20$  períodos, calculado por simulação, conforme Tabela 6, uma vez que o período de campanha em número de semanas não chega a 20, portanto, não há expectativa de um alarme falso em número menor que 20 estatísticas plotadas.



**Figura 4:** Cartas de Controle multivariadas de Benford para avaliação de suspeita de fraude - Prefeitura da Capital A - Eleições 2016

O candidato A1, em segundo lugar, apresentou em sua prestação de contas 3.735 linhas de movimentações financeiras em 13 semanas de lançamentos. O candidato A2, em primeiro lugar, apresentou em sua prestação de contas 899 linhas de movimentações

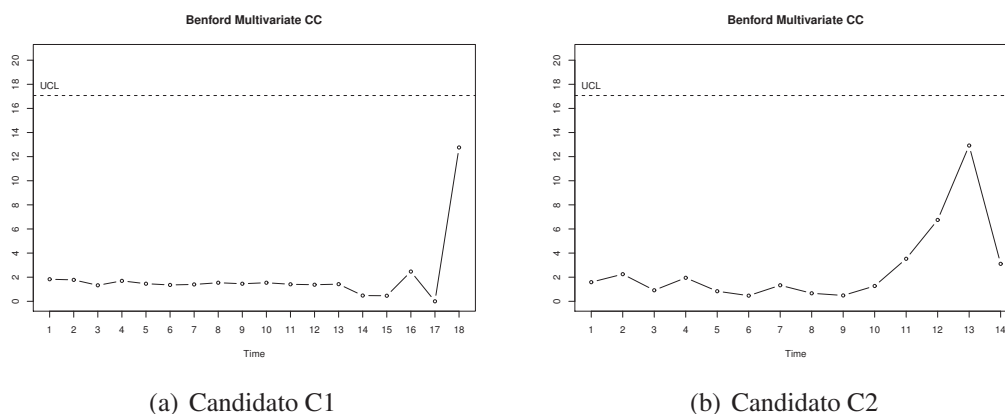
financeiras em 12 semanas de lançamentos. Verifica-se que, apesar das estatísticas do candidato A2 apresentarem tendência positiva nas últimas três semanas, não há efeito conclusivo em razão do término do período de análise. Conclui-se acerca dos dados analisados que tanto o candidato A1 como o candidato A2 apresentam todas as estatísticas dentro do limite de controle, ou seja, mostram um processo sob controle estatístico onde as estatísticas geradas aderem à frequência dos primeiros dígitos, não sinalizando suspeição sobre os dados das movimentações financeiras.



**Figura 5:** Cartas de Controle multivariadas de Benford para avaliação de suspeita de fraude - Prefeitura da Capital B - Eleições 2016

Considerando a Figura 5, o candidato B1, em segundo lugar, apresentou em sua prestação de contas 1.476 linhas de movimentações financeiras em 16 semanas de lançamentos. O candidato B2, em primeiro lugar, apresentou em sua prestação de contas 1.756 linhas de movimentações financeiras em 13 semanas de lançamentos. Verifica-se que as estatísticas do candidato B1 apresentam picos nas semanas 5, 13 e 14, demonstrando alguma instabilidade, porém não significativa. Na semana 16 o pico é devido à atribuição de valor igual a 0 (zero) na semana 15, que não apresentou movimentações. O candidato B2 mostra pico na estatística da semana 9 e tendência positiva nas últimas duas semanas, sem efeito conclusivo em razão do término do período de análise. Conclui-se acerca dos dados analisados que, tanto o candidato B1 como o candidato B2 apresentam todas as estatísticas dentro do limite de controle, ou seja, mostram o processo sob controle estatístico e que as estatísticas geradas aderem à frequência dos primeiros dígitos, não sinalizando suspeição sobre os dados das movimentações financeiras.

A Figura 6 mostra as Cartas de Controle da análise dos dados referentes ao pleito na Capital C.

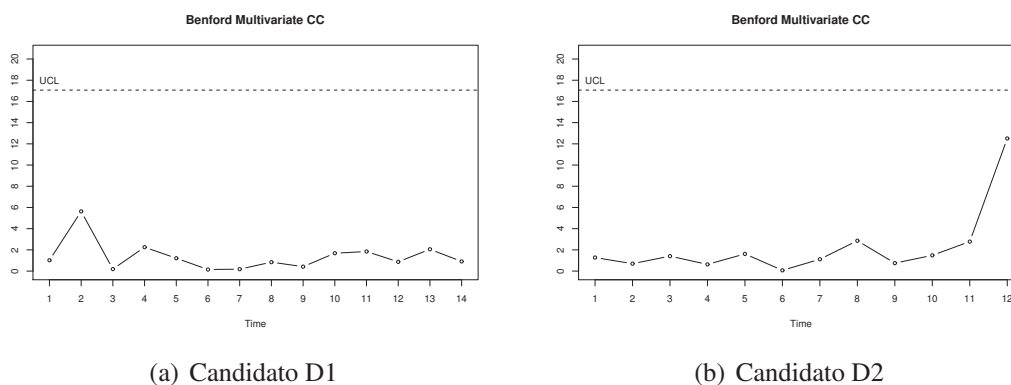


**Figura 6:** Cartas de Controle multivariadas de Benford para avaliação de suspeita de fraude - Prefeitura da Capital C - Eleições 2016

O candidato C1, em segundo lugar, apresentou em sua prestação de contas 18.344 linhas de movimentações financeiras em 18 semanas de lançamentos. O candidato C2, em primeiro lugar, apresentou em sua prestação de contas 2.000 linhas de movimentações financeiras em 14 semanas de lançamentos. Verifica-se que as estatísticas do candidato C1 apresentam estabilidade até a semana 16, na semana 17 foi atribuído valor 0 (zero) em razão de não haver lançamentos e um pico na semana 18 que demonstra uma tendência positiva sem confirmação em razão do término do período de análise. As estatísticas do candidato C2 mostram estabilidade até a semana 10 e tendência positiva da semana 11 a 13 e tendência negativa na última semana, apontando instabilidade inconclusa em razão do término do período de análise. Conclui-se acerca dos dados analisados que, tanto o candidato C1 como o candidato C2 apresentam todas estatísticas dentro do limite de controle, ou seja, mostram o processo sob controle estatístico e que as estatísticas geradas aderem à frequência dos primeiros dígitos, não sinalizando suspeição sobre os dados das movimentações financeiras.

Considerando os lançamentos de movimentação financeira da campanha municipal da Capital D das Eleições 2016, a Figura 7 mostra os resultados encontrados.

O candidato D1, em segundo lugar, apresentou em sua prestação de contas 503 linhas de movimentações financeiras em 14 semanas de lançamentos. O candidato D2

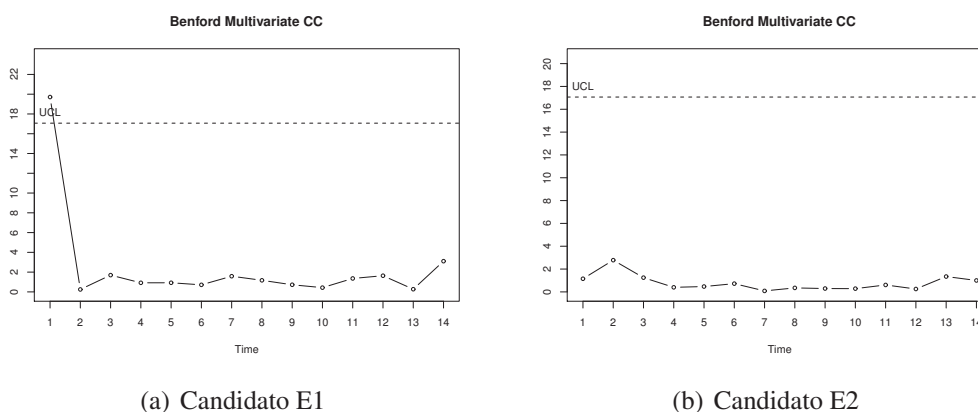


**Figura 7:** Cartas de Controle multivariadas de Benford para avaliação de suspeita de fraude - Prefeitura da Capital D - Eleições 2016

em primeiro lugar, apresentou em sua prestação de contas 1.992 linhas de movimentações financeiras em 12 semanas de lançamentos. Verifica-se que as estatísticas relativas ao candidato D1 apresentam relativa estabilidade em todo período, já o candidato D2 mostra estabilidade nas estatísticas até a semana 10 e tendência positiva da semana 11 até a 12, apontando tendência inconclusa em razão do término do período de análise. Conclui-se para os dados analisados que, tanto o candidato D1 como o candidato D2 apresentam todas estatísticas dentro do limite de controle, ou seja, mostram estatísticas que aderem à frequência dos primeiros dígitos e não sinalizam suspeição sobre os dados das movimentações financeiras.

