



XX PRÊMIO TESOIRO NACIONAL – 2015

TEMA: TÓPICOS ESPECIAIS

**AUDITORIA DOS GASTOS DIRETOS DO GOVERNO FEDERAL
USANDO *BOOTSTRAP* E A LEI DE NEWCOMB-BENFORD**

SUMÁRIO

1. INTRODUÇÃO.....	3
1.1 Motivação e Objetivos	3
1.2 Metodologia	4
1.3 Resumo	5
2. BREVE REVISÃO DA LITERATURA.....	6
2.1 Despesa Pública.....	6
2.2 Auditoria	7
2.3 Contabilometria	8
2.4 A Lei de Newcomb-Benford como Método Contabilométrico	9
2.5 A Lei de Newcomb-Benford no Brasil	11
3. DISTRIBUIÇÃO DE NEWCOMB-BENFORD.....	12
3.1 Conceito	12
3.2 Testes de Hipótese.....	16
3.2.1 Z-Teste	17
3.2.2 χ^2 -Teste.....	19
3.2.3 Teste de Kolmogorov-Smirnoff (K-S Teste).....	20
3.2.4 Críticas ao Excesso de Poder	21
3.2.5 <i>Bootstrap</i>	22
4. <i>BOOTSTRAP</i> APLICADO À LEI DE NEWCOMB-BENFORD.....	25
4.1 Metodologia <i>Bootstrap</i>	26
4.1.1 Cálculo dos Parâmetros da Distribuição de Newcomb-Benford.....	26
4.1.2 Cálculo dos Intervalos de Confiança (ICs) dos Parâmetros Exatos	28
4.1.3 Cálculo dos Intervalos de Confiança (ICs) das Estatísticas Reais	32
4.1.4 Critérios de Conformidade	32
4.2 Regressão <i>Bootstrap</i>	35
4.3 Cálculo de ICs para a FDA via <i>Bootstrap</i>	38
5. ANÁLISE DOS GASTOS DIRETOS DO GOVERNO FEDERAL.....	41
5.1 Metodologia	41
5.2 Base de Dados	42
5.3 Escopo	42
5.4 Análise de Conformidade.....	45

5.5	Comparação dos Resultados dos Testes	61
5.6	Resultados dos Testes vs Razão Entre Áreas da <i>fda</i>	62
6.	TOMADAS DE CONTAS ESPECIAIS (TCEs)	65
6.1	Conceito	65
6.2	Base de Dados	67
6.3	Análise de Desempenho das Predições	68
7.	CONSIDERAÇÕES FINAIS	72
8.	REFERÊNCIAS	77

1. INTRODUÇÃO

1.1 Motivação e Objetivos

A auditoria tem papel fundamental no julgamento da qualidade e da veracidade dos registros contábeis. No âmbito do setor público, a auditoria externa realizada por órgãos competentes é de vital importância para garantir o adequado emprego dos recursos públicos. Os avanços recentes no poder computacional começam a viabilizar a auditoria contínua, produzindo resultados simultâneos ou com pouco atraso em relação aos eventos auditados.

A contabilometria, introduzida pelo professor Sérgio de Ludícibus (1982), é uma disciplina de aplicação de métodos quantitativos na solução de problemas contábeis e, no contexto desenhado, desempenha um importante papel na análise tempestiva dos dados. Dentre os métodos contabilométricos, destaca-se a Lei de Newcomb-Benford, introduzida na contabilidade por Carslaw (1988) e utilizada pela primeira vez na detecção de fraudes através de testes de hipótese por Nigrini (1996). No Brasil, o modelo proposto por Nigrini (1996) foi introduzido por Santos et al. (2003), tornando-se referência na literatura nacional, e sendo utilizado por diversos autores (Santos et al., 2009; Costa et al., 2011; Cunha, 2013).

Entretanto, os testes sobre os quais o modelo se apóia foram criticados pelo seu “excesso de poder” (Nigrini, 2000; Geyer e Williamson, 2004; Cho e Gaines, 2007), isto é, pela sensibilidade excessiva e conseqüente sinalização de falsos positivos. Recentemente, Suh e Headrick (2010), assim como Suh, Headrick e Minaburo (2011), propuseram a aplicação de procedimentos de *bootstrap* (baseados na simulação

computacional) sobre a Lei de Newcomb-Benford, a fim de contornar as limitações dos testes tradicionais. No Brasil, contudo, não temos conhecimento da aplicação desses procedimentos no contexto da auditoria do setor público.

O presente trabalho visa preencher essa lacuna, tendo como objetivo principal aplicar as novas técnicas sobre um amplo conjunto de dados nacionais, qual seja, os gastos diretos do Governo Federal no período de 2011 a 2014 cujos elementos de classificação da despesa são passíveis de análise de conformidade com a Lei de Newcomb-Benford, totalizando um universo de mais de 17 milhões de despesas distribuídas ao longo de 26 órgãos e 8 elementos de despesa.

Além disso, o trabalho também tem por objetivo realizar uma análise de desempenho das predições obtidas, tomando como parâmetro de comparação as Tomadas de Contas Especiais instauradas pela Controladoria Geral da União.

Por fim, esse trabalho também propõe uma abordagem experimental para o teste de conformidade com a Lei de Newcomb-Benford baseado na estimação de intervalos de confiança para a função de distribuição acumulada dos dois primeiros dígitos através de um procedimento de *bootstrap*.

1.2 Metodologia

Para a consecução desse trabalho, coletamos os dados relativos aos gastos diretos do Governo Federal entre os anos de 2011 e 2014 e identificamos o órgão, a classificação por elemento, e os dois primeiros dígitos de cada uma das despesas realizadas. Então, selecionamos um subconjunto de elementos de despesa cujas

características são passíveis de análise de conformidade com a Lei de Newcomb-Benford, de acordo com critérios que serão detalhados oportunamente.

Sobre esses dados, aplicamos uma sequência de testes, incluindo tanto aqueles do modelo contabilométrico tradicional, bem como aqueles baseados em modernas técnicas de *bootstrap*. Apresentamos os resultados obtidos e, adicionalmente, calculamos uma métrica relacionada à abordagem experimental para o teste de conformidade proposta nesse trabalho.

Finalmente, coletamos também os dados relativos à instauração de Tomadas de Contas Especiais para o mesmo período e realizamos uma análise de desempenho das predições obtidas, calculando a acurácia, sensibilidade e especificidade das mesmas.

1.3 Resumo

No Capítulo 2, apresentamos uma breve revisão da literatura, elencando os conceitos de despesa pública, auditoria e contabilometria, com destaque para a aplicação da Lei de Newcomb-Benford como método contabilométrico de auditoria contábil.

No Capítulo 3, introduzimos a distribuição de Newcomb-Benford e sua origem, descrevendo os tradicionais testes de hipótese à ela associados, assim como as críticas à tais testes. Além disso, apresentamos a técnica *bootstrap* sobre a qual são baseados novos testes que se propõe a superar as limitações dos testes tradicionais. O Capítulo 4, então, detalha a aplicação dessa técnica no contexto da Lei de Newcomb-Benford, através da estimação de intervalos de confiança para os parâmetros populacionais da distribuição e para os coeficientes de regressão entre os dígitos. Além disso, também

apresenta os critérios para análise de conformidade e detalha a abordagem experimental para o teste de conformidade proposta nesse trabalho.

No Capítulo 5, realizamos a análise propriamente dita dos gastos diretos do Governo Federal, aplicando os testes apresentados sobre as despesas ao longo de cada ano, agrupadas por órgão e elemento de despesa e apresentando os resultados, tanto de forma extensiva quanto esquemática. Em seguida, comparamos os resultados dos diferentes testes e os relacionamos com aqueles obtidos a partir de nossa abordagem experimental.

No Capítulo 6, por sua vez, introduzimos as Tomadas de Contas Especiais, instrumento legal de apuração de responsabilidade por danos causados ao Erário, cujos processos de instauração servem de base para uma análise de desempenho das predições obtidas no capítulo anterior, e então apresentamos os resultados dessa análise.

2. BREVE REVISÃO DA LITERATURA

2.1 Despesa Pública

A despesa pública é dada pelo total dos gastos realizados pela Administração Pública visando o atendimento das necessidades, econômicas e sociais, da coletividade e o cumprimento das responsabilidades institucionais do setor público. O orçamento público, por sua vez, surge da necessidade de se organizar e controlar as receitas e despesas do Estado, e se justifica pela exigência de previsibilidade das ações estatais em prol da sociedade. No Brasil, a Lei 4.320/64 definiu princípios

orçamentários e o atual modelo de elaboração da proposta orçamentária. Dez anos depois, em 1974, foi instituída a classificação funcional-programática, pela a Portaria nº 9. A Portaria nº 42/1999 do MPOG atualizou os anexos da Lei 4.320/1964, substituindo a classificação funcional programática e estabelecendo a classificação das despesas por função e subfunção, vigente até os dias de hoje e de observância obrigatória por parte da União, Estados e Municípios. Por sua vez, a Portaria Interministerial nº 163/2001 atualizou o Anexo 4 da Lei 4.320/1964, substituindo a sua estrutura de classificação. Ficou estabelecido que, na classificação da despesa segundo sua natureza, o 1º dígito representa a categoria econômica, o 2º dígito, o grupo de despesa, o 3º e 4º dígitos, a modalidade de aplicação e, o 5º e 6º dígitos, o elemento de despesa. No presente trabalho, utilizamos o elemento de despesa para agrupar os valores a serem analisados.

2.2 Auditoria

A auditoria é a função organizacional de revisão, avaliação e emissão de pareceres relativos ao ciclo administrativo e atua sobre as informações contábeis das organizações. O processo de auditoria visa identificar o estado de consistência e exatidão dos dados, relatando os erros e inconsistências encontrados. Seu papel é complementar àquele da contabilidade, podendo ser entendida como uma técnica que visa obter elementos de convicção para o julgamento da qualidade e da veracidade dos registros contábeis, tendo como finalidade principal a confirmação desses (Franco e Marra, 1991). A auditoria pode ser classificada de acordo com sua forma de intervenção: interna ou externa. O foco do presente trabalho é na auditoria externa, aquela realizada pelas instituições ou órgãos reguladores e de fiscalização.

A auditoria contínua, por sua vez, caracteriza-se por produzir resultados simultâneos, ou com pouco atraso em relação ao evento sendo auditado (Vasarhelyi e Halper, 1991). Essa tempestividade, contudo, é dependente da existência de sistemas automatizados operando sobre dados armazenados eletronicamente. Em algumas áreas, como a detecção de fraudes em transações realizadas através de cartões de crédito, isso já é uma realidade. No setor público, infelizmente, o estado atual encontra-se bastante distante desse ideal. Nesse contexto, ressaltamos a importância da contabilometria enquanto uma disciplina de aplicação de métodos quantitativos sobre dados contábeis.

2.3 Contabilometria

Em 1982, a Revista Brasileira de Contabilidade publicou um artigo intitulado “Existirá a Contabilometria?”, de autoria do professor Sérgio de Ludícibus, onde a contabilometria foi definida como uma disciplina de aplicação de métodos quantitativos na solução de problemas contábeis, à semelhança do que ocorre na Econometria. Atualmente, a disciplina trata da mensuração de relações econômicas e financeiras a fim de detectar desvios ou distorções. De acordo com Marion e Silva (1986), a metodologia de trabalho da contabilometria consiste de diversas etapas: a) formulação da teoria; b) especificação do modelo matemático ou teórico (modelo real ou determinista); c) especificação do modelo contabilométrico da teoria (modelo probabilístico); d) hipóteses básicas; e) obtenção dos dados; f) estimativas dos parâmetros do modelo contabilométrico; g) teste de hipótese; h) previsão e i) utilização do modelo para fins de controle ou política financeira. Para Santos et al (2005), a contabilometria é uma

metodologia científica para a auditoria contábil, nos casos em que a análise qualitativa não assegura ao auditor uma tomada de decisão precisa.

Em um ambiente de rápido desenvolvimento tecnológico, onde a quantidade de informações disponíveis eletronicamente cresce em ritmo acelerado, inclusive no setor público brasileiro, a contabilometria se destaca como uma importante disciplina para a análise tempestiva desses dados. Para os órgãos de controle externo, a utilização de métodos contabilométricos aplicados à auditoria do setor público pode fornecer valiosas informações a respeito de pontos relevantes a serem auditados, tornando-se um importante subsídio para decisões relativas à focalização dos esforços de suas equipes de auditoria.

2.4 A Lei de Newcomb-Benford como Método Contabilométrico

Dentre os métodos contabilométricos utilizados para a detecção de erros ou fraudes contábeis, inclui-se a chamada Lei de Newcomb-Benford. Essa lei afirma que os dígitos em uma sequência de números que ocorre naturalmente não aparecem com iguais proporções entre si, como se poderia supor inicialmente. Segundo a Lei de Newcomb-Benford, os dígitos menores aparecem com maior probabilidade, obedecendo à uma razão logarítmica. No próximo capítulo, apresentamos a formulação da Lei de Newcomb-Benford em detalhes.

Esse fenômeno foi identificado inicialmente por Simon Newcomb (1881), astrônomo e matemático, que observou que as tábuas de logaritmos apresentavam maior desgaste nas primeiras páginas, denotando uma maior utilização dos logaritmos de números começando com 1 do que daqueles começando com 9. Em 1938, Frank

Benford redescobriu esse fenômeno de forma independente e produziu uma extensa análise a seu respeito, englobando mais de 20 mil conjuntos de dados de fontes diversas.

A intuição por trás da Lei de Newcomb-Benford é baseada na noção que, dada uma sequência de eventos que provoque a alteração dos valores observados, levará mais tempo para que um valor pequeno, por exemplo, igual a 1, cresça 100% e se iguale a 2, do que um valor maior, por exemplo, igual a 2, cresça 50% e se iguale a 3. Logo, se observarmos esses valores de maneira periódica, os dígitos associados aos valores menores aparecerão mais frequentemente que aqueles associados aos valores maiores.

Dessa forma, assumindo-se que a Lei de Newcomb-Benford governa a distribuição natural dos dígitos que compõem os números, eventuais desvios em relação às frequências previstas seriam indicativos de interferência do ser humano (Diniz, Corrar e Slomski, 2010). Entretanto, nem todos os desvios são, necessariamente, fruto de fraudes ou erros: fatores normativos ou legais, onde valores são deliberadamente definidos (por exemplo, números de telefone) podem se situar na origem desses desvios.

Carshaw (1988) foi o primeiro autor a aplicar a Lei de Newcomb-Benford na contabilidade. Em seu trabalho, o autor levantou a hipótese de que a distribuição de frequência para os segundos dígitos dos lucros de empresas da Nova Zelândia se afastava significativamente daquela preconizada por Newcomb-Benford. Além disso, Carshaw (1988) introduziu a utilização do z-teste e do χ^2 -teste para as diferenças entre

as probabilidades esperadas e observadas para os primeiros e segundos dígitos, separadamente.

Nigrini (1996) foi o primeiro autor a utilizar a Lei de Newcomb-Benford na identificação de sonegação de impostos e na detecção de fraudes, desenvolvendo um modelo contabilométrico baseado em testes de hipóteses. A partir da premissa que os indivíduos tendem a repetir suas ações e que são incapazes de forjar números em conformidade com a Lei de Newcomb-Benford, o autor investigou os dois primeiros dígitos de cheques fraudulentos e verificou que, de fato, esses números manipulados eram improváveis e incompatíveis com a distribuição esperada.

2.5A Lei de Newcomb-Benford no Brasil

No Brasil, o modelo contabilométrico nos moldes utilizados por Carslaw (1988) e Nigrini (1996), isto é, baseado na Lei de Newcomb-Benford e nos testes de hipótese (z-teste e do χ^2 -teste), foi introduzido por Santos et al. (2003) no contexto de auditoria do setor público. Nesse trabalho, os autores tinham por objetivo identificar desvios na distribuição do primeiro dígito a partir dos valores das notas de empenho de despesas efetuadas por 20 municípios paraibanos: em apenas três deles a distribuição dos dígitos se encontrava em conformidade, dado um nível de significância de 10%.