Por fim, as Cartas de Controle da análise dos dados referentes ao pleito na Capital E são apresentadas na Figura 8.

O candidato E1, em segundo lugar, apresentou em sua prestação de contas 1.082 linhas de movimentações financeiras. O candidato E2, em primeiro lugar, apresentou em sua prestação de contas 653 linhas de movimentações financeiras e ambos candidatos apresentaram 14 semanas de lançamentos. Verifica-se um ponto fora de controle nas estatísticas do candidato E1 na semana 1, esse ponto, ao ser analisado, mostrou ser fruto de apenas uma movimentação financeira nesta semana e, portanto, provoca um viés na análise. De resto, todos pontos demonstram estabilidade durante o período analisado. As estatísticas do candidato E2 mostram estabilidade em todo período analisado. Concluindo a respeito dos dados computados que tanto o candidato E1 como o candidato E2 apresentam todas estatísticas dentro do limite de controle, ou seja, mostram estatísticas



(a) Candidato E1

(b) Candidato E2

**Figura 8:** Cartas de Controle multivariadas de Benford para avaliação de suspeita de fraude - Prefeitura da Capital E - Eleições 2016

sob controle estatístico e aderem à frequência dos primeiros dígitos, não sinalizando suspeição sobre os dados das movimentações financeiras.

#### 4.2.2.2 Lançamentos de contas de uma instituição financeira

O segundo exemplo de análise envolve uma planilha de lançamentos de contas de uma instituição financeira. Esta planilha é resultado dos registros efetuados após uma auditoria financeira e os dados foram originariamente utilizados por Silva, Korzenowski e Vaccaro (2014).

Os dados coletados referem-se a um período de cinco meses ou 150 dias corridos. Como há mais de um registro por dia, tem-se na base um total de 12.766 linhas de lançamentos financeiros. Por serem dias corridos, finais de semana e feriados não possuem lançamentos, de modo que o intervalo de datas torna-se menor após a exclusão dos dias sem lançamentos. Esta conta, cujos dados foram fornecidos, apresentou suspeita de fraude de acordo com os critérios da instituição e do BACEN (2009), referentes à prevenção de lavagem de dinheiro.

Para realizar a análise gráfica, foi executado um ajuste na Carta de Controle. A aplicação original neste volume de registros resultou em pequena variabilidade em termos de proporção dos primeiros dígitos significativos. Assim, ao invés de monitorar a proporção dos dígitos significativos, foi monitorado o número total absoluto. Foi utilizado para comparação o número esperado dos primeiros dígitos significativos de Benford

1, 2, . . . , 9, no total de 12.766 lançamentos financeiros das 150 datas.

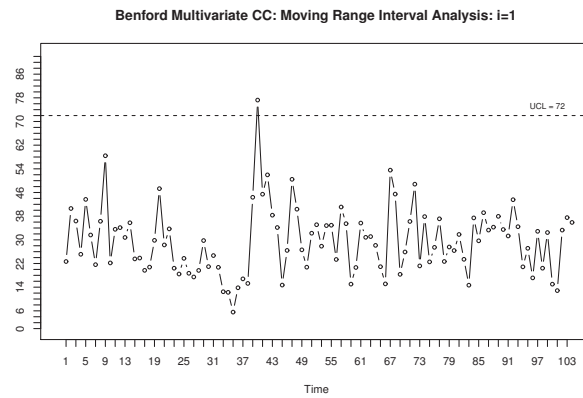
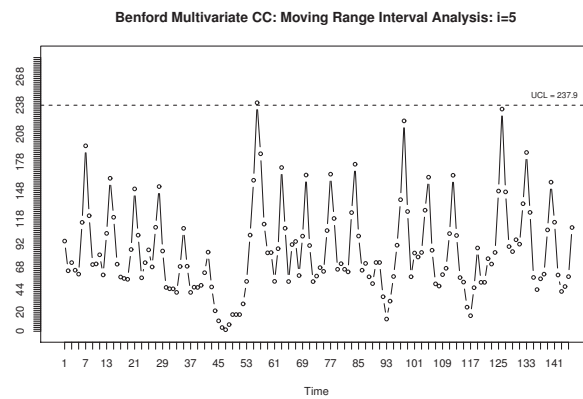
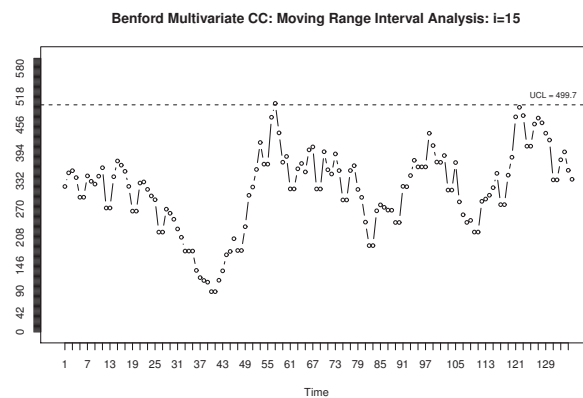
Neste caso, por tratar-se de uma análise retrospectiva, o limite de controle foi definido com 0,9973 de probabilidade, equivalente ao intervalo de  $\pm 3$  desvios-padrão  $\sigma$ , estabelecidos de acordo com a função de distribuição acumulada empírica – "ecdf", criada por Maechler (2008) e obtida com auxílio do pacote estatístico R. A Figura 9 apresenta três Cartas de Controle com intervalos de janelas de tempo móveis. Foram utilizadas janelas de tempo com tamanhos  $i=1$ ,  $i=5$  e  $i=15$  dias, permitindo a análise de forma diferenciada em cada uma delas, com objetivo de enriquecer as conclusões acerca das estatísticas geradas. As diferenças no limite de controle entre as três diferentes Cartas de Controle geradas são explicadas pelo tamanho da janela de tempo móvel e como isso impacta na função "ecdf".

Verifica-se nas três Cartas de Controle que as estatísticas apresentam-se com alta dispersão e cada uma delas com um ponto fora de controle. Na Carta de Controle da Figura 9(a)  $\Rightarrow i=1$ , a estatística da amostra 40 encontra-se além do limite de controle. Já na Carta de Controle da Figura 9(b)  $\Rightarrow i=5$ , a estatística da amostra 56 está fora de controle. Por fim, a Carta de Controle da Figura 9(c)  $\Rightarrow i=15$ , a estatística da amostra 57 encontra-se fora de controle.

Esperava-se uma suavização no comportamento da dispersão dos pontos plotados à medida que se aumenta o tamanho da janela de tempo  $i$ . Isto porque é esperado a partir do momento em que desvios em relação à distribuição de Benford em pontos específicos são mesclados com outras datas de lançamentos onde estas diferenças não são tão proeminentes. Porém, observa-se que da janela de tempo tamanho  $i = 1$  para a janela de tempo tamanho  $i = 5$  não se percebe a suavização na dispersão dos pontos plotados. Isto se deve ao fato da ocorrência de dias sem lançamentos pertencerem à janela de tempo.

Quando o tamanho da janela de tempo é aumentado de  $i = 5$  para  $i = 15$ , verifica-se então a suavização esperada. Note que apesar dessa alteração nos comportamentos observados da série de estatísticas de monitoramento, os pontos fora de controle concentram-se numa região temporal próxima em todas as três janelas de tempo apresentadas na Figura 9. Este fato indica para a ocasião de análise por parte dos auditores que esta região dos pontos é prioritária para avaliação detalhada por suspeitas de fraude. Percebe-se que dessa forma reduz-se o grupo de dados em que a auditoria deve realizar um exame detalhado. Este modelo de apresentação das Cartas de Controle permite

facilitar o trabalho dos auditores, tornando o processo mais ágil. Deste modo, permite imprimir maior velocidade ao trabalho de investigação, ao apontar o(s) período(s) onde devem receber atenção concentrada na aplicação dos procedimentos de análise em dados suspeitos.

(a) Intervalo da janela  $i = 1$ (b) Intervalo da janela  $i = 5$ (c) Intervalo da janela  $i = 15$ 

**Figura 9:** Cartas de Controle multivariadas de Benford para avaliação de suspeita de fraude - dados da Instituição Financeira



## 5 CONSIDERAÇÕES FINAIS

A lavagem de dinheiro representa perdas consideráveis para o sistema financeiro e causam sérios impactos de cunho social e econômico aos países, dentre eles, a desvalorização da moeda, pressão sobre o câmbio, a impunidade de criminosos e o consequente descrédito na comunidade internacional. Entidades internacionais como o BIS, da Suíça, em sua missão de ser o regulador dos bancos centrais, o FATF em suas recomendações de boas práticas, a SEC que administra e fiscaliza a aplicação do ato SOX nos Estados Unidos, no Brasil o COAF que atua como órgão regulador para a prevenção da lavagem de dinheiro e o combate ao terrorismo. Estas entidades surgiram para criar mecanismos e recomendações para que os países protejam a integridade de seus sistemas financeiros. Hierarquicamente as entidades financeiras tem a responsabilidade de implementar ações de controle com a finalidade de proteger os investidores, assegurar a retidão, confiabilidade e ética no cumprimento de regras de governança corporativa por meio das estratégias dos negócios em conformidade com a legislação. A forma mais prática de cumprir os regulamentos anti-lavagem de dinheiro é o aprimoramento contínuo, identificando as transações comerciais e financeiras normais dos clientes para que as atividades incomuns ou suspeitas sejam reconhecidas.

Assim, o objetivo geral da pesquisa consistiu em avaliar a efetividade do uso de cartas de controle multivariadas na detecção de presença de suspeitas de fraude financeira, com o fim de responder a questão: “Qual a efetividade do uso de Cartas de Controle multivariadas na detecção da presença de suspeitas de fraude financeira?” Como objetivos específicos, a pesquisa buscou verificar a efetividade de Cartas de Controle multivariadas na detecção da presença de movimentações financeiras com suspeição de fraude, com base na Lei de Benford; propor uma Carta de Controle multivariada para detecção de suspeitas de fraude financeira; e analisar a performance dos procedimentos citados nos objetivos específicos anteriores, no monitoramento, a partir da simulação de dados com a inclusão de movimentações fraudulentas.

Como apresentado, a pesquisa justificou-se em síntese por identificar que os recursos tecnológicos de *hardware* e de *software* baseado na Lei de Benford oferecem facilidade no processamento de grande massa de dados com carga computacional pouco onerosa. Os recursos de análise digitais apresentam-se confiáveis, precisos e estão incorporados em *softwares* populares. Os métodos de monitoramento de dados financeiros carecem

de um modelo computacional gerador de estatísticas que permita o aprendizado pela análise de padrões e de suas mudanças. A pesquisa buscou desenvolver um recurso estatístico de monitoramento de lançamentos financeiros em tempo real, capaz de identificar suspeita de fraudes por lavagem de dinheiro. A literatura pesquisada em lavagem de dinheiro indica que movimentações financeiras sem indício de fraude seguem a Lei de Benford e apontou uma lacuna de aplicação de um recurso de vigilância estatística para a análise de movimentações financeiras como suporte às atividades de auditoria.

A exemplo do controle estatístico de processos no meio industrial, foi desenvolvido um procedimento que consiste na implantação de uma Carta de Controle multivariada para o monitoramento de movimentações financeiras. O método de monitoramento aplicado foi a Carta de Controle  $\chi^2$ , uma vez que tenha sido admitida a aderência dos dados à Lei de Benford, ou seja, parâmetros conhecidos. Quando da aplicação de parâmetros estimados (parâmetros desconhecidos) na Fase I, foi feita a opção pela a Carta de Controle  $T^2$  de Hotelling. A geração de dados aleatórios teve a finalidade de certificação destes dados de modo padronizado de acordo com a frequência dos primeiros dígitos de Benford, para que possam ser utilizados na verificação da aderência de dados reais gerados, dessa forma caracterizando a aplicação do modelo de Carta de Controle proposto. Em geral, após uma mudança no processo, as Cartas de Controle levam um certo número de observações para detectar a mudança. Este tempo depende basicamente de dois fatores: o tipo de Carta de Controle e a taxa admissível de alarmes falsos. Quanto ao tipo de Carta de Controle, os modelos escolhidos são mais conservadores quanto ao tempo de detecção, podendo este tempo ser encurtado por outros modelos consagrados na literatura, como as Cartas de Controle MCuSum e WEWMA, por exemplo. Ainda, quanto ao número de observações até o sinal, os limites de controle exercem papel relevante, ao mesmo tempo em que influenciam a taxa de alarmes falsos. Para confirmar a capacidade de detecção do modelo de simulação foram adicionadas perturbações nos dados simulados e limites foram ajustados, a fim de atender diferentes particularidades operacionais na aplicação das Cartas de Controle.

Por simulação foi gerado um conjunto de dados históricos e partindo de uma matriz desses dados, obteve-se estimativas de média e da variabilidade para a proporção de cada dígito significativo. Com base nessa matriz de dados histórica, obteve-se as estatísticas de monitoramento, nas quais incrementou-se desvios  $\delta$  em torno da média das séries. O comportamento das Cartas de Controle seguiu um comportamento esperado

em aplicações deste tipo, reduzindo o número de observações necessárias até o sinal (alarme), conforme os desvios  $\delta$  inseridos foram aumentados. Entretanto, os valores de  $ARL_0$  observados na simulação não foram consistentes com a taxa teórica esperada de alarmes falsos. Novas simulações produzidas com desvio  $\delta = 0$  (zero) nos vetores médios foram executadas para determinação de limites de controle, buscando valores de  $ARL_0$  e respectivos limites de controle específicos. Com estes valores os limites de controle ficaram mais adequados às condições reais, esperadas para a aplicação das ferramentas estatísticas. Neste estudo, os limites de controle simulados foram utilizados no primeiro exemplo aplicado.

Os resultados obtidos por simulação atenderam aos dois primeiros objetivos específicos, tanto no que tange ao comportamento esperado das ferramentas estatísticas, quanto à capacidade de detecção dos sinais obtida nas séries de dados, uma vez adicionados os desvios  $\delta$  às médias em proporção de desvios-padrão  $\sigma$  e em se tratando da efetividade da aplicação de Cartas de Controle multivariadas na detecção de presença de movimentações financeiras com suspeição de fraude, com base na Lei de Benford e a proposição de uma Carta de Controle multivariada para detecção de suspeitas de fraude financeira.

A performance dos procedimentos propostos para o monitoramento evidenciou-se nas estratégias utilizadas nas simulações realizadas. Isso é verificado em um primeiro momento a partir do modelo com parâmetros conhecidos e, em um segundo momento, com a aprendizagem do comportamento da série da distribuição dos primeiros dígitos significativos a partir da série histórica. Dessa forma, atendendo a todos objetivos específicos e, por consequência, o geral da pesquisa.