O baixo nível de conformidade estimado por Santos et al. (2003) em seu trabalho pode ser entendido como um indicativo de ampla manipulação dos valores das notas de empenho avaliadas. Entretanto, diversos estudos (Cho e Gaines, 2007; Geyer e Williamson, 2004; Nigrini, 2000) apontam para problemas de “excesso de poder” dos testes de hipótese propostos por Carslaw (1988), de forma que esses testes seriam

propensos a falsos positivos se utilizados em conjuntos de dados relativamente grandes. Nesse caso, o baixo nível de conformidade estimado seria apenas o resultado do “excesso de poder” dos testes de hipótese utilizados pelos autores.

Embora o trabalho de Santos et al. (2003) tenha se tornado uma referência na literatura brasileira, o modelo contabilométrico baseado na Lei de Newcomb-Benford ainda está longe de ser adotado como ferramenta padrão de auditoria no setor público brasileiro. Dentre os trabalhos realizados no Brasil e baseados no mesmo modelo contabilométrico, podemos citar: Santos et al. (2009), aplicado ao trabalho de auditoria tributária do Imposto Sobre Serviços de Qualquer Natureza (ISSQN); Costa et al. (2011), aplicado às notas de empenho de 20 unidades gestoras de dois Estados brasileiros; e Cunha (2013), aplicado às planilhas orçamentárias de obras públicas.

3. DISTRIBUIÇÃO DE NEWCOMB-BENFORD

3.1 Conceito

Benford (1938) verificou que, em diversos casos, as probabilidades associadas aos primeiros nove dígitos seguiam uma distribuição logarítmica. Então, a partir de extensos estudos em mais de 20 mil conjuntos de dados distintos, o autor formalizou suas observações em fórmulas para as distribuições tanto dos primeiros (d_1) como dos segundos (d_2) dígitos, bem como para sua distribuição conjunta (d_1d_2), as quais apresentamos a seguir.

A distribuição exata dos primeiros dígitos obedece à seguinte fórmula:

$$P(d_1) = \log_{10} \left(1 + \frac{1}{d_1} \right) \quad d_1 = 1 \dots 9$$

Dessa forma, a tabela abaixo apresenta as probabilidades associadas à cada um dos 9 primeiros dígitos possíveis:

Tabela 3.1: Distribuição de Probabilidade dos Primeiros Dígitos

	Primeiros Dígitos (d_1)								
	1	2	3	4	5	6	7	8	9
$P(d_1)$	0,301	0,176	0,125	0,097	0,079	0,067	0,058	0,051	0,046

Fonte: Elaborado pelo autor

A distribuição exata dos segundos dígitos, por sua vez, obedece à seguinte fórmula:

$$P(d_2) = \sum_{d_1=1}^9 \log_{10} \left(1 + \frac{1}{d_1 d_2} \right) \quad d_2 = 0 \dots 9$$

Dessa forma, a tabela abaixo apresenta as probabilidades associadas à cada um dos 10 segundos dígitos possíveis:

Tabela 3.2: Distribuição de Probabilidade dos Primeiros Dígitos

	Segundos Dígitos (d_2)									
	0	1	2	3	4	5	6	7	8	9
$P(d_2)$	0,120	0,114	0,109	0,104	0,100	0,097	0,093	0,090	0,088	0,085

Fonte: Elaborado pelo autor

Podemos observar que, ao considerarmos os segundos dígitos, as probabilidades associadas aos mesmos tornam-se menos distintas entre si, em comparação àquelas calculadas para os primeiros dígitos. Ao considerarmos o terceiro, quarto e quinto dígitos, as probabilidades se aproximam cada vez mais de uma distribuição uniforme e, portanto, perdem a interessante propriedade exibida mais nitidamente pelos dois primeiros dígitos, qual seja, sua diferença de probabilidades.

A distribuição conjunta exata dos dois primeiros dígitos obedece à seguinte fórmula:

$$P(d_1d_2) = \log_{10}(1 + 1/d_1d_2) \quad d_1d_2 = 10 \dots 99$$

Dessa forma, a tabela abaixo apresenta as probabilidades associadas à cada uma das 90 combinações possíveis dos dois primeiros dígitos:

Tabela 3.3: Distribuição Conjunta dos Dois Primeiros Dígitos

Dois Primeiros Dígitos (d_1d_2)									
d_1d_2	$P(d_1d_2)$	d_1d_2	$P(d_1d_2)$	d_1d_2	$P(d_1d_2)$	d_1d_2	$P(d_1d_2)$	d_1d_2	$P(d_1d_2)$
10	0,041393	28	0,015240	46	0,009340	64	0,006733	82	0,005264
11	0,037789	29	0,014723	47	0,009143	65	0,006631	83	0,005201
12	0,034762	30	0,014240	48	0,008955	66	0,006531	84	0,005140
13	0,032185	31	0,013788	49	0,008774	67	0,006434	85	0,005080
14	0,029963	32	0,013364	50	0,008600	68	0,006340	86	0,005021
15	0,028029	33	0,012965	51	0,008433	69	0,006249	87	0,004963
16	0,026329	34	0,012589	52	0,008273	70	0,006160	88	0,004907
17	0,024824	35	0,012234	53	0,008118	71	0,006074	89	0,004853
18	0,023481	36	0,011899	54	0,007969	72	0,005990	90	0,004799
19	0,022276	37	0,011582	55	0,007825	73	0,005909	91	0,004746
20	0,021189	38	0,011281	56	0,007687	74	0,005830	92	0,004695
21	0,020203	39	0,010995	57	0,007553	75	0,005752	93	0,004645
22	0,019305	40	0,010724	58	0,007424	76	0,005677	94	0,004596
23	0,018483	41	0,010465	59	0,007299	77	0,005604	95	0,004548
24	0,017729	42	0,010219	60	0,007179	78	0,005532	96	0,004501
25	0,017033	43	0,009984	61	0,007062	79	0,005463	97	0,004454
26	0,016390	44	0,009760	62	0,006949	80	0,005395	98	0,004409
27	0,015794	45	0,009545	63	0,006839	81	0,005329	99	0,004365

Fonte: Elaborado pelo autor

A distribuição conjunta dos dois primeiros dígitos é de particular interesse na detecção de erros ou fraudes, pois reduz o escopo de valores “suspeitos” àqueles iniciando-se com determinados dois dígitos na sequência. Podemos notar, ainda, que essa distribuição também exibe uma grande diferença entre as probabilidades calculadas para cada combinação de dois dígitos, variando mais de 9 vezes entre seus extremos.

Uma das principais características da distribuição de Newcomb-Benford é a invariância de escala, conforme estabelecido por Pinkham (1961). Essa propriedade, única entre as distribuições, determina que, se um dado conjunto de números segue a distribuição de Newcomb-Benford, então o resultado da multiplicação de todos seus elementos por uma constante diferente de zero segue a mesma distribuição. Por consequência, pode-se escolher livremente a unidade de medida para representar um conjunto de dados para o qual desejamos testar se obedece à distribuição.

Mais recentemente, Hill (1996) demonstrou que a distribuição da distribuição de números aleatórios, calculada a partir de amostras aleatórias provenientes de diferentes distribuições é, de fato, a distribuição de Newcomb-Benford.

Entretanto, nem todo conjunto de dados é constituído de dígitos que obedecem à distribuição de Newcomb-Benford. Para tal, é necessário que alguns requisitos sejam preenchidos, quais sejam:

- a) os dados devem conter valores relativos à fenômenos similares, isto é, não podem conter, simultaneamente, por exemplo, valores do censo populacional e de notas de empenho;
- b) não deve haver valores mínimos ou máximos pré-definidos, isto é, sem valores iniciais arbitrários ou pontos de corte;
- c) os valores não devem ser arbitrariamente definidos, como por exemplo, números de telefone; e
- d) os dados devem conter uma quantidade maior de valores pequenos do que de valores grandes.

No âmbito da auditoria no setor público, as notas de empenho preenchem os quatro requisitos e, portanto, qualificam-se como um conjunto de dados válido para análise contabilométrica segundo a Lei de Newcomb-Benford.

3.2 Testes de Hipótese

A razão para a realização de um teste de hipótese é averiguar a significância estatística dos resultados obtidos. Ao formular uma hipótese para uma dada característica de uma população, testamos essa característica para uma amostra dela retirada. Se as diferenças observadas podem ser consideradas decorrentes de flutuações normais, dado um nível de significância previamente escolhido, os valores encontrados não são estatisticamente significativos e, portanto, concluímos que a amostra pertence à população de origem. Por outro lado, se tais diferenças são devidas à algum efeito real, e não ao acaso, os valores encontrados são estatisticamente significativos e, portanto, concluímos que a amostra não pertence à população de origem.

O nível de significância é o risco que se corre ao rejeitar a hipótese nula, dado que a mesma seja verdadeira (erro do tipo I) e geralmente é definido no intervalo entre 1% e 10%. No caso da Lei de Newcomb-Benford, a hipótese nula a ser testada é que o conjunto de dados considerado obedece à lei, isto é, que esse conjunto é uma amostra retirada da distribuição de Newcomb-Benford. Logo, o nível de significância adotado representa o risco de um falso positivo na detecção de erros ou fraudes, isto é, o risco de concluir que um conjunto de dados não obedece à lei quando, de fato, obedece.

Por outro lado, o poder do teste é a probabilidade de rejeitar a hipótese nula dado que a mesma seja, de fato, falsa. Quanto maior o poder do teste, portanto, menor a chance de um falso negativo, isto é, de não rejeitar a hipótese nula dado que a mesma seja falsa (erro tipo II). No caso da Lei de Newcomb-Benford, um falso negativo equivale a afirmar que não há erro ou fraude quando, de fato, há.

Ora, no âmbito da auditoria no setor público, é evidente a preferência pela abordagem prudencial: um falso positivo é preferível a um falso negativo, de forma que um erro ou fraude não passe despercebido pelos órgãos de controle. Para tanto, é preciso aumentar o poder do teste. Entretanto, a redução dos falsos negativos vem à custa do aumento dos falsos positivos. No limite, todos os casos seriam considerados potencialmente fraudulentos e a taxa de falsos negativos seria zero. Os falsos positivos, contudo, constituiriam a esmagadora maioria, assumindo que erros ou fraudes estão presentes apenas em um pequeno número de casos. O custo do falso positivo, portanto, é aquele de uma investigação desnecessária.

Alguns testes estatísticos, como o Z-Teste, quando aplicados em grandes amostras, podem exibir um “excesso de poder”, isto é, tornar-se excessivamente sensíveis à pequenos desvios e, portanto, propensos a rejeitar a hipótese nula quando a mesma é verdadeira, produzindo falsos positivos. No contexto da auditoria no setor público, ao analisarmos milhares de notas de empenho, devemos estar atentos à esse problema.

3.2.1 Z-Teste

O Z-Teste é utilizado como parâmetro de aceitação ou rejeição da hipótese nula (ausência de diferença) para a diferença entre a proporção observada na amostra (p_o) e

a proporção da população específica (p_e) quando a distribuição de amostragens desta proporção segue, aproximadamente, a distribuição normal. O teste foi utilizado por Carslaw (1988) como medida de significância estatística para a diferença entre a proporção observada na amostra e aquela da distribuição de Newcomb-Benford. Assim, o Z-Teste pode ser definido da seguinte forma:

$$\begin{cases} H_0: p_o = p_e \\ H_1: p_o \neq p_e \end{cases} \quad Z = \frac{|p_o - p_e| - \frac{1}{2n}}{\sqrt{\frac{p_e(1-p_e)}{n}}}$$

Onde n é o tamanho da amostra e $1/2n$ é o termo de correção de continuidade, utilizado apenas quando seu valor é inferior a $|p_o - p_e|$. Podemos notar que, tudo o mais constante, a estatística do teste aumenta conforme a raiz quadrada do tamanho da amostra, de forma que o Z-Teste é suscetível ao já mencionado problema de “excesso de poder” (Nigrini e Mittermaier, 1997; Nigrini, 2000). Assim, são consideradas aceitáveis até cinco rejeições da hipótese nula ao realizar o Z-Teste sobre os dois primeiros dígitos, dado um nível de significância de 5% (Nigrini, 2012).

A rejeição da hipótese nula ocorre quando o valor absoluto da estatística de teste Z supera o Z crítico para um dado nível de significância. Em um teste tradicional, para um nível de significância de 5%, isso equivale a rejeitar a hipótese nula para valores da estatística de teste, em módulo, superiores a 1,96. Entretanto, no contexto da distribuição de Newcomb-Benford, devemos levar em consideração que são realizados múltiplos testes (um para cada dígito ou par de dígitos), tornando-se necessário utilizar um nível de significância reduzido em cada um desses testes, de acordo com a

desigualdade de Bonferroni¹ (Hogg, McKean e Craig, 2005), para que a probabilidade total de rejeição se iguale ao nível de significância desejado. Assim, em cada teste, o nível de significância reduzido é obtido pela simples divisão do nível de significância desejado pela quantidade total de testes realizados (nove, no caso do teste do primeiro dígito; 90 no caso do teste dos dois primeiros dígitos). Por exemplo, para o teste do primeiro dígito a um nível de significância de 5%, realiza-se cada um dos nove testes ao nível de significância reduzido de 0,5556% (= 5% / 9), correspondendo à rejeição da hipótese nula para valores da estatística de teste, em módulo, superiores a 2,77.

3.2.2 χ^2 -Teste

Diferentemente do Z-Teste, que avalia localmente os desvios em relação à probabilidade esperada, o χ^2 -Teste é um teste global, avaliando os desvios em todos os dígitos simultaneamente e, portanto, medindo o grau de conformidade da distribuição observada (p_o) com a esperada (p_e) como um todo. O χ^2 -Teste é unilateral por natureza e pode ser definido da seguinte forma:

$$\begin{cases} H_0: p_o = p_e \\ H_1: p_o > p_e \end{cases} \quad \chi^2 = n \sum_{d=10}^{99} \frac{(p_o - p_e)^2}{p_e}$$

Onde n é o tamanho da amostra e p_o e p_e são, respectivamente, as probabilidades observada e esperada, para cada combinação dos dois primeiros dígitos (10 a 99). Assim, comparamos a estatística de teste obtida com o valor crítico da distribuição χ^2 com 89 graus de liberdade (considerando as 90 combinações) para um dado nível de significância. Caso o valor da estatística de teste encontrado supere o valor crítico,

¹ A desigualdade de Bonferroni é uma regra de probabilidade muito útil que estabelece que $P(a_1 a_2 \dots a_n) \geq P(a_1) + P(a_2) + \dots + P(a_n) - n + 1$. Essa desigualdade produz um intervalo de confiança sem o pressuposto de independência dos parâmetros.

rejeitamos a hipótese nula e concluímos que as probabilidades observadas não estão em conformidade com aquelas da distribuição de Newcomb-Benford, sugerindo a existência de erro ou fraude.

Da mesma forma que no Z-Teste, ao considerarmos tudo o mais constante, a estatística do teste aumenta conforme o tamanho da amostra, de forma que o χ^2 -Teste também é suscetível ao já mencionado problema de “excesso de poder” (Nigrini e Mittermaier, 1997; Nigrini, 2000).

3.2.3 Teste de Kolmogorov-Smirnoff (K-S Teste)

O Teste de Kolmogorov-Smirnoff, assim como o χ^2 -Teste, é um teste global e baseia-se na função densidade acumulada (fda), isto é, na soma cumulativa das probabilidades associadas aos dois primeiros dígitos, de 10 a 99. A estatística de teste é dada pelo maior valor absoluto das diferenças entre a função densidade acumulada observada (fda_o) e a função densidade acumulada esperada (fda_e), apurado em cada ponto da distribuição, isto é, de 10 a 99. Essa diferença, também chamada de “supremo” na estatística, é dada pela maior distância horizontal entre as linhas que representam as funções de densidade acumulada das distribuições sendo comparadas. Podemos expressá-la da seguinte forma:

$$\begin{cases} H_0: p_o = p_e \\ H_1: p_o > p_e \end{cases} \quad K - S = \max\{|fda_o - fda_e|_{d=10\dots99}\}$$

Onde p_o e p_e são, respectivamente, as probabilidades observada e esperada, para cada combinação dos dois primeiros dígitos (10 a 99) e fda_o e fda_e são suas respectivas funções de densidade acumulada.

O valor crítico, para um nível de significância de 5% e uma amostra de tamanho n é dado pela expressão:

$$K - S = \frac{1,36}{\sqrt{n}}$$

Caso o valor da estatística de teste encontrado supere o valor crítico, rejeitamos a hipótese nula e concluímos que as probabilidades observadas não estão em conformidade com aquelas da distribuição de Newcomb-Benford, sugerindo a existência de erro ou fraude. No K-S Teste, podemos verificar que o valor crítico diminui com a raiz quadrada do tamanho da amostra de forma que o teste também padece do problema de “excesso de poder” que os dois testes já apresentados.

3.2.4 Críticas ao Excesso de Poder

Retomamos aqui a problemática do “excesso de poder” do Z-Teste e do χ^2 -Teste, quando aplicados a grandes conjuntos de dados. Diversos estudos alertam para esse problema: Nigrini (2000) define essa situação como “excesso de poder”, termo que utilizamos no presente trabalho; Geyer e Williamson (2004) chamam a atenção para o fato que, quando a amostra é muito grande, é quase certo que o Z-Teste rejeitará a hipótese nula para um dado nível de significância; enquanto Cho e Gaines (2007) atentam para o fato que o χ^2 -Teste tem um enorme poder para amostras grandes, de forma que mesmo pequenas diferenças podem ser consideradas estatisticamente significativas. As implicações práticas dessas observações no contexto da auditoria no setor público já foram discutidas anteriormente nessa Seção.

Em face dessas limitações, Suh e Headrick (2010) propuseram um procedimento de *bootstrap* para determinar a conformidade de uma dada amostra à distribuição de Newcomb-Benford através da obtenção de intervalos de confiança para a média (μ), variância (σ^2), assimetria (γ) e curtose (δ) relativos aos dois primeiros dígitos, considerados separada e conjuntamente, e para o coeficiente de correlação de Pearson entre os dois dígitos. Assim, diferentemente dos testes tradicionais, cuja análise das diferenças é realizada dígito a dígito, o método proposto pelos autores se propõe a realizar uma análise verdadeiramente global. No próximo capítulo apresentamos em detalhes a metodologia desenvolvida pelos autores, a qual utilizamos no presente trabalho. Entretanto, primeiramente, convém introduzir o conceito de *bootstrap*.

3.2.5 Bootstrap

O *bootstrap* é um procedimento estatístico não-paramétrico que permite realizar inferências a respeito de um dado parâmetro, como a média populacional, sem a necessidade de pressupostos a respeito do tipo de distribuição associada à estimativa do parâmetro. Enquanto em testes tradicionais, como o Z-Teste, se assume um pressuposto a respeito da forma da distribuição amostral, no caso, o de uma distribuição normal, no *bootstrap* a distribuição amostral de um parâmetro pode ser entendida como a distribuição das estimativas amostrais desse parâmetro calculadas a partir de todas as amostras possíveis de um mesmo tamanho.

Se soubermos a distribuição das estimativas obtida através da repetição de um experimento, podemos calcular os erros-padrão e os intervalos de confiança. Nos testes tradicionais, como assumimos um pressuposto a respeito dessa distribuição, esses

valores são obtidos trivialmente. Entretanto, isso não se aplica aos casos onde a distribuição é desconhecida. Nesses casos, é necessário recorrer a dados simulados que, por hipótese, possuem a mesma distribuição dos dados reais e, então, aplicar o estimador a esses dados simulados, correspondendo ao sorteio de um elemento da distribuição amostral. Ao repetir esse procedimento muitas vezes, obtemos a distribuição amostral desejada.

O fato do procedimento proposto utilizar os próprios dados para determinar seu próprio grau de incerteza fez com que seu autor, Efron (1979), o chamasse de “*bootstrapping*”, em uma referência ao Barão de Munchhausen, cujo plano para sair de um pantâno consistia em puxar a si mesmo pra fora pelos seus cadarços (*bootstraps*, em inglês).

O procedimento do *bootstrap* consiste, portanto, no sorteio de uma amostra aleatória de N elementos, com reposição, a partir de um conjunto de dados de tamanho N , onde todos os elementos são equiprováveis. Dessa forma, a nova amostra possui o mesmo tamanho do conjunto original (N), mas seus elementos podem diferir, devido à amostragem com reposição. Para essa nova amostra, calcula-se a estatística amostral desejada. Repetindo esse processo T vezes, obtemos um conjunto de T estatísticas amostrais calculadas que constituem a distribuição das estimativas para essa estatística. Ordenando os valores, do mínimo ao máximo, os intervalos de confiança são obtidos, pelo método do percentil, observando-se os valores do $\alpha N/2$ -ésimo e do $(1 - \alpha/2)N$ -ésimo elementos, onde α é o nível de significância adotado.

Por exemplo, se temos um conjunto de 5.000 notas de empenho ($N = 5.000$) e desejamos obter a distribuição para a média dos seus primeiros dígitos, iniciamos gerando uma nova amostra, com reposição, de também 5.000 elementos e calculamos a média amostral dos seus primeiros dígitos. Repetimos esse procedimento 25.000 vezes ($T = 25.000$), resultando em 25.000 médias amostrais calculadas. Ordenando essas médias, e assumindo um nível de significância (α) igual a 1%, o intervalo de confiança será dado pelo 125º e 24.875º elementos, uma vez que $\alpha N/2 = 0,01 \cdot 25000/2 = 125$ e $(1 - \alpha/2)N = (1 - 0,01/2)25000 = 24875$. Assim, construímos um intervalo de confiança de 99% para a média amostral dos primeiros dígitos das 5.000 notas de empenho através do procedimento de *bootstrap*.

Embora o método do percentil seja o mais intuitivo para a criação de um intervalo de confiança através do *bootstrap*, o método chamado BC_a (*bias-corrected and accelerated*), o qual utilizamos nesse trabalho, produz melhores intervalos de confiança. Embora uma explicação mais detalhada do método esteja fora do escopo desse trabalho, apresentamos a formulação para os limites inferior (z_L) e superior (z_U) abaixo:

$$z_L = \frac{z_0 - z_{1-\alpha/2}}{1 - a(z_0 - z_{1-\alpha/2})} + z_0 \quad e \quad z_U = \frac{z_0 + z_{1-\alpha/2}}{1 - a(z_0 + z_{1-\alpha/2})} + z_0$$

Nas fórmulas, z_0 representa a mediana não-viesada e é definida através $P(Z \leq z_0) = p_0$, onde p_0 é a proporção das estatísticas calculadas nas reamostragens do *bootstrap* cujo valor é inferior ao da estatística calculada na amostra original. Por exemplo, supondo que a média amostral dos primeiros dígitos das 5.000 notas de empenho seja igual a 2,93, e que, das 25.000 médias amostrais calculadas via *bootstrap*, 12.137 delas

tenham valor inferior a 2,93, temos $p_0 = 0,4855$, correspondendo a $z_0 = -0,0364$. O termo $z_{1-\alpha/2}$ representa o valor crítico da distribuição normal para um dado nível de significância α . Assim, para α igual a 5%, temos $z_{1-\alpha/2} = 1,96$. Por sua vez, o termo a representa a assimetria dos dados, e é igual a $1/6$ do valor da assimetria (γ) da distribuição amostral obtida via *bootstrap*. Assim, dado μ , a média das T estatísticas amostrais calculadas², temos:

$$a = \frac{1}{6} \frac{\sum_{i=1}^T (x_i - \mu)^3}{(\sum_{i=1}^T (x_i - \mu)^2)^{3/2}} = \frac{1}{6} \gamma$$

Onde x_i é a i-ésima estatística amostral calculada via *bootstrap* e T é a quantidade total de amostras geradas via *bootstrap*.

Note que, se a mediana das estatísticas calculadas nas reamostragens for idêntica à mediana da estatística calculada sobre a amostra original, isto é, se $p_0 = 0,5$ e $z_0 = 0$; e se a distribuição das estatísticas calculadas nas reamostragens for simétrica, isto é, $\gamma = 0$ e $a = 0$, o método BC_a se equivale ao método do percentil.

4. BOOTSTRAP APLICADO À LEI DE NEWCOMB-BENFORD

Apresentamos nesse capítulo os detalhes da aplicação da técnica estatística do *bootstrap* à análise da distribuição dos dígitos baseada na Lei de Newcomb-Benford de duas formas distintas: *bootstrap* (tradicional) e regressão *bootstrap*.

² No exemplo, as 25.000 médias amostrais calculadas podem ser entendidas como 25.000 elementos de uma variável aleatória X e, portanto, podemos calcular a média de X, isto é, a média das médias amostrais calculadas.

4.1 Metodologia *Bootstrap*

O procedimento proposto por Suh e Headrick (2010) envolve a utilização do *bootstrap* para a determinação de intervalos de confiança para a média (μ), variância (σ^2), assimetria (γ) e curtose (δ) relativos aos dois primeiros dígitos, considerados separada e conjuntamente, e para o coeficiente de correlação de Pearson entre os dois dígitos.

4.1.1 Cálculo dos Parâmetros da Distribuição de Newcomb-Benford

O primeiro passo, portanto, é calcular esses parâmetros populacionais, assumindo uma distribuição de Newcomb-Benford. Para tal, precisamos calcular também os quatro primeiros momentos para os primeiros e segundos dígitos em separado e em conjunto, conforme segue:

$$\mu_r^{d_1} = \sum_{d_1=1}^9 d_1^r P_{d_1} ; \mu_r^{d_2} = \sum_{d_2=0}^9 d_2^r P_{d_2} ; \mu_r^{d_1 d_2} = \sum_{d_1 d_2=10}^{99} (d_1 d_2)^r P_{d_1 d_2}$$

Onde μ_r são os momentos; d_1 , d_2 , e $d_1 d_2$ são, respectivamente, os primeiros, segundos, e dois primeiros dígitos; e P_{d_1} , P_{d_2} e $P_{d_1 d_2}$ são, respectivamente, as probabilidades associadas aos primeiros, segundos e dois primeiros dígitos, de acordo com as Tabelas 3.1, 3.2 e 3.3.

A tabela abaixo apresenta os resultados encontrados para os quatro primeiros momentos:

Tabela 4.1: Momentos das Distribuições

Momento	Primeiros, Segundos e Dois Primeiros Dígitos		
	d_1	d_2	$d_1 d_2$
μ_1	3,440237	4,187390	38,5898

μ_2	17,891743	25,788011	2.111,0010
μ_3	115,082049	180,264352	141.424,5000
μ_4	823,273101	1.350,711251	10.569.883,0000

Fonte: Elaborado pelo autor.

Dados os valores calculados para os momentos, podemos calcular os parâmetros populacionais, conforme segue:

$$\begin{aligned}\mu &= \mu_1 \\ \sigma^2 &= \mu_2 - \mu_1^2 \\ \gamma &= \frac{(\mu_3 - 3\mu_2\mu_1 + 2\mu_1^3)}{\sigma^3} \\ \delta &= \frac{(\mu_4 - 4\mu_3\mu_1 - 3\mu_2^2 + 12\mu_2\mu_1^2 - 6\mu_1^4)}{\sigma^4}\end{aligned}$$

A tabela abaixo apresenta os resultados encontrados para os quatro parâmetros populacionais:

Tabela 4.2: Parâmetros Populacionais das Distribuições

Parâmetro	Primeiros, Segundos e Dois Primeiros Dígitos		
	d_1	d_2	d_1d_2
Média (μ)	3,440237	4,187390	38,5898
Variância (σ^2)	6,056513	8,253779	621,8317
Assimetria (γ)	0,795604	0,133107	0,7719
Curtose (δ)	-0,548225	-1,208395	-0,5465

Fonte: Elaborado pelo autor.

Além disso, calculamos também o coeficiente de correlação de Pearson entre o primeiro e o segundo dígitos através das seguintes fórmulas:

$$\begin{aligned}E[d_1d_2] &= \sum_{d_1=1}^9 \sum_{d_2=0}^9 (d_1)(d_2) \log_{10} \left(1 + \frac{1}{d_1d_2} \right) \\ \rho_{d_1d_2} &= \frac{(E[d_1d_2] - \mu_1\mu_2)}{(\sigma_1^2\sigma_2^2)^{1/2}}\end{aligned}$$

Para $E[d_1d_2]$, encontramos o valor de 14,801940 e para o coeficiente de correlação $\rho_{d_1d_2}$, encontramos o valor 0,05605634.

A esse conjunto de parâmetros calculados, chamamos de parâmetros exatos.

4.1.2 Cálculo dos Intervalos de Confiança (ICs) dos Parâmetros Exatos

O segundo passo consiste na aplicação do procedimento de *bootstrap* para estimar a distribuição amostral dos parâmetros para um dado tamanho de amostra N (correspondente ao tamanho do conjunto de dados o qual se deseja testar a conformidade) e, assim, calcular seus ICs. Portanto, sorteamos uma nova amostra aleatória de tamanho N , cujos elementos são números entre 10 e 99, inclusive, a partir das probabilidades dadas pela distribuição de Newcomb-Benford constantes da Tabela 3.3. Então calculamos as estimativas amostrais para a média (μ), variância (σ^2), assimetria (γ) e curtose (δ) para os primeiros dígitos e para os segundos dígitos dos elementos da amostra e também para os elementos em si, considerando os dois dígitos. Repetimos esse procedimento 25.000 vezes (T) e obtemos as distribuições amostrais de cada um dos parâmetros, tanto para os primeiros, quanto para os segundos e para os dois primeiros dígitos. Então, para cada distribuição, construímos os ICs utilizando o método BC_a.

Os ICs construídos para os parâmetros exatos, portanto, são originários de uma verdadeira distribuição de Newcomb-Benford e são adotados como referência de comparação para testar se um determinado conjunto de dados está em conformidade ou não com a distribuição.

Logo, a fim de facilitar essa comparação, e dado o elevado custo computacional do procedimento, elaboramos uma tabela contendo os limites mínimos e máximos dos ICs de cada um dos parâmetros, considerando um nível de significância de 1%, para tamanhos de amostra variando entre 30 e 1.000.000 de elementos, para os primeiros e segundos dígitos, e entre 1.000 e 1.000.000 para os dois primeiros dígitos. Segundo Suh e Headrick (2010), como são quatro os parâmetros a serem testados, o nível de significância total de 5% corresponde, pela desigualdade de Bonferroni, a 1,25% por teste, o qual os autores optaram por arredondar para 1% dado que esse é um nível de significância largamente utilizado. Os autores recomendam esse nível conservador a fim de prevenir a ocorrência de falsos positivos.