Como visto no primeiro exemplo de aplicação do modelo em dados reais, foi utilizado o limite de controle corresponde ao valor de  $ARL_0 = 20$ , uma vez que tratou-se de uma série curta e com poucos lançamentos financeiros nas contas por semana (critério de agrupamento de lançamentos para obtenção dos dígitos adotado neste caso) na maior parte dos períodos. Os dados da prestação das contas da campanha eleitoral das Eleições Municipais em capitais dos Estados no Brasil em 2016 exemplificou a aplicabilidade do modelo na prática. No segundo exemplo, na aplicação em lançamentos de contas de uma instituição financeira, alterou-se a estratégia de análise, adotou-se não a proporção dos primeiros dígitos significativos, mas a frequência esperada em termos absolutos. Para tanto, os limites de controle das estatísticas foram definidos através da

função de distribuição cumulativa empírica durante a Fase I de monitoramento, ou seja, a fase de análise retrospectiva.

Com a exemplificação apresentada, torna-se clara a contribuição trazida pelo trabalho, tanto para a academia quanto para a sociedade.

A contribuição acadêmica identifica-se pelo desenvolvimento de um modelo de aplicação das Cartas de Controle multivariadas e da Lei de Benford, por uma abordagem inovadora de controle estatístico de processos voltado à área financeira, fazendo uso de recurso computacional acessível, de fácil processamento, confiável e preciso que permite aprimoramento por novas abordagens na academia.

No que tange à contribuição à sociedade, se dá pela oportunidade de aplicação do modelo por entidades financeiras no monitoramento de padrões das movimentações financeiras de seus clientes, para identificação de suspeitas de fraude. E pela comunidade, a possibilidade de análise dos dados de organizações civis e estatais, divulgados nos canais de informação, de modo a proporcionar a prática cidadã pela constatação da idoneidade dos fatos e dos dados.

O estudo avaliou ainda a distribuição dos dois primeiros dígitos sob os enfoques apresentados, porém, a modelagem ficou prejudicada em função do aumento de variáveis em análise, resultado comum quando trata-se de monitoramento de processos multivariados. Entretanto, sugere-se algumas abordagens para estudo do segundo dígito:

- Análise separada do primeiro e do segundo dígitos, obtendo-se para cada dígito significativo (primeiro e segundo) um vetor de monitoramento;
- Monitoramento conjunto dos primeiro e segundo dígitos por Carta de Controle multivariada não paramétrica; e
- Análise conjunta e individual dos dígitos através do desenvolvimento de estratégias de autoinicialização, onde se faz a estimativa da variabilidade por sequência de diferenças.

Ainda como sugestão para a continuidade da pesquisa, aprofundar a análise dos dígitos por janelas móveis de tempo de lançamentos, como apresentado no segundo exemplo aplicado nesta pesquisa, empregando neste modelo de janelas móveis as sugestões de análise individual e conjunta do primeiro e segundo dígitos significativos, bem como estratégias não-paramétricas e de auto-inicialização.

## REFERÊNCIAS

- ALISSON, E. FUNDAÇÃO DE AMPARO A PESQUISA DO ESTADO DE SÃO PAULO (FAPESP). **Pesquisadores criam métodos estatísticos para prever fraudes em operações financeiras**. São Paulo, 27 out. 2015. Disponível em: <[http://agencia.fapesp.br/pesquisadores\\_criam\\_metodos\\_estatisticos\\_para\\_prever\\_fraudes\\_em\\_operacoes\\_financeiras/22120/](http://agencia.fapesp.br/pesquisadores_criam_metodos_estatisticos_para_prever_fraudes_em_operacoes_financeiras/22120/)>. Acesso em: 26 nov. 2015.
- ALT, F. B.; GRIMSHAW, S. D. Multivariate Quality Control. In: **Encyclopedia of Operations Research and Management Science**. Springer, 2013. p. 1014–1018.
- AMORIM, E. N. C.; CARDOZO, M. A.; VICENTE, E. F. Os impactos da implementação de controles internos, auditoria e compliance no combate e prevenção à lavagem de dinheiro no Brasil. **Enfoque: Reflexão Contábil**, v. 31, n. 3, p. 23–35, 2012.
- ANDRIOTTI, J. L. S. **Controle de qualidade em prospecção geoquímica**. 2. ed. São Paulo:Record, 2007.
- BACEN. BANCO CENTRAL DO BRASIL (BACEN). **Circular nº 3.461, de 24 de Julho de 2009**. Consolida as regras sobre os procedimentos a serem adotados na prevenção e combate às atividades relacionadas com os crimes previstos na Lei nº 9.613, de 3 de março de 1998. Disponível em: <<https://www.bcb.gov.br/pre/normativos/busca/normativo.asp?tipo=circ&ano=2009&numero=3461>>. Acesso em: 13 jan. 2017.
- BENFORD, F. The law of anomalous numbers. **Proceedings of the American Philosophical Society**, p. 551–572, 1938.
- BERTONI, F. F.; CARVALHO, D. **Criminal compliance e lavagem de dinheiro**. Disponível em: <[http://ebooks\\_pucrs.br\\_edipucrs\\_anais\\_cienciascriminais\\_IV/08\\_pdf](http://ebooks_pucrs.br_edipucrs_anais_cienciascriminais_IV/08_pdf)>. Acesso em: 10 dez. 2015.
- BERTRAND, J. W. M.; FRANSOO, J. C. Operations management research methodologies using quantitative modeling. **International Journal of Operations & Production Management**, v. 22, n. 2, p. 241–264, 2002.
- BHATTACHARYA, S.; XU, D.; KUMAR, K. An ANN-based auditor decision support system using Benford's law. **Decision support systems**, v. 50, n. 3, p. 576–584, 2011.
- CARDOZO, J. S. S. Origem e conceitos de auditoria. **Revista de Contabilidade do Mestrado em Ciências Contábeis da UERJ**, v. 2, n. 2, p. 27–36, 2014.

- CARNEIRO, A. **Auditoria e controle de sistemas de informação**. FCA, 2009.
- CHESNEAUX, J. Devemos fazer tabula rasa do passado?: sobre história e os historiadores. trad. marcos a. da silva: são paulo: ática, 1995. **Fundamentos**, v. 109, 1995.
- CHUNG, I.; JOO, J.; KIM, S.; CHO, H.; AHN, N. Verification Method to Detect the Fake Test Data in Military Supplies. **Journal of the Korea Academia-Industrial cooperation Society**, v. 17, n. 3, p. 231–240, 2016.
- CLIPPE, P.; AUSLOOS, M. Benford's law and Theil transform of financial data. **Physica A: Statistical Mechanics and its Applications**, v. 391, n. 24, p. 6556–6567, 2012.
- COAF. CONSELHO DE CONTROLE DE ATIVIDADES FINANCEIRAS (COAF). **Lavagem de dinheiro e financiamento do terrorismo**. Brasília, 2015. Disponível em: <<http://www.coaf.fazenda.gov.br/acesso-a-informacao/Institucional/competencias>>. Acesso em: 15 dez. 2015.
- COAF. CONSELHO DE CONTROLE DE ATIVIDADES FINANCEIRAS (COAF). **Informação institucional publicações e competências**. Brasília, 2015. Disponível em: <<http://www.coaf.fazenda.gov.br/menu/pld-ft/publiccoes>>. Acesso em: 15 dez. 2015.
- COAF. CONSELHO DE CONTROLE DE ATIVIDADES FINANCEIRAS (COAF). **Pessoas obrigadas - comunicações**. Brasília, 2016. Disponível em: <[http://www.coaf.fazenda.gov.br/backup/Pessoas\\_Obrigadas/perguntas-e-respostas#comunicacoes](http://www.coaf.fazenda.gov.br/backup/Pessoas_Obrigadas/perguntas-e-respostas#comunicacoes)>. Acesso em: 22 mar. 2016.
- COAF. CONSELHO DE CONTROLE DE ATIVIDADES FINANCEIRAS (COAF). **Pessoas obrigadas - Art. 9º Lei nº 9.613**. Brasília, 2016. Disponível em: <<http://www.coaf.fazenda.gov.br/links-externos/pessoas-obrigadas-art.-9o-da-lei-9.613>>. Acesso em: 10 mai. 2016.
- COAF. CONSELHO DE CONTROLE DE ATIVIDADES FINANCEIRAS (COAF). **Pessoas obrigadas - comunicações de não ocorrência ou declaração negativa**. Brasília, 2016. Disponível em: <[http://www.coaf.fazenda.gov.br/backup/Pessoas\\_Obrigadas/perguntas-e-respostas#Comunicacoes\\_de\\_ao\\_ocorrencia](http://www.coaf.fazenda.gov.br/backup/Pessoas_Obrigadas/perguntas-e-respostas#Comunicacoes_de_ao_ocorrencia)>. Acesso em: 22 mar. 2016.
- COELHO, I. INSTITUTO DE AUDITORES INDEPENDENTES DO BRASIL (IBRACON). **Fraude não é erro e auditoria não faz seguro**. São Paulo, Nov 2015. Disponível em: <<http://www.ibracon.com.br/ibracon/Portugues/detArtigo.php?cod=8>>. Acesso em: 24 nov. 2015.