Tabela 4.3: ICs dos Parâmetros Exatos para os Primeiros Dígitos (d_1)

Intervalos de Confiança – Primeiros Dígitos (d_1) – $\alpha = 1\%$									
N	μ		σ		γ		δ		
	inferior	superior	inferior	superior	inferior	superior	inferior	superior	
30	2,4333	4,7851	2,9789	9,7656	0,2011	2,2910	-1,2054	8,0409	
40	2,5500	4,5500	3,2494	9,1100	0,2491	1,9288	-1,1877	5,2543	
50	2,6400	4,4400	3,6084	8,9291	0,2914	1,6612	-1,1761	4,9546	
60	2,7000	4,3333	3,7233	8,5105	0,3299	1,5499	-1,1580	3,9428	
70	2,7286	4,2571	3,9184	8,4271	0,3528	1,5212	-1,1297	3,3382	
80	2,8000	4,2375	4,0630	8,2244	0,3805	1,4224	-1,1214	3,0017	
90	2,8222	4,1444	4,1515	8,0272	0,4028	1,3929	-1,0973	2,6661	
100	2,8600	4,1400	4,2414	8,0280	0,4200	1,3469	-1,0908	2,0675	
110	2,8727	4,1000	4,3111	7,8700	0,4352	1,3269	-1,0656	1,7560	
120	2,8917	4,0417	4,3364	7,7571	0,4491	1,3015	-1,0505	2,1376	
130	2,9154	4,0231	4,4630	7,7517	0,4630	1,2664	-1,0430	1,5924	
140	2,9357	3,9985	4,5357	7,6798	0,4726	1,2316	-1,0332	1,9186	
150	2,9533	4,0000	4,5362	7,6051	0,4771	1,2398	-1,0212	1,1753	
200	3,0200	3,9150	4,7702	7,4039	0,5207	1,1565	-0,9871	0,6192	
250	3,0640	3,8720	4,9024	7,2881	0,5462	1,1069	-0,9641	0,4314	
300	3,0958	3,8300	4,9764	7,1813	0,5643	1,0790	-0,9303	0,3610	
350	3,1200	3,7958	5,0570	7,0826	0,5769	1,0494	-0,9134	0,2341	
400	3,1375	3,7700	5,1273	6,9873	0,5918	1,0367	-0,8875	0,1847	
450	3,1600	3,7556	5,1914	6,9209	0,5992	1,0186	-0,8825	0,0743	
500	3,1680	3,7280	5,2436	6,9161	0,6145	1,0103	-0,8698	0,0455	
600	3,1950	3,7200	5,2812	6,8201	0,6222	0,9872	-0,8474	-0,0214	
700	3,2129	3,6957	5,3423	6,7517	0,6321	0,9694	-0,8337	-0,0848	
800	3,2213	3,6750	5,4065	6,7307	0,6477	0,9607	-0,8198	-0,1178	

900	3,2392	3,6578	5,4371	6,6666	0,6502	0,9470	-0,8068	-0,1499
999	3,2465	3,6496	5,4716	6,6548	0,6626	0,9419	-0,7957	-0,1762
1.000	3,2445	3,6420	5,4472	6,6385	0,6655	0,9444	-0,7877	-0,1639
2.000	3,3015	3,5870	5,6467	6,4873	0,6997	0,8992	-0,7340	-0,2972
3.000	3,3273	3,5600	5,7224	6,4081	0,7188	0,8794	-0,7017	-0,3507
4.000	3,3408	3,5395	5,7555	6,3487	0,7293	0,8671	-0,6798	-0,3789
5.000	3,3500	3,5290	5,7914	6,3168	0,7360	0,8592	-0,6686	-0,4023
6.000	3,3585	3,5234	5,8095	6,2974	0,7398	0,8547	-0,6599	-0,4095
7.000	3,3654	3,5170	5,8303	6,2774	0,7452	0,8500	-0,6486	-0,4239
8.000	3,3695	3,5104	5,8434	6,2609	0,7469	0,8452	-0,6456	-0,4359
9.000	3,3752	3,5088	5,8643	6,2583	0,7500	0,8438	-0,6403	-0,4413
10.000	3,3761	3,5062	5,8668	6,2443	0,7511	0,8399	-0,6357	-0,4447
20.000	3,3958	3,4856	5,9277	6,1892	0,7653	0,8274	-0,6117	-0,4788
30.000	3,4040	3,4771	5,9489	6,1666	0,7708	0,8217	-0,6001	-0,4916
40.000	3,4084	3,4715	5,9624	6,1512	0,7741	0,8179	-0,5934	-0,4994
50.000	3,4125	3,4693	5,9715	6,1399	0,7762	0,8158	-0,5896	-0,5051
60.000	3,4146	3,4661	5,9804	6,1324	0,7777	0,8136	-0,5859	-0,5090
70.000	3,4166	3,4643	5,9848	6,1262	0,7794	0,8127	-0,5816	-0,5114
80.000	3,4184	3,4628	5,9905	6,1252	0,7805	0,8111	-0,5809	-0,5144
90.000	3,4191	3,4613	5,9934	6,1192	0,7810	0,8102	-0,5791	-0,5163
100.000	3,4206	3,4608	5,9987	6,1179	0,7817	0,8094	-0,5775	-0,5177
200.000	3,4257	3,4543	6,0145	6,0991	0,7857	0,8058	-0,5695	-0,5272
500.000	3,4312	3,4493	6,0298	6,0836	0,7893	0,8018	-0,5617	-0,5349
1.000.000	3,4336	3,4466	6,0375	6,0754	0,7912	0,8001	-0,5577	-0,5387

Fonte: Elaborado pelo autor.

Tabela 4.4: ICs dos Parâmetros Exatos para os Segundos Dígitos (d_2)

Intervalos de Confiança – Primeiros Dígitos (d_2) – $\alpha = 1\%$								
N	μ		σ		γ		δ	
	inferior	superior	inferior	superior	inferior	superior	inferior	superior
30	2,9000	5,5667	5,3289	12,3180	-0,5545	0,8716	-1,5315	1,0579
40	3,0500	5,4000	5,6122	11,6752	-0,4338	0,7900	-1,4970	0,7361
50	3,1600	5,2400	5,8816	11,3206	-0,3645	0,7011	-1,4884	-0,0729
60	3,2667	5,1667	6,0989	11,0733	-0,3483	0,6286	-1,4732	-0,2885
70	3,3286	5,1143	6,2485	10,7730	-0,3050	0,5909	-1,4565	-0,3951
80	3,3750	5,0125	6,3436	10,6677	-0,2823	0,5647	-1,4438	-0,4355
90	3,4222	4,9889	6,4105	10,3857	-0,2478	0,5311	-1,4320	-0,5524
100	3,4300	4,9300	6,5004	10,3824	-0,2337	0,5169	-1,4280	-0,5951
110	3,5000	4,9111	6,5627	10,1763	-0,2090	0,5075	-1,4184	-0,6746
120	3,5333	4,8833	6,6779	10,1687	-0,1991	0,4882	-1,4125	-0,6800
130	3,5462	4,8385	6,6956	10,0633	-0,1867	0,4655	-1,4082	-0,7251
140	3,5857	4,8214	6,7549	9,9990	-0,1753	0,4574	-1,4031	-0,7750
150	3,6000	4,7933	6,7862	9,9405	-0,1632	0,4428	-1,3995	-0,7981
200	3,6650	4,7150	7,0228	9,6914	-0,1264	0,4015	-1,3819	-0,8757
250	3,7160	4,6600	7,1210	9,4905	-0,1043	0,3724	-1,3665	-0,9360
300	3,7667	4,6233	7,2545	9,4324	-0,0797	0,3557	-1,3571	-0,9734
350	3,7914	4,5800	7,2811	9,3111	-0,0608	0,3366	-1,3449	-0,9824
400	3,8250	4,5675	7,3551	9,2669	-0,0481	0,3233	-1,3411	-1,0079
450	3,8395	4,5326	7,4061	9,1967	-0,0374	0,3127	-1,3317	-1,0157

500	3,8540	4,5180	7,4234	9,1506	-0,0304	0,3060	-1,3264	-1,0325
600	3,8967	4,5067	7,5238	9,0736	-0,0210	0,2861	-1,3200	-1,0527
700	3,9086	4,4762	7,5532	8,9720	-0,0058	0,2752	-1,3101	-1,0602
800	3,9250	4,4517	7,6079	8,9682	0,0062	0,2670	-1,3074	-1,0783
900	3,9508	4,4334	7,6491	8,9362	0,0129	0,2573	-1,3000	-1,0822
999	3,9534	4,4214	7,6586	8,8619	0,0174	0,2521	-1,2957	-1,0895
1.000	3,9510	4,4160	7,6618	8,8573	0,0165	0,2505	-1,2944	-1,0927
2.000	4,0208	4,3535	7,8315	8,6929	0,0497	0,2173	-1,2727	-1,1275
3.000	4,0513	4,3220	7,9186	8,6051	0,0668	0,2027	-1,2609	-1,1454
4.000	4,0690	4,3018	7,9588	8,5514	0,0745	0,1923	-1,2552	-1,1553
5.000	4,0826	4,2920	7,9873	8,5238	0,0803	0,1849	-1,2510	-1,1617
6.000	4,0915	4,2840	8,0174	8,5101	0,0844	0,1807	-1,2477	-1,1661
7.000	4,0986	4,2778	8,0320	8,4830	0,0892	0,1782	-1,2438	-1,1693
8.000	4,1020	4,2675	8,0413	8,4661	0,0926	0,1760	-1,2412	-1,1712
9.000	4,1119	4,2667	8,0591	8,4531	0,0938	0,1713	-1,2401	-1,1748
10.000	4,1156	4,2620	8,0694	8,4439	0,0950	0,1690	-1,2380	-1,1759
20.000	4,1323	4,2388	8,1229	8,3908	0,1076	0,1601	-1,2305	-1,1859
30.000	4,1448	4,2323	8,1476	8,3670	0,1108	0,1543	-1,2263	-1,1898
40.000	4,1512	4,2239	8,1607	8,3483	0,1148	0,1518	-1,2235	-1,1923
50.000	4,1547	4,2208	8,1709	8,3385	0,1162	0,1494	-1,2219	-1,1938
60.000	4,1561	4,2172	8,1758	8,3299	0,1182	0,1483	-1,2208	-1,1953
70.000	4,1592	4,2157	8,1838	8,3266	0,1190	0,1472	-1,2199	-1,1962
80.000	4,1611	4,2131	8,1869	8,3205	0,1201	0,1462	-1,2195	-1,1972
90.000	4,1633	4,2120	8,1902	8,3187	0,1206	0,1453	-1,2191	-1,1980
100.000	4,1643	4,2104	8,1911	8,3135	0,1213	0,1446	-1,2183	-1,1984
200.000	4,1709	4,2043	8,2110	8,2961	0,1247	0,1415	-1,2153	-1,2011
500.000	4,1766	4,1975	8,2260	8,2798	0,1279	0,1384	-1,2127	-1,2038
1.000.000	4,1800	4,1948	8,2347	8,2726	0,1294	0,1368	-1,2116	-1,2052

Fonte: Elaborado pelo autor.

Tabela 4.5: ICs dos Parâmetros Exatos para os Dois Primeiros Dígitos (d_1d_2)

Intervalos de Confiança – Dois Primeiros Dígitos (d_1d_2) – $\alpha = 1\%$								
N	μ		σ		γ		δ	
	inferior	superior	inferior	superior	inferior	superior	inferior	superior
1.000	36,5938	40,6281	559,2820	681,5606	0,6443	0,9185	-0,7798	-0,1635
2.000	37,1666	40,0843	579,1456	666,3730	0,6768	0,8763	-0,7271	-0,3031
3.000	37,4530	39,7881	587,4419	658,4911	0,6959	0,8552	-0,6978	-0,3539
4.000	37,5833	39,5956	590,7047	651,4603	0,7060	0,8439	-0,6768	-0,3805
5.000	37,6700	39,4869	594,4151	648,7819	0,7120	0,8343	-0,6671	-0,4054
6.000	37,7395	39,4208	596,6074	646,3503	0,7170	0,8297	-0,6575	-0,4110
7.000	37,8290	39,3611	598,5476	644,5461	0,7215	0,8253	-0,6453	-0,4231
8.000	37,8746	39,2988	599,8943	642,6063	0,7241	0,8213	-0,6422	-0,4356
9.000	37,9203	39,2756	602,1888	642,8859	0,7275	0,8189	-0,6370	-0,4406
10.000	37,9402	39,2611	602,4236	641,2057	0,7283	0,8165	-0,6330	-0,4434
20.000	38,1380	39,0514	608,3510	635,3092	0,7412	0,8031	-0,6090	-0,4787
30.000	38,2227	38,9598	610,8717	633,1155	0,7472	0,7974	-0,5974	-0,4908
40.000	38,2664	38,9055	612,1864	631,5620	0,7507	0,7940	-0,5911	-0,4983
50.000	38,3073	38,8860	613,0438	630,4354	0,7526	0,7917	-0,5864	-0,5040
60.000	38,3337	38,8511	614,0415	629,7694	0,7543	0,7898	-0,5834	-0,5077

70.000	38,3524	38,8341	614,4141	628,9623	0,7557	0,7886	-0,5799	-0,5102
80.000	38,3683	38,8174	615,0475	628,8814	0,7568	0,7872	-0,5788	-0,5132
90.000	38,3784	38,8028	615,3317	628,2582	0,7575	0,7862	-0,5766	-0,5155
100.000	38,3911	38,7972	615,8604	628,0992	0,7581	0,7857	-0,5752	-0,5165
200.000	38,4455	38,7326	617,4683	626,1785	0,7621	0,7818	-0,5673	-0,5257
500.000	38,4980	38,6806	619,1061	624,6430	0,7657	0,7780	-0,5598	-0,5336
1.000.000	38,5227	38,6542	619,9068	623,7873	0,7675	0,7763	-0,5559	-0,5372

Fonte: Elaborado pelo autor.

4.1.3 Cálculo dos Intervalos de Confiança (ICs) das Estatísticas Reais

No terceiro passo, aplicamos o procedimento de *bootstrap* sobre o conjunto de dados reais o qual desejamos testar a conformidade. Dado esse conjunto de dados, de tamanho N , obtemos os primeiros dois dígitos de todos os seus elementos e procedemos exatamente da mesma forma que no passo anterior, obtendo as distribuições amostrais das mesmas estatísticas (μ , σ^2 , γ e δ) para os primeiros, segundos e dois primeiros dígitos, e seus respectivos ICs pelo método BC_a . Esses são os ICs das estatísticas reais (em referência aos dados reais). Além disso, diferentemente do passo anterior, também aplicamos o procedimento de *bootstrap* e de construção dos ICs ao coeficiente de correlação de Pearson (ρ) entre os primeiros e os segundos dígitos.

4.1.4 Critérios de Conformidade

No quarto passo, elencamos os critérios de decisão para inferência a respeito da conformidade ou não de um conjunto de dados com a distribuição de Newcomb-Benford.

Primeiramente, definimos a hipótese nula tal que a estatística real (isto é, calculada a partir dos dados reais) é igual ao parâmetro exato correspondente (isto é, calculado a

partir da verdadeira distribuição de Newcomb-Benford). Para que a hipótese nula seja aceita, é necessário que pelo menos uma das seguintes condições seja satisfeita:

- existe intersecção do IC calculado para a estatística real (Item 4.1.3) e do IC calculado para o parâmetro exato correspondente (Item 4.1.2);
- o IC calculado para a estatística real contém o parâmetro exato correspondente (Item 4.1.1).

Definidas as hipóteses nulas e seus critérios de aceitação, passamos aos critérios de decisão em si:

- i) Para conjuntos de dados com mais de 1.000 elementos, testamos as hipóteses nulas baseadas nas estatísticas e parâmetros referentes aos dois primeiros dígitos, tomados em conjunto. Três resultados são possíveis:
 - a. se todas as quatro hipóteses nulas, relativas às quatro estatísticas (média, variância, assimetria e curtose) forem aceitas, podemos concluir que os dados testados estão em conformidade com a Lei de Newcomb-Benford e não contêm anomalias (erros ou fraudes);
 - b. se todas as quatro hipóteses nulas forem rejeitadas, concluímos que os dados testados provavelmente contêm erros ou fraudes e devem ser investigados de forma mais aprofundada;
 - c. se algumas hipóteses forem aceitas e outras rejeitadas, o resultado do teste é dito inconsistente.
- ii) Para conjuntos de dados com menos de 1.000 elementos, ou cujos testes para os dois primeiros dígitos foram considerados inconsistentes (subitem i-c,

acima), testamos as hipóteses nulas baseadas nas estatísticas e parâmetros referentes tanto aos primeiros, quanto aos segundos dígitos, totalizando, portanto, oito testes de hipóteses. Novamente, três resultados são possíveis:

- a. se todas as oito hipóteses nulas forem aceitas, os dados testados não contêm anomalias;
 - b. se todas as oito hipóteses nulas forem rejeitadas, os dados testados provavelmente contêm erros ou fraudes e devem ser investigados;
 - c. se algumas hipóteses forem aceitas e outras rejeitadas, o resultado do teste é dito inconsistente.
- iii) Para os casos em que os testes dos primeiros e segundos dígitos foram considerados inconsistentes (subitem ii-c, acima), verificamos se o IC do coeficiente de correlação de Pearson (calculado no Item 4.1.3) contém o respectivo parâmetro exato (calculado no Item 4.1.1 e equivalente a 0,05605634):
- a. caso afirmativo, concluímos que os dados testados não contêm anomalias;
 - b. caso contrário, os dados testados provavelmente contêm erros ou fraudes e devem ser investigados.

Para realizar os testes descritos acima, podemos fazer uso da Tabela 4.3, que contém os ICs para os parâmetros exatos. Assim, para um conjunto de dados reais de tamanho N , cujas estatísticas desejamos testar, buscamos na tabela os limites para o ICs calculados para o tamanho de amostra mais próximo de N . Note que os intervalos se estreitam enquanto o tamanho da amostra cresce, de forma que, se desejamos

adotar uma postura conservadora, buscamos na tabela os valores correspondentes ao tamanho de amostra disponível imediatamente superior a N.