CREPALDI, S. A. Auditoria contábil. **São Paulo: Atlas**, 2004.

CROWDER, S. V. Design of exponentially weighted moving average schemes. **Journal of Quality Technology**, v. 21, n. 3, p. 155–162, 1989.

CUNHA, F. C. R.; BUGARIN, M. S. Lei de Benford e Auditoria de Obras Públicas: uma análise de sobrepreço na reforma do maracanã. **Revista do TCU**, n. 131, p. 46–53, 2014.

CURY, L.; CAPOBIANCO, L. **Princípios da História das Tecnologias da Informação e Comunicação Grandes Invenções**. 2014.

CÁCERES, J. L. H.; GARCIA, J. L. P.; ORTIZ, C. M. M.; DOMÍNGUEZ, L. G. First digit distribution in some biological data sets. Possible explanation for departures from Benford's Law. **Revista Electrónica Biomédica / Electronic Journal Biomedic**, v. 1, p. 27 – 35, 2008. Disponível em: <<http://www.biomed.uninet.edu/2008/n1/caceres.html/>>. Acesso em: 07 dez. 2015.

DANIALI, G. E-money laundering prevention. **New Marketing Research Journal**, v. 4, 2014.

DÍAZ, J.; GALLART, J.; RUIZ, M. On the ability of the Benford's Law to detect earthquakes and discriminate seismic signals. **Seismological Research Letters**, 2014.

DÍAZ, M. d. C. M.; VERA, I. A. Orígenes y clasificación de la auditoría de la información. **Revista Cubana de Información en Ciencias de la Salud**, v. 14, n. 5, p. 15, 2006.

DRAKE, P. D.; NIGRINI, M. J. Computer assisted analytical procedures using Benford's law. **Journal of Accounting Education**, v. 18, n. 2, p. 127–146, 2000.

DUCZMAL, L.; KULLDORFF, M.; HUANG, L. Evaluation of spatial scan statistics for irregularly shaped clusters. **Journal of Computational and Graphical Statistics**, 2012.

DURTSCHI, C.; HILLISON, W.; PACINI, C. The effective use of Benford's Law to assist in detecting fraud in accounting data. **Journal of Forensic Accounting - R.T. Edwards, Inc.**, v. V, p. 17–34, 2004.

ENCCLA. BRASIL - MINISTÉRIO DA JUSTIÇA (ENCCLA). **Estratégia nacional de combate a corrupção e lavagem de dinheiro**. Brasília, 2004. Disponível em: <<http://enccla.camara.leg.br/acoes/metas-de-2004/>>. Acesso em: 06 dez. 2015.

FATF. FINANCIAL ACTION TASK FORCE IN MONEY LAUNDERING (FATF). **History**. Paris, France, 1989. Disponível em: <<http://http://www.fatf-gafi.org/about/historyofthefatf/>>. Acesso em: 06 dez. 2015.

FATF. FINANCIAL ACTION TASK FORCE IN MONEY LAUNDERING (FATF). **Annual Report 1999-2000**. Paris, France, 2000. Disponível em: <<http://www.fatf-gafi.org/media/fatf/documents/reports/1999%202000%20ENG.pdf>>. Acesso em: 06 dez. 2015.

FATF. FINANCIAL ACTION TASK FORCE IN MONEY LAUNDERING (FATF). **Anti-money laundering and terrorist financing measures and financial inclusion**. Paris, France, 2011. Disponível em: <<http://www.fatf-gafi.org/media/fatf/content/images/AML%20CFT%20measures%20and%20financial%20inclusion.pdf>>. Acesso em: 08 ago. 2016.

FATF. FINANCIAL ACTION TASK FORCE IN MONEY LAUNDERING (FATF). **FATF steps up the fight against money laundering and terrorist financing**. Paris, France, 2012. Disponível em: <<http://www.fatf-gafi.org/publications/fatfrecommendations/documents/fatfstepsupthefightagainstmoneylaunderingandterroristfinancing.html>>. Acesso em: 08 ago. 2016.

FATF. FINANCIAL ACTION TASK FORCE IN MONEY LAUNDERING (FATF). **International standards on combating money laundering and the financing of terrorism & proliferation - FATF recommendations**. Paris, France, 2012. updated June 2016. Disponível em: <<http://www.fatf-gafi.org/publications/fatfrecommendations/documents/fatf-recommendations.html>>. Acesso em: 06 ago. 2016.

FATF. FINANCIAL ACTION TASK FORCE IN MONEY LAUNDERING (FATF). **Methodology for assessing compliance with the FATF recommendations and the effectiveness of AML/CFT systems**. Paris, France, 2013. updated February 2016. Disponível em: <[www.fatf-gafi.org/publications/fatfrecommendations/documents/fatfissuesnewmechanismtostrengthenmoneylaunderingandterroristfinancingcompliance.html](http://www.fatf-gafi.org/publications/fatfrecommendations/documents/fatfissuesnewmechanismtostrengthenmoneylaunderingandterroristfinancingcompliance.html)>. Acesso em: 06 mar. 2016.

FATF. FINANCIAL ACTION TASK FORCE IN MONEY LAUNDERING (FATF). **Procedures for the FATF fourth round of AML/CFT mutual evaluations**. Paris, France, 2013. updated October 2015. Disponível em: <[www.fatf-gafi.org/publications/fatfrecommendations/documents/fatfissuesnewmechanismtostrengthenmoneylaunderingandterroristfinancingcompliance.html](http://www.fatf-gafi.org/publications/fatfrecommendations/documents/fatfissuesnewmechanismtostrengthenmoneylaunderingandterroristfinancingcompliance.html)>. Acesso em: 06 mar. 2016.

FATF. FINANCIAL ACTION TASK FORCE IN MONEY LAUNDERING (FATF). **International standards on combating money laundering and the financing of terrorism & proliferation**. Paris, France, 2015. updated October 2015. Disponível em: <[www.fatf-gafi.org/recommendations.html](http://www.fatf-gafi.org/recommendations.html)>. Acesso em: 06 dez. 2015.



FATF. FINANCIAL ACTION TASK FORCE IN MONEY LAUNDERING (FATF). **G20 Finance ministers and central bank governors call on FATF to intensify its efforts**. Shanghai, China, 2016. Disponível em: <<http://www.fatf-gafi.org/publications/fatfrecommendations/documents/g20-resolved-to-combat-terrorist-financing-call-on-fatf-to-intensify-efforts.html>>. Acesso em: 21 mar. 2016.

FATF. FINANCIAL ACTION TASK FORCE IN MONEY LAUNDERING (FATF). **Outcomes of the plenary meeting of the FATF, Paris, 17-19 February 2016**. Paris, France, 2016. Disponível em: <<http://www.fatf-gafi.org/publications/fatfgeneral/documents/outcomes-plenary-february-2016.html>>. Acesso em: 26 mar. 2016.

FATF. FINANCIAL ACTION TASK FORCE IN MONEY LAUNDERING (FATF). **Outcomes of the plenary meeting of the FATF, Korea, 22-24 June 2016**. Busan, Korea, 2016. Disponível em: <<http://www.fatf-gafi.org/publications/fatfgeneral/documents/plenary-outcomes-june-2016.html>>. Acesso em: 06 out. 2016.