4.2 Regressão *Bootstrap*

O método de regressão *bootstrap* (*Bootstrap REGression*, em inglês), ou BREG, estende o método apresentado na seção anterior ao aplicar as técnicas de *bootstrap* e de construção de ICs sobre os coeficientes estimados de uma regressão linear. Iniciamos a descrição do método BREG pela expressão que define uma regressão linear:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

Onde Y_i é a variável dependente, X_i é a variável regressora, ε_i é o termo de erro e $i = 1 \dots N$. Ao estimar os parâmetros de uma regressão linear através de mínimos quadrados ordinários (MQO), assumimos o pressuposto que o termo de erro é normalmente distribuído. A maioria dos dados reais, entretanto, não atende a esse pressuposto, de forma que o *bootstrap* pode ser utilizado, dado que o método dispensa a necessidade de pressupostos a respeito da distribuição dos erros (Efron e Tibshirani, 1998).

No BREG, assim como no *bootstrap* tradicional, sorteamos uma amostra aleatória de tamanho N, com reposição, a partir do conjunto de dados original de tamanho N de elementos representados pelos respectivos pares de variáveis dependente e regressora, ou seja, $(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)$, onde todos os elementos são equiprováveis. Para cada amostra, no BREG, estimamos os parâmetros $\widehat{\beta}_0$ e $\widehat{\beta}_1$ da regressão linear. Repetindo esse processo T vezes, obtemos um conjunto de T estimativas para ambos os parâmetros que constituem suas respectivas distribuições.

As formas de construção dos intervalos de confiança, seja pelo método dos percentis ou pelo método BC_a, são as mesmas já apresentadas.

Suh, Headrick e Minaburo (2011), então, introduzem o BREG no contexto da Lei de Newcomb-Benford. Os modelos de regressão aplicados à distribuição de Newcomb-Benford são expressos em termos dos primeiros (d_1) e segundos dígitos (d_2), cada qual como variável dependente e regressora, conforme segue:

$$d_{1i} = \alpha_0 + \alpha_1 d_{2i} + \varepsilon_{1i}$$

$$d_{2i} = \beta_0 + \beta_1 d_{1i} + \varepsilon_{2i}$$

Os parâmetros α_0 , α_1 , β_0 e β_1 são chamados “parâmetros de Benford” quando as probabilidades de d_1 e d_2 são dadas pela distribuição de Newcomb-Benford constante das Tabelas 3.1 e 3.2. Esses parâmetros podem ser expressos da seguinte forma (Pedhazur, 1997):

$$\alpha_1 = \rho_{d_1 d_2} \frac{\sigma_{d_1}}{\sigma_{d_2}} ; \alpha_0 = \mu_{d_1} - \alpha_1 \mu_{d_2}$$

$$\beta_1 = \rho_{d_1 d_2} \frac{\sigma_{d_2}}{\sigma_{d_1}} ; \beta_0 = \mu_{d_2} - \beta_1 \mu_{d_1}$$

Onde $\rho_{d_1 d_2}$ é o coeficiente de correlação de Pearson entre os primeiros e os segundos dígitos, μ_{d_1} e μ_{d_2} são médias e σ_{d_1} e σ_{d_2} são variâncias associadas, respectivamente, com os primeiros e segundos dígitos. A partir dos valores constantes na Tabela 4.2, calculamos os valores para os parâmetros de Benford e definimos as duas hipóteses nulas para o teste de conformidade:

$$H_{0,d_1} : \alpha_0 = 3,239164 ; \alpha_1 = 0,048019$$

$$H_{0,d_2} : \beta_0 = 3,962262 ; \beta_1 = 0,065439$$

Se ambas as hipóteses nulas forem aceitas, concluímos que os dados seguem a distribuição de Newcomb-Benford e, portanto, não contêm erros ou fraudes. Para o teste de conformidade, novamente adotamos um nível de significância de 5% que, conforme a desigualdade de Bonferroni, deve ser reduzido. Apesar de existirem quatro valores a serem testados, existe uma forte correlação entre as estimativas do intercepto e do coeficiente angular, isto é, $\rho_{\alpha_0\alpha_1} \approx \rho_{\beta_0\beta_1} \approx -0,82$. Por essa razão, realizamos o ajuste por um fator igual a dois, ao invés de quatro, resultando em um nível de significância de 2,5%.

Os demais procedimentos seguem a mesma lógica já detalhada na Seção 4.1:

- i) cálculo dos intervalos de confiança para os parâmetros de Benford (α_0 , α_1 , β_0 e β_1) exatos a partir das probabilidades de d_1 e d_2 através da técnica de *bootstrap* e do método BC_a;
- ii) cálculo dos intervalos de confiança para as estimativas dos parâmetros de Benford reais a partir do conjunto de dados reais através da técnica de *bootstrap* e do método BC_a;
- iii) aceitação de cada hipótese nula caso exista intersecção dos ICs ou o IC para a estimativa dos parâmetros reais contenha o respectivo parâmetro exato;
- iv) conclusão pela conformidade (não-conformidade) dos dados reais com a Lei de Newcomb-Benford se todas quatro hipóteses nulas forem aceitas (rejeitadas), da mesma forma que os subitens i-a, i-b e i-c do Item 4.1.4).

Mais uma vez, elaboramos uma tabela contendo os limites mínimos e máximos dos ICs de cada um dos quatro parâmetros de Benford, considerando um nível de significância de 2,5%, para tamanhos de amostra variando entre 1.000 e 1.000.000 de elementos, a qual apresentamos abaixo:

Tabela 4.6: ICs dos Parâmetros de Benford Exatos

Intervalos de Confiança – Parâmetros de Benford – $\alpha = 2,5\%$								
N	α_0		α_1		β_0		β_1	
	inferior	superior	inferior	superior	inferior	superior	inferior	superior
1.000	2,9395	3,5542	-0,0118	0,1091	3,6160	4,3148	-0,0162	0,1488
2.000	3,0308	3,4603	0,0049	0,0904	3,7173	4,2133	0,0068	0,1236
3.000	3,0626	3,4156	0,0131	0,0835	3,7636	4,1646	0,0176	0,1138
4.000	3,0876	3,3947	0,0179	0,0780	3,7882	4,1401	0,0244	0,1064
5.000	3,1076	3,3791	0,0211	0,0748	3,8093	4,1178	0,0289	0,1019
6.000	3,1136	3,3656	0,0230	0,0727	3,8189	4,1021	0,0315	0,0993
7.000	3,1225	3,3533	0,0249	0,0709	3,8291	4,0970	0,0338	0,0970
8.000	3,1332	3,3458	0,0264	0,0688	3,8402	4,0875	0,0360	0,0939
9.000	3,1381	3,3410	0,0274	0,0677	3,8439	4,0778	0,0375	0,0925
10.000	3,1478	3,3415	0,0283	0,0665	3,8568	4,0772	0,0383	0,0906
20.000	3,1714	3,3068	0,0343	0,0613	3,8849	4,0408	0,0467	0,0836
30.000	3,1852	3,2970	0,0369	0,0590	3,8997	4,0277	0,0503	0,0803
40.000	3,1905	3,2881	0,0385	0,0579	3,9072	4,0185	0,0524	0,0789
50.000	3,1963	3,2829	0,0393	0,0565	3,9115	4,0114	0,0536	0,0770
60.000	3,2008	3,2790	0,0400	0,0556	3,9175	4,0079	0,0545	0,0758
70.000	3,2022	3,2756	0,0409	0,0554	3,9210	4,0045	0,0557	0,0754
80.000	3,2058	3,2728	0,0414	0,0549	3,9234	4,0014	0,0563	0,0747
90.000	3,2072	3,2709	0,0417	0,0543	3,9258	3,9990	0,0568	0,0741
100.000	3,2086	3,2696	0,0420	0,0540	3,9268	3,9964	0,0573	0,0737
200.000	3,2175	3,2608	0,0437	0,0522	3,9375	3,9871	0,0595	0,0712
500.000	3,2255	3,2525	0,0453	0,0508	3,9464	3,9779	0,0617	0,0692
1.000.000	3,2294	3,2489	0,0461	0,0499	3,9512	3,9734	0,0628	0,0681

Fonte: Elaborado pelo autor.

4.3 Cálculo de ICs para a FDA via *Bootstrap*

Nessa seção apresentamos uma ideia experimental e, até onde sabemos, inovadora, que poderia se constituir como alternativa ao K-S Teste, cuja estatística de teste é baseada na diferença entre as funções densidade acumulada (*fda*) das distribuições. A intuição por trás do teste é que, se a *fda* observada difere

significativamente da *fda* esperada, a primeira supostamente não teria sido obtida através do mesmo processo gerador de dados da última. Em outras palavras, as *fdas* corresponderiam a amostras provenientes de distribuições distintas.

Em nosso caso, sabemos que a *fda* esperada é dada pela distribuição de Newcomb-Benford, mas desconhecemos a distribuição que gerou a *fda* observada e desejamos determinar se a mesma, também, é fruto de uma realização a partir da distribuição de Newcomb-Benford. Então, a partir de uma amostra gerada a partir da verdadeira distribuição de Newcomb-Benford para os dois primeiros dígitos (Tabela 3.3), realizamos 25.000 reamostragens e, para cada uma, calculamos a respectiva função densidade acumulada (*fda*), obtendo os ICs através do método dos percentis. Intuitivamente, os ICs contém, para um dado tamanho de amostra e nível de significância, as *fdas* cujo processo gerador subjacente é, de fato, a distribuição de Newcomb-Benford. Assim, verificamos se a *fda* observada está contida no respectivo IC e, caso afirmativo, inferimos que ambas as distribuições são idênticas.

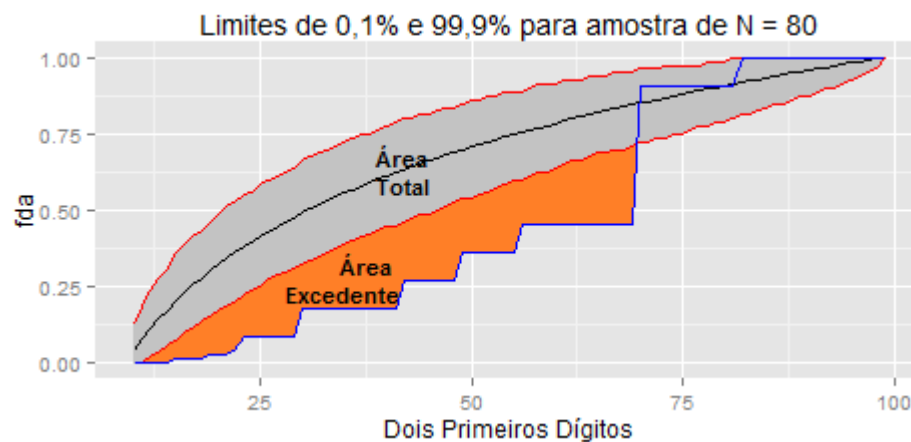
Como mencionado, essa abordagem é experimental e carece de uma investigação mais aprofundada para fins de sua validação como instrumento de análise estatística. Entretanto, com o intuito de motivação para tal, apresentamos os resultados obtidos a partir dessa abordagem e os comparamos com as demais. A métrica utilizada é a razão entre duas áreas:

- a área contida entre a *fda* observada e o limite inferior de 0,1% da *fda* esperada, quando a *fda* observada se encontra abaixo desse limite (em laranja na Figura 4.1); somada à área contida entre a *fda* observada e o limite

superior de 99,9% da *fda* esperada, quando a *fda* observada se encontra acima desse limite (inexistente na Figura 4.1); totalizando o que chamamos de área excedente; e

- a área contida entre os limites inferior e superior da *fda* esperada, a qual chamamos de área total (em cinza, na Figura 4.1).

Figura 4.1: Exemplo de Área Total e Área Excedente da *fda*



Fonte: Elaborado pelo autor.

A intuição por trás dessa razão é a mesma do K-S Teste, isto é, quanto maior a distância entre a *fda* observada e a *fda* esperada, maior será a área excedente. A área total, por sua vez, diminui conforme aumenta o tamanho da amostra, uma vez que os limites se aproximam. Portanto, quanto maior o tamanho da amostra, menor deve ser a área excedente a fim de que a razão entre elas se mantenha.

Se a *fda* observada estiver totalmente contida entre os limites, a razão será nula e, portanto, concluímos que a *fda* observada foi obtida a partir de uma verdadeira distribuição de Newcomb-Benford.

No presente trabalho, comparamos as razões calculadas com os resultados dos demais testes a fim de estimar para quais valores dessa razão, além do valor nulo, podemos chegar a mesma conclusão.

5. ANÁLISE DOS GASTOS DIRETOS DO GOVERNO FEDERAL

Nessa seção iniciamos a análise de conformidade dos gastos diretos do Governo Federal com a Lei de Newcomb-Benford. Para tanto, aplicamos tanto os testes tradicionais, Z-Teste e χ^2 -Teste, quanto os testes baseados em técnicas de *bootstrap* sobre os valores das despesas, separadas por Órgão e Elemento de Despesa.

5.1 Metodologia

Para cada conjunto de dados, isto é, de valores de despesas, extraímos os dois primeiros dígitos de cada valor e, sobre esses dados transformados, aplicamos uma série de testes, a saber:

- i) Z-Teste para cada combinação dos dois primeiros dígitos (10 a 99);
- ii) χ^2 -Teste para a distribuição dos dois primeiros dígitos (10 a 99);
- iii) Testes de hipóteses para os ICs dos parâmetros populacionais através técnica de *bootstrap* segundo Suh e Headrick (2010);
- iv) Testes de hipóteses para os ICs dos coeficientes de regressão através da técnica BREG segundo Suh, Headrick e Minaburo (2011); e
- v) Razão entre as áreas excedente e total para a *fda*.

Os resultados dos testes são apresentados de forma extensiva e também de forma esquemática, para facilitar a comparação de seus resultados.

5.2 Base de Dados

A base de dados utilizada foi construída a partir dos dados de recursos gastos diretamente pelo Governo Federal obtidos no Portal de Transparência, no caminho Download de Dados / Despesas – Gastos Diretos / Pagamentos (<http://transparencia.gov.br/downloads/mensal.asp?c=GastosDiretos>). No referido sítio, encontram-se os “pagamentos para aquisição e contratação de obras e compras governamentais, dentre outros, realizados pelos órgãos e entidades da Administração Pública Federal, que executam as despesas pelo Sistema Integrado de Administração Financeira do Governo Federal (SIAFI)”.

Os arquivos são mensais e contém diversas informações, das quais utilizaremos apenas cinco: Código Órgão Superior, Nome Órgão Superior, Código Elemento Despesa, Nome Elemento Despesa e Valor Pagamento. Assim, construímos os conjuntos de dados a serem analisados agrupando os valores de pagamento de todos os meses de cada ano por órgão superior e elemento de despesa.

5.3 Escopo

No presente trabalho, foram considerados os gastos diretos do Governo Federal realizados de 2011 a 2014, para todos os órgãos do Governo Federal e para um subconjunto definido de elementos de despesa. A tabela abaixo apresenta os códigos e nomes dos órgãos:

Tabela 5.1: Órgãos do Governo Federal

Código	Órgão
20000	PRESIDENCIA DA REPUBLICA
20102	GABINETE DA VICE-PRESIDENCIA DA REPUBLICA
20113	MINISTERIO DO PLANEJAMENTO, ORCAMENTO E GESTAO
22000	MINIST. DA AGRICUL., PECUARIA E ABASTECIMENTO
24000	MINISTERIO DA CIENCIA E TECNOLOGIA
25000	MINISTERIO DA FAZENDA
26000	MINISTERIO DA EDUCACAO
28000	MINISTERIO DO DESENV, IND. E COMERCIO EXTERIOR
30000	MINISTERIO DA JUSTICA
32000	MINISTERIO DE MINAS E ENERGIA
33000	MINISTERIO DA PREVIDENCIA SOCIAL
35000	MINISTERIO DAS RELACOES EXTERIORES
36000	MINISTERIO DA SAUDE
38000	MINISTERIO DO TRABALHO E EMPREGO
39000	MINISTERIO DOS TRANSPORTES
41000	MINISTERIO DAS COMUNICACOES
42000	MINISTERIO DA CULTURA
44000	MINISTERIO DO MEIO AMBIENTE
49000	MINISTERIO DO DESENVOLVIMENTO AGRARIO
51000	MINISTERIO DO ESPORTE
52000	MINISTERIO DA DEFESA
53000	MINISTERIO DA INTEGRACAO NACIONAL
54000	MINISTERIO DO TURISMO
55000	MINISTERIO DO DESENV. SOCIAL E COMBATE A FOME
56000	MINISTERIO DAS CIDADES
58000	MINISTERIO DA PESCA E AQUICULTURA

Fonte: Elaborado pelo autor.