FAVARETTTO, F. Verificação da qualidade de dados através da lei de Benford. In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO — XXVII ENEGEP, 2007. **Anais...** 2007. Disponível em: <[http://www.abepro.org.br/biblioteca/ENEGEP2007\\_TR640478\\_9130.pdf](http://www.abepro.org.br/biblioteca/ENEGEP2007_TR640478_9130.pdf)>. Acesso em: 06 dez. 2015.

FORSTER, R. P. **Auditoria Contábil em Entidades do Terceiro Setor. Uma Aplicação da Lei Newcomb-Benford**. 66 f. 2010. Dissertação (Mestrado em Ciências Contábeis) - Programa de Pós-Graduação em Ciências Contábeis, Universidade de Brasília (UNB), Brasília, 2010.

FRANCIS, J. R. What do we know about audit quality? **The British accounting review**, v. 36, n. 4, p. 345–368, 2004.

FRISÉN, M. **Financial surveillance**. John Wiley & Sons, 2008. v. 71.

FRY, D. E.; PINE, M.; JONES, B. L.; MEIMBAN, R. J. Control charts to identify adverse outcomes in elective colon resection. **The American Journal of Surgery**, v. 203, n. 3, p. 392–396, 2012.

GAO, S.; XU, D. Conceptual modeling and development of an intelligent agent-assisted decision support system for anti-money laundering. **Expert Systems with Applications**, v. 36, n. 2, p. 1493–1504, 2009.

- GASB. GOVERNAMENTAL ACCOUNTING STANDARD BOARD (GASB). **Implementation Guides**. USA, 2016. Disponível em: <<http://www.gasb.org/jsp/GASB/Page/GASBSectionPage&cid=1176163026371>>. Acesso em: 21 ago. 2016.
- GERTZ, M.; GULDENTOPS, E.; STROUS, L. **Integrity, internal control and security in information systems**. Springer Science, 2013.
- GOMES, E. D.; ARAÚJO, A. F. d.; BARBOZA, R. J. Auditoria: alguns aspectos a respeito de sua origem. **Revista Científica Eletrônica de Ciências Contábeis**. Garça, 2009.
- HALES, D. N.; SRIDHARAN, V.; RADHAKRISHNAN, A.; CHAKRAVORTY, S. S.; SIHA, S. M. Testing the accuracy of employee-reported data: an inexpensive alternative approach to traditional methods. **European Journal of Operational Research**, v. 189, n. 3, p. 583–593, 2008.
- HEVNER, V. A. R.; MARCH, S. T.; PARK, J.; RAM, S. Design science in information systems research. **MIS quarterly**, v. 28, n. 1, p. 75–105, 2004.
- HILL, T. P. A statistical derivation of the significant-digit law. **Statistical Science**, v. 10, n. 4, p. 354–363, 1995. Disponível em: <<http://www.jstor.org/stable/2246134>>. Acesso em: 06 dez. 2015.
- HILL, T. P. The First Digit Phenomenon A century-old observation about an unexpected pattern in many numerical tables applies to the stock market, census statistics and accounting data. **American Scientist**, v. 86, n. 4, p. 358–363, 1998.
- HJALMARS, U.; KULLDORFF, M.; GUSTAFSSON, G.; NAGARWALLA, N. Childhood leukaemia in Sweden: using gis and a spatial scan statistic for cluster detection. **Statistics in medicine**, v. 15, n. 7-9, p. 707–715, 1996.
- ISA, Y. M.; SANUSI, Z. M.; HANIFF, M. N.; BARNES, P. A. Money Laundering Risk: from the bankers’ and regulators perspectives. **Procedia Economics and Finance**, v. 28, p. 7–13, 2015.
- IUDÍCIBUS, S. d.; MARTINS, E.; CARVALHO, L. N. Contabilidade: Aspectos Relevantes da Epopéia de sua Evolução. **Revista Contabilidade & Finanças**, v. 16, n. 38, p. 7–19, 2005.
- JENSEN, W. A.; JONES-FARMER, L. A.; CHAMP, C. W.; WOODALL, W. H. Effects of parameter estimation on control chart properties: a literature review. **Journal of Quality Technology**, v. 38, n. 4, p. 349, 2006.

JOHNSON, G. G.; WEGGENMANN, J. Exploratory Research Applying Benford's Law to Selected Balances in the Financial Statements of State Governments. **Academy of Accounting and Financial Studies Journal**, v. 17, n. 3, p. 31, 2013.

KELTON, W. D.; LAW, A. M. **Simulation modeling and analysis**. McGraw Hill Boston, 2000.

KORZENOWSKI, A. L.; SIMÕES, W. L.; POLLMANN, C.; CHIWIACOWSKY, L. D. **Data simulation road map for Control Charts proposition and performance evaluation**. 2017, *Unpublished*.

KULLDORFF, M. A spatial scan statistic. **Communications in Statistics-Theory and methods**, v. 26, n. 6, p. 1481–1496, 1997.

LAGIOIA, U. C. T.; ARAÚJO, I. J. C.; FILHO, F. A. B.; BARROS, M. A. B.; SANTOS, S. G. O. A. Aplicabilidade da Lei de Newcomb Benford nas fiscalizações do imposto sobre serviços-ISS. **Revista Contabilidade & Finanças**, v. 22, n. 56, p. 203–224, 2011.

LOWRY, C. A.; MONTGOMERY, D. C. A review of multivariate control charts. **III transactions**, v. 27, n. 6, p. 800–810, 1995.

LUCAS, J. M.; SACCUCCI, M. S. Exponentially weighted moving average control schemes: properties and enhancements. **Technometrics**, v. 32, n. 1, p. 1–12, 1990.

MAEHLER, M. **Empirical cumulative distribution function 'ecdf' object**. The R project Free software environment for statistical computing and graphics. Disponível em: <<https://stat.ethz.ch/R-manual/R-devel/library/stats/html/ecdf.html>>. Acesso em: 13 jan. 2017.

MARCONI, M. d. A.; LAKATOS, E. M. **Fundamentos da metodologia científica**. 7. ed. São Paulo:Atlas, 2010.

MARTINS, E.; MARTINS, V. A.; MARTINS, É. A. Normatização contábil: ensaio sobre sua evolução e o papel do cpc. **Revista de Informação Contábil**, v. 1, n. 1, p. 7–30, 2007.

MARTINS, I. **Auditoria dos sistemas de informação das instituições financeiras**. 2013. Tese, 2013.

MASON, R. L.; TRACY, N. D.; YOUNG, J. C. A practical approach for interpreting multivariate T2 control chart signals. **Journal of Quality Technology**, v. 29, n. 4, p. 396, 1997.

MEBANE JR, W. R. Election forensics: the second-digit benford's law test and recent american presidential elections. In: ELECTION FRAUD CONFERENCE, 2006. **Anais...** 2006.

MIGUEL, P. A. C.; FLEURY, A. C. C. **Metodologia de pesquisa em engenharia de produção e gestão de operações**. 2. ed. Rio de Janeiro: Elsevier, 2012.

MILAN, M.; FERNANDES, R. A. T. Qualidade das operações de preparo de solo por controle estatístico de processo. **Scientia Agrícola**, v. 59, n. 02, p. 340–350, 2002.

MINAMISAVA, R.; NOUER, S. S.; MORAIS NETO, O. L. de; MELO, L. K. Spatial clusters of violent deaths in a newly urbanized region of Brazil: highlighting the social disparities. **International journal of health geographics**, v. 8, n. 1, p. 1, 2009.

MONTGOMERY, D. C. **Introdução ao controle estatístico da qualidade**. Tradução de Ana Maria Lima de Farias, Vera Regina Lima de Farias e Flores. 7. ed. Rio de Janeiro : LTC, 2016.

MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the theory of statistics**. New York (USA), McGraw-Hill, 1974.

MORO, S. F. Autonomia do crime de lavagem e prova indiciária. **Revista CEJ**, v. 12, n. 41, p. 11–14, 2008.

MORO, S. F. **Crime de lavagem de dinheiro**. 1. ed. São Paulo : Saraiva, 2010.

MOTTA-JR, E. P. **Investigação de modelo de auditoria contínua para tribunais de contas**. 2010. 133 p. Dissertação (Mestrado Profissional em Ciências da Computação) — Pós-Graduação em Ciência da Computação, Universidade Federal de Pernambuco, Recife, 2010.

NIGRINI, M. **Forensic analytics: methods and techniques for forensic accounting investigations**. John Wiley & Sons, 2011. v. 558.

NIGRINI, M. **Benford's Law: applications for forensic accounting, auditing, and fraud detection**. John Wiley & Sons, 2012. v. 586.

NIGRINI, M. J. Continuous auditing. **Ernst & Young Center for Auditing Research and Advanced Technology University of Kansas**, 2000.

NIGRINI, M. J.; MITTERMAIER, L. J. The use of Benford's law as an aid in analytical procedures. **Auditing**, v. 16, n. 2, p. 52, 1997.

NORDSTROM, F.; AF WETTERSTEDT, S.; JOHNSON, S.; CEBERG, C.; BACK, S. A. Control chart analysis of data from a multicenter monitor unit verification study. **Radiotherapy and Oncology**, v. 102, n. 3, p. 364–370, 2012.

OLIVEIRA, A. **Método de auditoria a sistemas de informação**. 2006.

ORITA, M.; HAGIWARA, Y.; MORITOMO, A.; TSUNOYAMA, K.; WATANABE, T.; OHNO, K. Agreement of drug discovery data with Benford's law. **Expert Opinion on Drug Discovery**, v. 8, n. 1, p. 1–5, 2013.

PACHECO FILHO, A. B. O projeto básico como elemento de responsabilidade na gestão pública. **Revista do TCU**, n. 99, p. 73–76, 2004.

PERICCHI, L.; TORRES, D. Quick Anomaly Detection by the Newcomb—Benford Law, with Applications to Electoral Processes Data from the USA, Puerto Rico and Venezuela. **Statistical Science**, p. 502–516, 2011.

PINHO, C. Implementation of Internal Control Systems to Combat Money Laundering and Financing of Terrorism—An Applied Exploratory Study to the Angolan Financial Sector. **International Journal of Business and Social Research**, v. 5, n. 7, p. 38–50, 2015.

PIRENNE, H. **História econômica e social da Idade Média**. São Paulo, Mestre Jou, 1968.

PLACKETT, R. L. Karl Pearson and the chi-squared test. **International Statistical Review/Revue Internationale de Statistique**, p. 59–72, 1983.

PORTER, W. T.; BURTON, J. C.; RAMÍREZ, R. Q. **Auditoría: un análisis conceptual**. Diana México DF, 1980.

PRODANOV, C. C.; FREITAS, E. C. **Metodologia do Trabalho Científico: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico -2ª edição**. Editora Feevale, 2013.

RIBEIRO, J. d. M.; FILHO, J. M. D. Aplicação de métodos quantitativos em auditoria: propostas para otimizar procedimentos e reduzir riscos. **Revista de Contabilidade e Organizações**, v. 1, n. 1, p. 43, 2007.

ROBERTS, S. Control chart tests based on geometric moving averages. **Technometrics**, v. 1, n. 3, p. 239–250, 1959.

ROCHA, F. R.; CYMROT, R. **Aplicação da lei de Benford em dados provenientes de Controle Estatístico de Processos**. In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO – XXIX ENEGEP, 2009. Salvador, Bahia. Anais eletrônicos, 2009. Disponível em: <[http://www.abepro.org.br/biblioteca/enegep2009\\_TN\\_STO\\_092\\_623\\_13677.pdf](http://www.abepro.org.br/biblioteca/enegep2009_TN_STO_092_623_13677.pdf)>, Acesso em: 10 dez. 2015.

- SAMBRIDGE, M.; TKALČIĆ, H.; JACKSON, A. Benford's law in the natural sciences. **Geophysical research letters**, v. 37, n. 22, 2010.
- SANTOS, J.; DINIZ, J. A.; CORRAR, L. J. O Foco é a Teoria Amostral nos Campos da Auditoria Contábil Tradicional e da Auditoria Digital: testando a lei de newcomb-benford para o primeiro dígito nas contas públicas. **BBR-Brazilian Business Review**, v. 2, n. 1, p. 71–89, 2005.
- SCHNEIDER, F. Turnover of organized crime and money laundering: some preliminary empirical findings. **Public choice**, v. 144, n. 3-4, p. 473–486, 2010.
- SELLKE, T.; BAYARRI, M.; BERGER, J. O. Calibration of  $\rho$  values for testing precise null hypotheses. **The American Statistician**, v. 55, n. 1, p. 62–71, 2001.
- SHEWHART, W. A. **Economic control of quality of manufactured product**. ASQ Quality Press, 1931.
- SILVA, A.; KORZENOWSKI, A. L.; VACCARO, G. L. R. Uma aplicação da lei de Benford na identificação de padrões estatisticamente assinaláveis de suspeitas de fraude por lavagem de dinheiro. **Espacios**, v. 35, n. 7, p. 43–62, 2014.
- STOERK, T. Statistical corruption in Beijing's air quality data has likely ended in 2012. **Atmospheric Environment**, v. 127, p. 365–371, 2016.
- SUSPIRO, A. OBSERVADOR. **Os grandes escândalos empresariais dos últimos 20 anos**. Portugal, 03 Out. 2015. Disponível em: <<https://observador.pt/especiais/os-grandes-escandalos-empresariais-dos-ultimos-20-anos>>. Acesso em: 17 out. 2016.
- TEAM, R. C. **The R project for statistical computing**. Free software environment for statistical computing and graphics. Disponível em: <<https://www.r-project.org/>>. Acesso em: 20 mar. 2016.
- TOLEDO, P.; RIQUELME, S.; CAMPOS, J. Earthquake source parameters that display the first digit phenomenon. **Nonlinear Processes in Geophysics**, v. 22, n. 5, p. 625, 2015.
- TSE. BRASIL - TRIBUNAL SUPERIOR ELEITORAL (TSE). **Divulgação de candidaturas e das contas eleitorais das eleições municipais de 2016**. Brasília, 2017. Disponível em: <<http://divulgacandcontas.tse.jus.br/divulga/#/>>. Acesso em: 07 jan. 2017.
- TSE. BRASIL - TRIBUNAL SUPERIOR ELEITORAL (TSE). **Consulta por região/UF/município do ano de 2016. Estatística do eleitorado**. Brasília, 2017.

Disponível em:

<<http://www.tse.jus.br/eleitor/estatisticas-de-eleitorado/consulta-quantitativo>>.

Acesso em: 07 jan. 2017.

UMIT, F.; CIGDEM, A. Multivariate Quality Control: a historical perspective. **Yildiz Technical University**, p. 54–65, 2001.

VEYNE, P. História da vida privada: do império romano ao ano mil. **Tradução de Hildegard Feist. Coleção dirigida por Philippe Ariès e Georges Duby**. v. 1. São Paulo: Companhia das Letras, 1990.

WOODALL, W. H.; MONTGOMERY, D. C. Some current directions in the theory and application of statistical process monitoring. **Journal of Quality Technology**, v. 46, n. 1, p. 78–94, 2014.

YUEN, J. Computer Assisted Auditing Techniques Money Laundering Detection. **ACC 626 Computer Technology**, v. 5, n. 4, p. 316–334, 2014.