Os elementos de despesa considerados são apresentados na tabela abaixo:

Tabela 5.2: Elementos de Despesa

Código	Elemento Despesa
30	Material de Consumo
32	Material de Distribuição Gratuita
34	Outras Despesas de Pessoal decorrentes de Contratos de Terceirização
35	Serviços de Consultoria
37	Locação de Mão-de-Obra
39	Outros Serviços de Terceiros - Pessoa Jurídica
51	Obras e Instalações
52	Equipamentos e Material Permanente

Fonte: Elaborado pelo autor.

Definimos esse subconjunto de elementos de despesa pois estamos interessados em analisar os gastos diretos relacionados à aquisição de bens e serviços e tais valores preenchem os requisitos elencados na Seção 3.1 para serem submetidos à análise de conformidade com a Lei de Newcomb-Benford.

As tabelas abaixo apresentam a quantidade de despesas, isto é, o número de elementos presente em cada ano, conforme o órgão ou elemento de despesa em que a mesma foi realizada:

Tabela 5.3: Quantidade de Despesas por Órgão e Ano

Órgão	2011	2012	2013	2014	Total
20000	99.790	97.948	98.657	52.725	349.120
20102	453	-	-	-	453
20113	56.156	59.690	56.500	57.263	229.609
22000	244.976	227.728	226.541	212.508	911.753
24000	97.047	99.100	96.088	93.622	385.857
25000	298.259	292.876	288.162	298.401	1.177.698
26000	1.047.400	1.131.004	1.177.927	1.293.600	4.649.931
28000	19.634	18.884	19.389	20.267	78.174
30000	158.498	155.540	161.683	153.192	628.913
32000	73.197	69.060	64.267	63.411	269.935
33000	165.736	162.732	174.111	182.840	685.419
35000	8.754	8.739	8.356	8.423	34.272
36000	327.947	310.292	300.455	305.051	1.243.745
38000	44.378	45.386	45.771	47.731	183.266
39000	208.810	205.988	218.964	221.785	855.547
41000	22.013	21.520	19.374	18.412	81.319
42000	51.483	50.023	50.001	53.978	205.485
44000	77.060	79.594	78.665	80.312	315.631
49000	43.145	40.296	41.942	47.380	172.763
51000	5.809	4.027	3.795	3.682	17.313
52000	974.125	1.005.749	1.021.455	1.066.554	4.067.883
53000	46.775	49.322	50.387	53.389	199.873
54000	6.066	7.518	7.482	8.387	29.453
55000	3.799	7.968	11.462	5.217	28.446
56000	44.008	53.091	42.388	39.749	179.236
58000	3.727	4.609	5.204	5.900	19.440
Total	4.129.045	4.208.684	4.269.026	4.393.779	17.000.534

Fonte: Elaborado pelo autor.

Tabela 5.4: Quantidade de Despesas por Elemento de Despesa e Ano

Elemento de Despesa	2011	2012	2013	2014	Total
30	1.162.097	1.158.140	1.148.522	1.190.235	4.658.994
32	17.206	18.856	21.700	17.608	75.370
34	6.330	6.217	3.936	2.781	19.264
35	27.158	24.355	19.104	17.514	88.131
37	422.023	458.311	501.086	558.181	1.939.601
39	2.103.420	2.185.957	2.255.677	2.268.680	8.813.734
51	157.409	122.144	99.281	100.410	479.244
52	233.402	234.704	219.720	238.370	926.196
Total	4.129.045	4.208.684	4.269.026	4.393.779	17.000.534

Fonte: Elaborado pelo autor.

Podemos observar que a metade da quantidade de despesas se concentra nos ministérios da Educação e da Defesa, seguidos dos ministérios da Saúde e da Fazenda. Em relação aos elementos de despesa, Outros Serviços de Terceiros – Pessoa Jurídica concentra mais da metade da quantidade de despesas realizadas, enquanto Material de Consumo responde por aproximadamente um quarto do total.

5.4 Análise de Conformidade

Uma vez definido os conjuntos de dados reais objeto do presente estudo, aplicamos a seqüência de testes elencados na Seção 5.1 a cada um dos conjuntos, isto é, aos valores de despesas realizadas por um órgão ao longo de um ano e agrupados por elemento de despesa. Apresentamos os resultados dos testes em tabelas distintas para cada ano, onde: ‘S’ significa que o teste aponta para indícios de erro ou fraude; ‘?’, que o resultado do teste é inconsistente; ‘N’, que não há indícios de erro ou fraude e ‘nd’, que o teste não está disponível devido ao tamanho reduzido da amostra. O resultado final é dado pela combinação dos resultados ‘Param.Total’ e ‘Boot.Regr.’, sendo idêntico ao primeiro nos casos em que ambos combinam ou naqueles em que o último é inconsistente ou não disponível; e inconsistente nos demais casos.

Tabela 5.5: Resultado dos Testes para Índícios de Erro ou Fraude – 2011

Órgão	Elem.	Qtd.Dsp.	Z- Teste	χ^2 - Teste	Param. d_1/d_2	Param. d_1/d_2	Param. Total	Boot. Regr.	Razão	Final
20000	30	4719	N	S	N	N	N	N	0,00%	N
20000	35	54	N	S	nd	N	N	nd	0,09%	N
20000	37	26400	S	S	S	S	S	S	172,01%	S
20000	39	67018	S	S	S	S	S	?	63,87%	S
20000	51	276	N	N	nd	N	N	nd	0,00%	N
20000	52	1323	N	N	N	N	N	N	0,00%	N
20102	30	90	N	S	nd	N	N	nd	10,98%	N
20102	37	44	N	S	nd	N	N	nd	9,11%	N
20102	39	319	N	S	nd	N	N	nd	0,24%	N
20113	30	13131	S	S	S	S	S	S	165,93%	S
20113	35	253	S	S	nd	S	S	nd	8,31%	S
20113	37	8446	S	S	?	N	N	N	40,49%	N
20113	39	33552	S	S	?	N	N	N	12,96%	N
20113	51	170	N	S	nd	N	N	nd	2,40%	N
20113	52	604	N	S	nd	N	N	nd	0,00%	N
22000	30	54451	S	S	?	S	S	?	33,61%	S
22000	32	9840	N	S	N	N	N	N	13,49%	N
22000	35	190	N	S	nd	N	N	nd	2,07%	N
22000	37	20229	S	S	S	N	S	?	126,12%	S
22000	39	147335	S	S	?	S	S	?	12,87%	S
22000	51	4100	N	N	N	N	N	N	0,00%	N
22000	52	8831	S	S	N	N	N	N	3,45%	N
24000	30	23442	S	S	N	N	N	N	2,30%	N
24000	34	429	N	S	nd	N	N	nd	4,23%	N
24000	35	924	N	S	nd	N	N	nd	0,00%	N
24000	37	8789	S	S	?	S	S	?	9,89%	S
24000	39	57673	S	S	N	N	N	S	3,01%	?
24000	51	1025	N	S	N	N	N	N	0,00%	N
24000	52	4765	N	S	?	N	N	N	25,67%	N
25000	30	35089	S	S	N	N	N	N	10,02%	N
25000	34	3829	S	S	N	N	N	N	0,26%	N
25000	35	262	S	S	nd	N	N	nd	6,52%	N
25000	37	50061	S	S	?	S	S	S	99,67%	S
25000	39	200913	S	S	S	N	S	?	98,64%	S
25000	51	2343	N	N	N	N	N	N	0,00%	N
25000	52	5762	N	S	N	N	N	N	0,03%	N
26000	30	409841	S	S	N	N	N	N	3,21%	N
26000	32	3119	N	S	N	N	N	N	0,00%	N
26000	35	1618	S	S	?	N	N	N	17,74%	N
26000	37	102816	S	S	?	N	N	?	140,23%	N
26000	39	369574	S	S	?	S	S	?	32,87%	S
26000	51	40750	N	S	N	N	N	N	24,12%	N
26000	52	119682	S	S	?	N	N	N	25,23%	N
28000	30	1543	N	N	N	N	N	N	0,00%	N
28000	35	40	N	S	nd	N	N	nd	12,16%	N
28000	37	3588	S	S	?	N	N	N	10,08%	N
28000	39	13956	S	S	?	N	N	N	64,51%	N

28000	51	72	N	N	nd	N	N	nd	0,00%	N
28000	52	435	N	N	nd	N	N	nd	0,00%	N
30000	30	27393	N	S	N	N	N	N	0,06%	N
30000	35	107	N	S	nd	N	N	nd	0,42%	N
30000	37	32364	S	S	?	N	N	N	37,46%	N
30000	39	92938	S	S	?	N	N	?	12,65%	N
30000	51	1223	N	N	N	N	N	N	0,86%	N
30000	52	4473	N	S	N	N	N	N	3,69%	N
32000	30	12135	S	S	N	S	S	S	11,24%	S
32000	35	693	S	S	nd	S	S	nd	21,18%	S
32000	37	9712	S	S	S	S	S	?	78,22%	S
32000	39	48107	S	S	N	S	S	?	8,95%	S
32000	51	563	N	S	nd	N	N	nd	0,00%	N
32000	52	1987	N	S	?	S	S	N	19,69%	?
33000	30	6908	N	S	N	N	N	N	0,04%	N
33000	32	726	N	S	nd	N	N	nd	0,00%	N
33000	37	41481	S	S	?	S	S	S	83,98%	S
33000	39	110139	S	S	S	S	S	S	477,88%	S
33000	51	2642	N	N	N	N	N	N	0,00%	N
33000	52	3826	N	S	N	N	N	N	0,16%	N
35000	30	894	N	N	nd	N	N	nd	0,00%	N
35000	37	60	N	S	nd	N	N	nd	6,13%	N
35000	39	7516	S	S	S	S	S	N	41,53%	?
35000	52	242	N	N	nd	N	N	nd	0,00%	N
36000	30	163416	S	S	S	S	S	?	103,91%	S
36000	32	411	N	N	nd	N	N	nd	0,00%	N
36000	34	918	N	S	nd	N	N	nd	13,76%	N
36000	35	1140	S	S	N	N	N	N	12,89%	N
36000	37	27675	S	S	S	S	S	?	157,14%	S
36000	39	126006	S	S	N	N	N	N	3,92%	N
36000	51	2729	N	S	N	N	N	N	6,70%	N
36000	52	5652	N	S	N	N	N	N	0,00%	N
38000	30	5080	N	N	N	N	N	N	0,05%	N
38000	34	230	N	S	nd	N	N	nd	0,68%	N
38000	35	40	N	S	nd	N	N	nd	27,19%	N
38000	37	10207	S	S	?	S	S	?	92,71%	S
38000	39	27973	S	S	?	N	N	?	31,83%	N
38000	51	39	N	N	nd	N	N	nd	0,00%	N
38000	52	809	N	S	nd	N	N	nd	0,00%	N
39000	30	8309	S	S	S	N	S	N	74,77%	?
39000	34	77	S	S	nd	S	S	nd	47,43%	S
39000	35	12634	N	S	N	N	N	S	10,84%	?
39000	37	13384	S	S	?	S	S	?	78,65%	S
39000	39	98597	S	S	N	N	N	N	0,59%	N
39000	51	75293	N	S	N	N	N	N	0,00%	N
39000	52	516	N	S	nd	N	N	nd	0,01%	N
41000	30	771	N	N	nd	N	N	nd	0,00%	N
41000	37	401	N	S	nd	N	N	nd	1,07%	N
41000	39	20442	S	S	?	N	N	N	20,11%	N
41000	51	41	N	N	nd	N	N	nd	0,00%	N
41000	52	358	N	N	nd	N	N	nd	0,00%	N

42000	30	3895	N	S	N	N	N	N	3,06%	N
42000	32	251	N	S	nd	N	N	nd	0,00%	N
42000	35	5308	S	S	S	S	S	?	73,58%	S
42000	37	11813	S	S	?	N	N	?	73,68%	N
42000	39	28098	S	S	?	S	S	?	45,25%	S
42000	51	1322	N	S	N	N	N	N	2,07%	N
42000	52	796	N	N	nd	N	N	nd	0,00%	N
44000	30	3471	N	S	N	N	N	N	0,43%	N
44000	34	71	N	S	nd	N	N	nd	14,49%	N
44000	35	303	N	S	nd	N	N	nd	4,10%	N
44000	37	26939	S	S	S	S	S	?	120,05%	S
44000	39	44900	S	S	?	N	N	N	19,06%	N
44000	51	124	N	N	nd	N	N	nd	0,00%	N
44000	52	1252	N	S	N	N	N	N	4,82%	N
49000	30	5002	N	S	N	N	N	N	0,03%	N
49000	35	61	N	S	nd	N	N	nd	0,78%	N
49000	37	9951	S	S	?	N	N	N	28,99%	N
49000	39	24884	S	S	N	N	N	N	4,26%	N
49000	51	2678	N	N	N	N	N	N	0,00%	N
49000	52	569	N	N	nd	N	N	nd	0,00%	N
51000	30	203	N	S	nd	N	N	nd	0,00%	N
51000	35	319	N	S	nd	N	N	nd	0,14%	N
51000	37	580	S	S	nd	N	N	nd	1,27%	N
51000	39	4621	N	S	N	N	N	N	0,00%	N
51000	52	86	N	N	nd	N	N	nd	0,00%	N
52000	30	371786	S	S	N	N	N	N	5,04%	N
52000	32	2859	N	S	N	N	N	N	0,00%	N
52000	35	464	N	N	nd	N	N	nd	0,00%	N
52000	37	7936	S	S	N	N	N	N	2,38%	N
52000	39	510588	S	S	?	S	S	S	48,35%	S
52000	51	11901	N	S	N	N	N	N	0,00%	N
52000	52	68591	S	S	N	N	N	?	1,29%	N
53000	30	6197	S	S	N	N	N	?	7,50%	N
53000	34	744	N	S	nd	N	N	nd	1,57%	N
53000	35	1676	N	S	S	S	S	N	52,93%	?
53000	37	6702	S	S	?	N	N	?	106,25%	N
53000	39	22750	S	S	N	N	N	N	6,36%	N
53000	51	7817	N	N	N	N	N	N	0,00%	N
53000	52	889	N	N	nd	N	N	nd	0,00%	N
54000	30	249	N	N	nd	N	N	nd	0,00%	N
54000	35	883	S	S	nd	N	N	nd	25,11%	N
54000	37	436	S	S	nd	S	S	nd	8,89%	S
54000	39	4479	N	S	N	N	N	N	0,13%	N
55000	30	151	N	S	nd	N	N	nd	0,00%	N
55000	37	642	S	S	nd	S	S	nd	46,43%	S
55000	39	2967	S	S	N	N	N	N	5,71%	N
55000	52	39	N	N	nd	N	N	nd	0,00%	N
56000	30	3408	N	S	?	S	S	N	17,42%	?
56000	34	32	N	S	nd	N	N	nd	0,82%	N
56000	35	147	S	S	nd	N	N	nd	17,30%	N
56000	37	609	S	S	nd	N	N	nd	41,03%	N

56000	39	35858	S	S	N	N	N	N	28,90%	N
56000	51	2226	N	S	?	S	S	S	12,74%	S
56000	52	1728	N	S	N	N	N	N	0,00%	N
58000	30	523	N	N	nd	N	N	nd	0,44%	N
58000	37	758	S	S	nd	S	S	nd	18,71%	S
58000	39	2217	S	S	N	N	N	N	10,73%	N
58000	51	61	N	S	nd	N	N	nd	0,03%	N
58000	52	168	N	N	nd	N	N	nd	0,00%	N

Fonte: Elaborado pelo autor.