## APÊNDICE A CÓDIGOS NO SOFTWARE R

### A.1 Código geral: Arquivo main.R

```

1 # Codigo para simulacao: Dissertacao Davenilcio Souza
2 # 07/7/2016
3 # Para carregar utilize { source("main.R") } no diretorio de trabalho
4 # Inicio do codigo
5 # Carregando pacotes necessarios
6 require(MASS)
7 # Benford distribution
8 dbenford ← function(x) {
9   log10(1+1/x)
10 }
11 pbenford ← function(q) {
12   cumprobs ← cumsum(dbenford(1:9))
13   return(cumprobs[q])
14 }
15 qbenford ← function(p) {
16   cumprobs ← cumsum(dbenford(1:9))
17   cumprobs[9] ← 1
18   quantiles ← sapply(p, function(x) {10-sum(cumprobs >= x)})
19   return(quantiles)
20 }
21 rbenford ← function(n) {
22   sample(1:9, size = n, replace = TRUE, prob = dbenford(1:9))
23 }
24 # Benford distribution for two digits
25 dbenford2 ← function(x) {
26   log10(1+1/x)
27 }
28 pbenford2 ← function(q) {
29   cumprobs ← cumsum(dbenford2(10:99))
30   return(cumprobs[q-9])
31 }
32 qbenford2 ← function(p) {
33   cumprobs ← cumsum(dbenford2(10:99))
34   quantiles ← sapply(p, function(x) {90-sum(cumprobs >= x)})
35   return(quantiles+10)

```



```

36 }
37 rbenford2 ← function(n){
38   sample(10:99, size = n, replace = TRUE, prob = dbenford2(10:99))
39 }
40
41 # Inicio da geracao de dados
42 set.seed(2016)
43 # Simulando dados para estimar a variabilidade do simulador
44 # Objetivo: ter uma estimativa da matrix de covariancia da sequencia
    de vetores que aderem ao modelo de Benford
45 # Criando uma matrix para armazenar os resultados da simulacao
46 tamserie ← 500 #Tamanho da serie
47 dados ← matrix(data=0,nrow=1,ncol=9)
48 # Para gerar a base de dados, primeiro temos que gerar um periodo de
    tempo de lancamentos.
49 # Vamos considerar que uma semana possui lancamentos que se
    distribuem como uma normal com media /mu e desvio \sigma
50 for(i in 1:tamserie){
51   NumLanc ← trunc(rnorm(n=1,mean=200,sd=15))+1
52   Lanc ← qbenford(runif(n=NumLanc))
53   newdados ← prop.table(tabulate(Lanc, nbins = 9))
54   dados ← rbind(dados,newdados)
55 }
56 dados ← dados[-1,]
57 # Estimando parametros
58 ebenford.mean ← apply(X=dados, MARGIN=2, FUN=mean)
59 ebenford.mean ← ebenford.mean/sum(ebenford.mean)
60 ebenford.cov ← diag(apply(X=dados, MARGIN=2, FUN=var))
61 # Definindo parametros da simulacao para obtencao dos ARLs
62 delta ← c( 0, 0.5, 1, 1.5, 2, 3, 5) # desvio da
    media inserido na serie para avaliar o ARL
63 nalarmes ← 50000
64 # carrega codigo das cartas qui-quadrado
65 source("quiquadrado.R")
66 source("T2.R")

```

## A.2 Carta $\chi^2$ : Arquivo quiquadrado.R

```

1 # Analise - Carta Qui-quadrado
2 set.seed(42)

```

```

3 i ← 0 # linha da matriz de resultados
4 resultados ← matrix(data=0,nrow=length(delta),ncol=6)
5 for (j in delta){
6   i ← i+1
7   alarmes ← 0 #conta o numero de alarmes
8   tamserie ← 0 #conta o tamanho da serie
9   LC ← qchisq(p=.9973,df=8) #Define o limite de controle
10  ptm ← proc.time() #salva o tempo inicial para o limite de
    processamento
11  tempo ← as.numeric((proc.time() - ptm)[3])
12  quiquadrado ← c(0) #Salva a estatistica qui-quadrado
13  while( (alarmes < nalarmes+1) ){
14    NumLanc ← trunc(rnorm(n=1,mean=200,sd=15))+1 #Define o numero de
    lancamentos
15    Lanc ← qbenford(runif(n=NumLanc)) #Define os valores
16    dados ← prop.table(tabulate(Lanc, nbins = 9))
17    dados[1] ← dados[1] - j*sqrt(diag(ebenford.cov))[1]
18    ifelse(dados[1] < 0.01, dados[1] ← 0.01, dados[1] ← dados
    [1])
19    dados[-1] ← dados[-1]+(1-sum(dados))*dados[-1]/sum(dados
    [-1]) #Calcula o qui-quadrado e testa o alarme
20    nquiquadrado ← t(dados-dbenford(seq(1:9)))*%*%solve(ebenford.
    cov)*%*(dados-dbenford(seq(1:9)))
21    if(nquiquadrado > LC){
22      alarmes ← alarmes + 1
23    }
24    quiquadrado ← c(quiquadrado,nquiquadrado)
25    tamserie ← tamserie + 1
26    tempo ← as.numeric((proc.time() - ptm)[3])
27  }
28  quiquadrado ← quiquadrado[-1]
29  x ← which(quiquadrado > LC)
30  RL ← x[2:length(x)]-x[1:length(x)-1]
31  resultados[i,1] ← j
32  resultados[i,2] ← mean(RL)
33  resultados[i,3] ← sd(RL)
34  resultados[i,4] ← tamserie
35  resultados[i,5] ← alarmes-1
36  resultados[i,6] ← tempo
37 }

```

```

38 cat("Resultados da simulacao para a analise na carta qui-quadrado -
      distribuicao conhecida \n\n")
39 resultados ← as.data.frame(resultados)
40 names(resultados) ← c("delta", "ARL", "sd(RL)", "size", "alarms", "time")
41 print(resultados)
42 write.csv(resultados, "qui-quadrado.csv")

```

### A.3 Carta $T^2$ de Hotteling: Arquivo T2.R

```

1 # Analise - Carta T2
2 set.seed(42)
3 i ← 0 # linha da matriz de resultados
4 resultados ← matrix(data=0, nrow=length(delta), ncol=6)
5 for (j in delta){
6   i ← i+1
7   alarmes ← 0 #conta o numero de alarmes
8   tamserie ← 0 #conta o tamanho da serie
9   LC ← qchisq(p=.9973, df=8) #Define o limite de controle
10  ptm ← proc.time() #salva o tempo inicial para o limite de
    processamento
11  tempo ← as.numeric((proc.time() - ptm)[3])
12  T2 ← c(0) #Salva a estatistica T2
13  while( (alarmes < nalarmes+1) ){
14    NumLanc ← trunc(rnorm(n=1, mean=200, sd=15))+1 #Define o numero de
        lancamentos
15    Lanc ← qbenford(runif(n=NumLanc)) #Define os valores
16    dados ← prop.table(tabulate(Lanc, nbins = 9))
17    dados[1] ← dados[1] - j*sqrt(diag(ebenford.cov))[1]
18    ifelse(dados[1] < 0.01, dados[1] ← 0.01, dados[1] ← dados
        [1])
19    dados[-1] ← dados[-1]+(1-sum(dados))*dados[-1]/sum(dados
        [-1])
20    #Calcula o T2 e testa o alarme
21    nT2 ← t(dados-ebenford.mean)%*%solve(ebenford.cov)%*%(dados-
        ebenford.mean)
22    # T2 calculado com vetor de medias estimadas
23    if(nT2 > LC){
24      alarmes ← alarmes + 1
25    }
26    T2 ← c(T2, nT2)

```

```
27     tamserie ← tamserie + 1
28     tempo ← as.numeric((proc.time() - ptm)[3])
29     }
30 T2 ← T2[-1]
31 x ← which(T2 > LC)
32 RL ← x[2:length(x)]-x[1:length(x)-1]
33 resultados[i,1] ← j
34 resultados[i,2] ← mean(RL)
35 resultados[i,3] ← sd(RL)
36 resultados[i,4] ← tamserie
37 resultados[i,5] ← alarmes-1
38 resultados[i,6] ← tempo
39 }
40 cat("Resultados da simulacao para a analise na carta T2 -
     distribuicao conhecida \n\n")
41 resultados ← as.data.frame(resultados)
42 names(resultados) ← c("delta", "ARL", "sd(RL)", "size", "alarms", "time")
43 print(resultados)
44 write.csv(resultados, "T2.csv")
```