Tabela 5.6: Resultado dos Testes para Indícios de Erro ou Fraude – 2012

Órgão	Elem.	Qtd.Dsp.	Z- Teste	χ^2 - Teste	Param. d ₁ d ₂	Param. d ₁ /d ₂	Param. Total	Boot. Regr.	Razão	Final
20000	30	4798	N	S	N	N	N	N	0,00%	N
20000	35	133	N	S	nd	N	N	nd	0,00%	N
20000	37	7329	S	S	S	S	S	?	134,81%	S
20000	39	84405	S	S	?	N	N	N	46,87%	N
20000	51	123	N	N	nd	N	N	nd	0,00%	N
20000	52	1160	N	N	N	N	N	N	0,04%	N
20113	30	16348	S	S	S	S	S	?	121,90%	S
20113	35	263	N	S	nd	N	N	nd	0,31%	N
20113	37	7436	S	S	?	S	S	?	38,27%	S
20113	39	35183	S	S	N	S	S	N	17,13%	?
20113	51	68	N	N	nd	N	N	nd	0,00%	N
20113	52	392	N	N	nd	N	N	nd	0,00%	N
22000	30	51712	S	S	?	S	S	?	28,96%	S
22000	32	10287	S	S	?	N	N	N	12,35%	N
22000	35	149	N	S	nd	N	N	nd	0,00%	N
22000	37	21746	S	S	S	S	S	?	104,78%	S
22000	39	133167	S	S	S	S	S	?	58,98%	S
22000	51	3408	N	N	N	N	N	N	4,69%	N
22000	52	7259	N	N	N	N	N	N	0,00%	N
24000	30	22471	S	S	N	N	N	N	13,97%	N
24000	34	628	N	S	nd	N	N	nd	17,53%	N
24000	35	1872	S	S	?	S	S	?	86,99%	S
24000	37	9558	S	S	?	N	N	N	57,44%	N
24000	39	58988	S	S	?	S	S	?	13,18%	S
24000	51	799	N	S	nd	N	N	nd	0,12%	N
24000	52	4784	S	S	?	S	S	N	81,51%	?
25000	30	34029	S	S	?	N	N	N	30,54%	N
25000	34	2960	S	S	N	N	N	N	0,53%	N
25000	35	420	N	S	nd	N	N	nd	1,73%	N
25000	37	52907	S	S	?	N	N	?	117,33%	N
25000	39	194811	S	S	?	N	N	N	34,01%	N
25000	51	2230	N	N	N	N	N	N	0,08%	N
25000	52	5519	N	S	N	N	N	N	0,00%	N
26000	30	418137	S	S	?	N	N	N	27,14%	N
26000	32	4507	N	S	N	N	N	N	0,01%	N
26000	34	87	S	S	nd	N	N	nd	0,03%	N

26000	35	1519	S	S	N	N	N	N	11,72%	N
26000	37	134403	S	S	S	S	S	?	133,19%	S
26000	39	409343	S	S	N	S	S	?	16,77%	S
26000	51	41481	N	S	N	N	N	N	8,24%	N
26000	52	121527	S	S	N	S	S	N	2,01%	?
28000	30	1316	N	S	N	N	N	N	0,35%	N
28000	35	31	N	S	nd	N	N	nd	11,47%	N
28000	37	3562	S	S	N	N	N	N	15,00%	N
28000	39	13360	S	S	?	S	S	N	33,10%	?
28000	51	92	N	N	nd	N	N	nd	0,09%	N
28000	52	523	N	N	nd	N	N	nd	0,00%	N
30000	30	25667	N	S	?	N	N	N	11,17%	N
30000	35	83	N	S	nd	N	N	nd	1,56%	N
30000	37	35738	S	S	?	S	S	S	75,66%	S
30000	39	88652	S	S	N	S	S	N	1,59%	?
30000	51	1026	N	N	N	N	N	N	1,39%	N
30000	52	4374	N	S	N	N	N	N	12,34%	N
32000	30	10753	S	S	N	S	S	S	15,55%	S
32000	35	583	S	S	nd	S	S	nd	16,98%	S
32000	37	11927	S	S	?	S	S	?	44,89%	S
32000	39	43964	S	S	N	N	N	N	3,00%	N
32000	51	283	N	N	nd	N	N	nd	0,00%	N
32000	52	1550	N	S	N	N	N	N	0,00%	N
33000	30	6469	N	S	N	N	N	N	0,00%	N
33000	32	1175	S	S	N	N	N	N	4,77%	N
33000	35	61	S	S	nd	N	N	nd	27,09%	N
33000	37	46002	S	S	?	S	S	S	105,03%	S
33000	39	104464	S	S	S	S	S	?	397,44%	S
33000	51	2163	N	S	?	N	N	?	31,52%	N
33000	52	2398	S	S	N	N	N	N	2,08%	N
35000	30	940	N	S	nd	N	N	nd	0,00%	N
35000	35	38	N	S	nd	N	N	nd	0,27%	N
35000	37	84	S	S	nd	N	N	nd	1,81%	N
35000	39	7471	N	S	?	N	N	N	8,52%	N
35000	52	206	N	N	nd	N	N	nd	0,00%	N
36000	30	152863	S	S	?	N	N	N	53,58%	N
36000	32	72	N	S	nd	N	N	nd	0,00%	N
36000	34	1242	N	S	N	N	N	N	1,30%	N
36000	35	563	N	S	nd	N	N	nd	14,83%	N
36000	37	27748	S	S	?	S	S	?	58,32%	S
36000	39	120292	S	S	?	N	N	N	12,52%	N
36000	51	2270	N	S	N	N	N	N	0,01%	N
36000	52	5242	N	S	N	N	N	N	0,45%	N
38000	30	5032	N	S	N	N	N	N	0,00%	N
38000	34	65	N	S	nd	N	N	nd	17,56%	N
38000	35	52	S	S	nd	N	N	nd	25,78%	N
38000	37	13480	S	S	?	N	N	?	101,97%	N
38000	39	26011	S	S	?	S	S	?	9,71%	S
38000	52	732	N	N	nd	N	N	nd	0,00%	N
39000	30	6837	S	S	?	S	S	?	71,22%	S
39000	34	144	S	S	nd	S	S	nd	74,88%	S

39000	35	11526	N	S	N	N	N	N	16,90%	N
39000	37	14608	S	S	?	S	S	S	75,45%	S
39000	39	127185	N	S	N	N	N	N	18,59%	N
39000	51	44594	N	S	N	N	N	N	0,00%	N
39000	52	1094	N	S	N	N	N	N	0,00%	N
41000	30	784	N	N	nd	N	N	nd	0,00%	N
41000	35	35	N	S	nd	N	N	nd	1,97%	N
41000	37	889	S	S	nd	S	S	nd	57,27%	S
41000	39	19423	S	S	?	N	N	N	38,88%	N
41000	51	41	N	N	nd	N	N	nd	0,00%	N
41000	52	348	N	S	nd	N	N	nd	0,00%	N
42000	30	3785	N	S	N	N	N	N	3,70%	N
42000	32	347	N	N	nd	N	N	nd	0,00%	N
42000	35	3713	S	S	?	S	S	S	33,71%	S
42000	37	13173	S	S	?	N	N	N	77,61%	N
42000	39	27789	S	S	?	N	N	N	23,81%	N
42000	51	581	N	N	nd	N	N	nd	0,00%	N
42000	52	635	N	S	nd	N	N	nd	2,29%	N
44000	30	3634	N	S	N	N	N	N	0,00%	N
44000	34	108	N	S	nd	S	S	nd	18,54%	S
44000	35	214	N	S	nd	N	N	nd	0,03%	N
44000	37	29109	S	S	?	S	S	S	122,83%	S
44000	39	44881	S	S	?	N	N	?	15,25%	N
44000	51	163	N	N	nd	N	N	nd	0,00%	N
44000	52	1485	N	S	N	N	N	N	0,00%	N
49000	30	4054	N	S	N	N	N	N	0,18%	N
49000	35	61	N	N	nd	N	N	nd	5,41%	N
49000	37	10454	S	S	?	N	N	N	28,12%	N
49000	39	23936	N	S	N	N	N	?	6,23%	N
49000	51	1336	N	N	N	N	N	N	0,00%	N
49000	52	455	N	N	nd	N	N	nd	0,00%	N
51000	30	127	N	N	nd	N	N	nd	0,00%	N
51000	35	175	N	S	nd	N	N	nd	1,34%	N
51000	37	567	S	S	nd	N	N	nd	12,64%	N
51000	39	3109	N	S	N	N	N	N	15,13%	N
51000	52	49	N	N	nd	N	N	nd	0,00%	N
52000	30	377420	S	S	?	N	N	N	5,95%	N
52000	32	2468	N	S	N	N	N	N	0,02%	N
52000	35	379	N	S	nd	N	N	nd	3,64%	N
52000	37	7441	S	S	?	N	N	N	10,37%	N
52000	39	534104	S	S	N	S	S	?	24,77%	S
52000	51	12179	N	N	N	N	N	N	0,00%	N
52000	52	71758	S	S	N	N	N	?	1,31%	N
53000	30	6376	S	S	N	N	N	?	8,29%	N
53000	34	983	S	S	nd	S	S	nd	79,23%	S
53000	35	1284	N	S	N	N	N	N	0,23%	N
53000	37	7516	S	S	S	S	S	?	105,07%	S
53000	39	24499	S	S	N	N	N	N	14,20%	N
53000	51	7279	N	S	N	N	N	N	0,38%	N
53000	52	1385	N	S	N	N	N	N	0,00%	N
54000	30	241	N	N	nd	N	N	nd	0,00%	N

54000	35	1073	N	N	N	N	N	N	1,51%	N
54000	37	392	S	S	nd	N	N	nd	5,36%	N
54000	39	5756	N	S	N	N	N	N	0,10%	N
54000	52	56	N	N	nd	N	N	nd	0,00%	N
55000	30	143	N	S	nd	N	N	nd	1,01%	N
55000	37	449	S	S	nd	N	N	nd	1,27%	N
55000	39	7328	N	S	?	S	S	?	26,76%	S
55000	52	48	N	N	nd	N	N	nd	0,00%	N
56000	30	3590	N	S	N	N	N	N	0,24%	N
56000	35	128	S	S	nd	N	N	nd	0,20%	N
56000	37	639	S	S	nd	N	N	nd	74,52%	N
56000	39	45246	S	S	?	S	S	S	102,96%	S
56000	51	2014	S	S	?	S	S	S	30,07%	S
56000	52	1474	S	S	N	N	N	N	10,02%	N
58000	30	614	N	S	nd	N	N	nd	4,84%	N
58000	37	1154	S	S	?	N	N	N	25,05%	N
58000	39	2590	N	S	N	N	N	N	0,04%	N
58000	52	251	N	S	nd	N	N	nd	2,18%	N

Fonte: Elaborado pelo autor.

Tabela 5.7: Resultado dos Testes para Índícios de Erro ou Fraude – 2013

Órgão	Elem.	Qtd.Dsp.	Z- Teste	χ^2 - Teste	Param. d_1/d_2	Param. d_1/d_2	Param. Total	Boot. Regr.	Razão	Final
20000	30	4582	N	S	N	N	N	N	0,01%	N
20000	35	320	N	S	nd	N	N	nd	3,04%	N
20000	37	6666	S	S	?	N	N	?	31,15%	N
20000	39	85434	S	S	?	N	N	N	44,27%	N
20000	51	111	N	S	nd	N	N	nd	0,00%	N
20000	52	1544	N	S	N	N	N	N	0,77%	N
20113	30	16918	S	S	?	S	S	S	94,28%	S
20113	35	472	S	S	nd	N	N	nd	4,04%	N
20113	37	6256	S	S	N	N	N	N	8,45%	N
20113	39	32428	S	S	?	S	S	N	27,00%	?
20113	51	74	N	N	nd	N	N	nd	0,00%	N
20113	52	352	N	N	nd	N	N	nd	0,00%	N
22000	30	46435	S	S	?	S	S	?	56,21%	S
22000	32	11926	S	S	?	N	N	N	53,98%	N
22000	35	171	N	S	nd	N	N	nd	1,72%	N
22000	37	23670	S	S	?	S	S	S	66,52%	S
22000	39	134225	S	S	?	N	N	N	15,65%	N
22000	51	3077	N	N	N	N	N	N	0,00%	N
22000	52	7037	N	S	N	N	N	N	2,95%	N
24000	30	21441	S	S	?	N	N	N	31,74%	N
24000	34	263	N	S	nd	N	N	nd	9,36%	N
24000	35	1463	S	S	?	N	N	?	60,25%	N
24000	37	9802	S	S	?	N	N	N	9,55%	N
24000	39	58536	S	S	S	S	S	?	266,01%	S
24000	51	648	N	S	nd	N	N	nd	0,00%	N
24000	52	3935	N	S	?	N	N	?	35,82%	N

25000	30	30863	S	S	S	N	S	?	94,88%	S
25000	34	1336	S	S	N	N	N	N	0,00%	N
25000	35	576	S	S	nd	N	N	nd	3,00%	N
25000	37	57921	S	S	?	N	N	?	95,79%	N
25000	39	190786	S	S	?	S	S	?	25,47%	S
25000	51	2146	N	N	N	N	N	N	0,26%	N
25000	52	4534	N	S	N	N	N	N	0,00%	N
26000	30	435008	S	S	N	S	S	N	5,96%	?
26000	32	6475	S	S	?	N	N	N	28,28%	N
26000	34	84	S	S	nd	N	N	nd	0,86%	N
26000	35	1529	S	S	N	N	N	N	8,59%	N
26000	37	153752	S	S	?	S	S	N	57,27%	?
26000	39	433558	S	S	N	S	S	?	7,86%	S
26000	51	38037	N	S	N	N	N	N	0,02%	N
26000	52	109484	S	S	N	N	N	N	2,67%	N
28000	30	1273	N	N	N	N	N	N	0,00%	N
28000	35	33	N	S	nd	N	N	nd	7,31%	N
28000	37	3611	S	S	?	S	S	?	30,83%	S
28000	39	13887	S	S	?	N	N	N	40,00%	N
28000	51	108	N	N	nd	N	N	nd	0,00%	N
28000	52	477	N	S	nd	N	N	nd	0,00%	N
30000	30	25089	S	S	N	N	N	N	16,64%	N
30000	35	63	N	S	nd	N	N	nd	2,25%	N
30000	37	41808	S	S	N	N	N	?	22,48%	N
30000	39	89036	S	S	?	S	S	S	18,10%	S
30000	51	1489	N	N	N	N	N	N	0,00%	N
30000	52	4198	N	S	N	N	N	N	0,00%	N
32000	30	9718	S	S	?	N	N	?	23,43%	N
32000	35	535	S	S	nd	N	N	nd	13,44%	N
32000	37	12463	S	S	?	S	S	S	90,13%	S
32000	39	39735	S	S	N	S	S	?	6,99%	S
32000	51	355	N	S	nd	N	N	nd	0,00%	N
32000	52	1461	N	S	N	N	N	N	10,00%	N
33000	30	5986	N	S	N	N	N	N	0,00%	N
33000	32	740	N	N	nd	N	N	nd	0,00%	N
33000	35	64	N	S	nd	N	N	nd	13,93%	N
33000	37	48404	S	S	S	S	S	?	151,82%	S
33000	39	114125	S	S	?	S	S	?	188,04%	S
33000	51	3051	N	S	?	N	N	N	34,36%	N
33000	52	1741	N	S	N	N	N	N	0,02%	N
35000	30	731	N	S	nd	N	N	nd	0,02%	N
35000	35	32	N	S	nd	N	N	nd	0,00%	N
35000	37	92	S	S	nd	N	N	nd	1,91%	N
35000	39	7261	N	S	N	N	N	N	0,17%	N
35000	51	59	N	N	nd	N	N	nd	0,00%	N
35000	52	181	N	N	nd	N	N	nd	0,00%	N
36000	30	138179	S	S	?	N	N	N	30,87%	N
36000	34	1490	N	S	S	N	S	N	39,84%	?
36000	35	888	N	S	nd	N	N	nd	1,72%	N
36000	37	31260	S	S	?	S	S	S	49,25%	S
36000	39	119945	S	S	?	N	N	N	22,05%	N

36000	51	2869	N	N	N	N	N	N	0,00%	N
36000	52	5810	N	S	N	N	N	N	0,00%	N
38000	30	4732	N	S	N	N	N	N	0,53%	N
38000	37	15739	S	S	?	N	N	N	34,68%	N
38000	39	24780	S	S	?	N	N	N	27,04%	N
38000	52	496	N	S	nd	N	N	nd	0,00%	N
39000	30	7069	S	S	?	N	N	N	35,70%	N
39000	34	73	N	S	nd	N	N	nd	1,31%	N
39000	35	6577	N	S	N	N	N	N	1,19%	N
39000	37	13216	S	S	?	S	S	S	94,32%	S
39000	39	167603	S	S	?	S	S	N	36,98%	?
39000	51	23504	N	S	N	N	N	N	1,08%	N
39000	52	922	N	S	nd	N	N	nd	0,05%	N
41000	30	573	N	S	nd	N	N	nd	0,00%	N
41000	35	43	N	S	nd	N	N	nd	1,98%	N
41000	37	415	N	S	nd	N	N	nd	0,86%	N
41000	39	17931	S	S	?	N	N	N	21,19%	N
41000	52	412	N	S	nd	N	N	nd	0,01%	N
42000	30	3384	N	S	?	N	N	N	4,35%	N
42000	35	2906	S	S	?	S	S	?	36,55%	S
42000	37	15231	S	S	?	N	N	?	117,50%	N
42000	39	27352	S	S	?	N	N	N	28,21%	N
42000	51	483	N	N	nd	N	N	nd	3,51%	N
42000	52	619	N	S	nd	N	N	nd	0,00%	N
44000	30	3112	N	S	N	N	N	N	1,04%	N
44000	34	626	S	S	nd	S	S	nd	71,91%	S
44000	35	228	N	S	nd	N	N	nd	0,29%	N
44000	37	30534	S	S	?	S	S	S	147,82%	S
44000	39	42397	S	S	N	S	S	?	14,37%	S
44000	51	656	N	N	nd	N	N	nd	0,00%	N
44000	52	1112	N	S	N	N	N	N	0,51%	N
49000	30	4028	N	S	N	N	N	N	0,00%	N
49000	35	59	N	S	nd	N	N	nd	6,33%	N
49000	37	10852	S	S	?	N	N	N	9,72%	N
49000	39	25331	S	S	N	N	N	N	2,41%	N
49000	51	989	N	N	nd	N	N	nd	0,00%	N
49000	52	683	N	N	nd	N	N	nd	0,00%	N
51000	30	72	N	N	nd	N	N	nd	0,00%	N
51000	35	244	N	S	nd	N	N	nd	0,50%	N
51000	37	597	S	S	nd	N	N	nd	0,67%	N
51000	39	2849	N	S	N	N	N	N	1,91%	N
51000	52	33	N	N	nd	N	N	nd	0,00%	N
52000	30	377788	S	S	?	S	S	N	14,96%	?
52000	32	2439	N	S	N	N	N	N	0,00%	N
52000	35	345	N	S	nd	N	N	nd	0,73%	N
52000	37	8092	S	S	N	N	N	N	2,00%	N
52000	39	548994	S	S	?	S	S	?	36,08%	S
52000	51	12309	N	N	N	N	N	N	2,46%	N
52000	52	71488	S	S	N	N	N	N	0,68%	N
53000	30	6750	S	S	?	N	N	?	30,80%	N
53000	32	64	N	N	nd	N	N	nd	0,00%	N

22000	52	6879	S	S	N	N	N	N	0,03%	N
24000	30	18851	S	S	N	N	N	N	9,17%	N
24000	34	205	S	S	nd	N	N	nd	29,09%	N
24000	35	1468	N	S	S	N	S	N	47,46%	?
24000	37	10256	S	S	?	N	N	N	21,18%	N
24000	39	57971	S	S	?	N	N	N	17,87%	N
24000	51	559	N	S	nd	N	N	nd	0,33%	N
24000	52	4312	N	S	?	N	N	N	39,49%	N
25000	30	24900	N	S	N	N	N	N	0,00%	N
25000	34	1303	S	S	N	N	N	N	3,85%	N
25000	35	406	N	S	nd	N	N	nd	0,15%	N
25000	37	60152	S	S	?	S	S	?	148,09%	S
25000	39	204514	S	S	N	S	S	?	14,19%	S
25000	51	3198	N	S	N	N	N	N	1,55%	N
25000	52	3928	N	S	N	N	N	N	1,12%	N
26000	30	460109	S	S	N	N	N	?	1,90%	N
26000	32	6713	N	S	N	N	N	N	14,73%	N
26000	34	56	S	S	nd	N	N	nd	8,28%	N
26000	35	1522	S	S	?	N	N	N	43,26%	N
26000	37	188957	S	S	?	N	N	?	112,13%	N
26000	39	473702	S	S	N	S	S	?	16,68%	S
26000	51	42409	N	S	N	N	N	N	0,46%	N
26000	52	120132	S	S	N	N	N	N	2,47%	N
28000	30	1209	N	N	N	N	N	N	0,00%	N
28000	35	36	N	S	nd	N	N	nd	0,16%	N
28000	37	4311	S	S	?	S	S	S	76,17%	S
28000	39	14204	N	S	?	N	N	?	76,19%	N
28000	51	91	N	N	nd	N	N	nd	0,00%	N
28000	52	416	N	S	nd	N	N	nd	0,49%	N
30000	30	28124	S	S	?	N	N	?	97,29%	N
30000	35	38	N	S	nd	N	N	nd	0,00%	N
30000	37	37242	S	S	S	N	S	?	165,36%	S
30000	39	80057	S	S	N	N	N	N	10,74%	N
30000	51	1087	N	N	N	N	N	N	16,84%	N
30000	52	6644	S	S	N	N	N	N	14,75%	N
32000	30	9077	S	S	N	S	S	?	16,34%	S
32000	35	500	N	S	nd	N	N	nd	18,25%	N
32000	37	12701	S	S	?	N	N	?	93,62%	N
32000	39	38816	S	S	?	N	N	N	18,84%	N
32000	51	497	N	S	nd	N	N	nd	0,00%	N
32000	52	1820	N	S	N	N	N	N	0,08%	N
33000	30	6234	N	S	N	N	N	N	0,14%	N
33000	32	1008	N	S	N	N	N	N	0,00%	N
33000	37	54937	S	S	S	N	S	?	232,69%	S
33000	39	115133	S	S	?	S	S	S	187,18%	S
33000	51	2176	N	S	?	N	N	N	20,61%	N
33000	52	3352	S	S	N	N	N	N	0,55%	N
35000	30	635	N	S	nd	N	N	nd	0,00%	N
35000	35	53	N	S	nd	N	N	nd	0,89%	N
35000	37	201	S	S	nd	N	N	nd	8,34%	N
35000	39	7371	S	S	S	S	S	?	87,01%	S

35000	51	39	N	N	nd	N	N	nd	0,00%	N
35000	52	124	N	S	nd	N	N	nd	0,00%	N
36000	30	145010	S	S	?	S	S	?	24,89%	S
36000	32	95	N	N	nd	N	N	nd	0,00%	N
36000	34	1146	N	S	N	N	N	N	5,00%	N
36000	35	1358	N	S	N	N	N	N	19,12%	N
36000	37	34582	S	S	?	S	S	?	69,88%	S
36000	39	113303	S	S	N	N	N	?	4,77%	N
36000	51	3487	N	N	N	N	N	N	0,00%	N
36000	52	6070	N	S	N	N	N	N	0,00%	N
38000	30	5260	N	S	?	N	N	N	39,13%	N
38000	37	17337	S	S	S	S	S	?	114,56%	S
38000	39	24655	S	S	?	N	N	N	24,13%	N
38000	51	31	N	N	nd	N	N	nd	0,00%	N
38000	52	448	N	N	nd	N	N	nd	0,00%	N
39000	30	8845	S	S	?	N	N	?	93,27%	N
39000	34	43	N	S	nd	N	N	nd	0,00%	N
39000	35	5065	N	S	N	N	N	N	0,00%	N
39000	37	16386	S	S	S	N	S	?	216,50%	S
39000	39	172573	S	S	N	N	N	N	6,34%	N
39000	51	18043	S	S	N	N	N	N	0,33%	N
39000	52	830	N	S	nd	N	N	nd	7,20%	N
41000	30	590	N	S	nd	N	N	nd	0,00%	N
41000	35	32	N	S	nd	N	N	nd	2,67%	N
41000	37	664	S	S	nd	N	N	nd	18,35%	N
41000	39	16789	S	S	?	N	N	N	42,19%	N
41000	52	309	N	S	nd	N	N	nd	0,14%	N
42000	30	3346	N	S	N	N	N	N	5,06%	N
42000	35	2564	S	S	?	S	S	?	22,33%	S
42000	37	17678	S	S	?	N	N	?	127,07%	N
42000	39	28895	S	S	?	N	N	N	54,75%	N
42000	51	601	N	N	nd	N	N	nd	0,00%	N
42000	52	871	N	S	nd	N	N	nd	0,10%	N
44000	30	2402	N	S	N	N	N	N	2,97%	N
44000	35	253	N	S	nd	N	N	nd	0,75%	N
44000	37	30357	S	S	?	S	S	?	244,04%	S
44000	39	45876	S	S	?	N	N	N	36,15%	N
44000	51	512	N	S	nd	N	N	nd	0,00%	N
44000	52	884	N	S	nd	N	N	nd	8,80%	N
49000	30	3330	N	S	N	N	N	N	0,00%	N
49000	35	96	N	N	nd	N	N	nd	0,00%	N
49000	37	10239	S	S	S	N	S	?	149,42%	S
49000	39	31509	S	S	?	S	S	?	46,56%	S
49000	51	1132	N	N	N	N	N	N	0,00%	N
49000	52	1074	N	N	N	N	N	N	0,04%	N
51000	30	106	N	N	nd	N	N	nd	0,00%	N
51000	35	289	N	S	nd	N	N	nd	1,50%	N
51000	37	640	S	S	nd	N	N	nd	0,14%	N
51000	39	2629	N	S	N	N	N	N	0,35%	N
52000	30	396539	S	S	N	S	S	?	5,34%	S
52000	32	2280	N	S	N	N	N	N	0,11%	N

52000	35	982	N	S	nd	N	N	nd	0,00%	N
52000	37	8947	S	S	N	N	N	N	7,33%	N
52000	39	569729	S	S	?	N	N	N	36,63%	N
52000	51	12953	N	N	N	N	N	N	0,00%	N
52000	52	75124	S	S	N	N	N	N	3,46%	N
53000	30	5789	S	S	?	N	N	?	23,04%	N
53000	32	31	N	S	nd	N	N	nd	0,00%	N
53000	35	767	N	S	nd	N	N	nd	7,25%	N
53000	37	8559	S	S	?	S	S	S	64,70%	S
53000	39	26671	S	S	?	S	S	S	23,09%	S
53000	51	9777	N	S	N	N	N	N	0,00%	N
53000	52	1795	N	N	N	N	N	N	0,00%	N
54000	30	194	N	N	nd	N	N	nd	0,00%	N
54000	35	985	S	S	nd	N	N	nd	7,13%	N
54000	37	602	S	S	nd	N	N	nd	28,26%	N
54000	39	6552	S	S	N	N	N	N	10,60%	N
54000	52	54	N	S	nd	N	N	nd	0,00%	N
55000	30	225	N	S	nd	N	N	nd	0,24%	N
55000	32	31	N	N	nd	N	N	nd	0,00%	N
55000	35	87	N	S	nd	N	N	nd	3,68%	N
55000	37	383	S	S	nd	S	S	nd	0,32%	S
55000	39	4365	N	S	N	N	N	N	9,48%	N
55000	52	126	N	S	nd	N	N	nd	11,85%	N
56000	30	3435	N	S	N	N	N	N	8,51%	N
56000	35	154	S	S	nd	S	S	nd	21,55%	S
56000	37	1730	S	S	S	S	S	?	107,82%	S
56000	39	32518	N	S	N	N	N	N	8,62%	N
56000	51	793	N	S	nd	N	N	nd	0,33%	N
56000	52	1119	N	S	N	N	N	N	1,06%	N
58000	30	598	N	S	nd	N	N	nd	0,05%	N
58000	37	1941	S	S	N	N	N	N	11,01%	N
58000	39	3169	N	S	N	N	N	N	4,87%	N
58000	52	166	N	N	nd	N	N	nd	0,00%	N

Fonte: Elaborado pelo autor.

Apresentamos, então, no quadro a seguir, os resultados referentes à coluna 'Final' de forma esquemática, para facilitar a visualização da distribuição dos indícios de erro ou fraude sinalizados ao longo dos órgãos, elementos e anos em que a despesa foi realizada.

Quadro 5.1: Resumo dos Resultados

Órgão / Elem.	2011											2012										
	52	51	39	37	35	34	32	30	Qtd.	52	51	39	37	35	34	32	30	Qtd.				
			X	X					2				X	X				1				
20000									0									—				
20102									2									—				
20113					X			X	2							X		2				
22000			X	X				X	3								X	3				
24000				X					1					X				2				
25000			X	X					2									0				
26000			X						1			X	X					2				
28000									0									0				
30000									0			X	X					1				
32000			X	X	X		X		4					X		X		3				
33000			X	X					2			X	X					2				
35000									0									0				
36000				X				X	2				X					1				
38000				X					1			X						1				
39000				X		X			2				X		X		X	3				
41000									0				X					1				
42000			X						2					X				1				
44000				X					1				X		X			2				
49000									0									0				
51000									0									0				
52000			X						1			X						1				
53000									0					X				2				
54000				X					1									0				
55000				X					1			X						1				
56000		X							1		X							2				
58000				X					1									0				
Total	0	1	8	13	3	1	0	4	30	0	1	8	12	3	3	0	4	31				

Fonte: Elaborado pelo autor.

Órgão / Elem.	2013												2014												Qtd. Total
	52	51	39	37	35	34	32	30	Qtd.	52	51	39	37	35	34	32	30	Qtd.							
20000									0										0	3					
20102									—										—	0					
20113							X		1										1	6					
22000				X				X	2			X							3	11					
24000									1										0	4					
25000			X				X		2		X		X						2	6					
26000			X						1		X								1	5					
28000				X					1			X							1	2					
30000			X						1		X		X						1	3					
32000			X	X					2		X						X		1	10					
33000			X	X					2		X		X						2	8					
35000									0		X								1	1					
36000				X					1				X						2	6					
38000									0				X						1	3					
39000				X					1			X							1	7					
41000									0										0	1					
42000					X				1					X					1	5					
44000			X	X		X			3				X						1	7					
49000									0		X		X						2	2					
51000									0										0	0					
52000			X						1								X		1	4					
53000				X					1				X						2	5					
54000									0										0	1					
55000									0				X						1	3					
56000			X						1				X						2	6					
58000									0										0	1					
Total	0	0	9	8	1	1	0	3	22	0	0	7	13	2	0	0	5	27	110						

Fonte: Elaborado pelo autor.

Podemos verificar que os Ministérios da Agricultura, Pesca e Abastecimento (22000) e de Minas e Energia (32000) apresentam as maiores quantidades de sinalizações. O Ministério do Esporte (51000), por sua vez, não apresentou nenhuma sinalização ao longo dos quatro anos. Em relação aos elementos de despesa, a grande maioria das

sinalizações ocorreu na Locação de Mão de Obra (37) e Outros Serviços de Terceiros – PJ (39). Do total de 808 combinações, os resultados sugerem indícios de erro ou fraude em 110 delas, ou seja, em 13,61% do total.

5.5 Comparação dos Resultados dos Testes

A partir dos resultados dos diferentes testes, elaboramos uma matriz de confusão³ a fim de verificar o nível de acurácia entre os testes, indicado pelo percentual dado na linha VP/VN (Verdadeiro Positivo/Verdadeiro Negativo). Então, assumindo que o resultado dado pelo teste indicado nas colunas seja o resultado verdadeiro, calculamos as proporções de Falso Positivo (FP) e de Falso Negativo (FN) incorridas pelo teste indicado nas linhas. Um valor elevado de FP significa que o teste indicado na linha é mais sensível que aquele indicado na coluna. O contrário é verdadeiro quando encontramos um valor elevado de FN na tabela.

Tabela 5.9: Matriz de Confusão

Teste	Result.	Regressão			Param.d ₁ d ₂			Param.d ₁ /d ₂		χ ² -Teste	
		NÃO	?	SIM	NÃO	?	SIM	NÃO	SIM	NÃO	SIM
Z-Teste	NÃO	146	7	2	135	17	3	335	6	101	240
	SIM	114	92	30	75	130	31	167	112	0	279
Z-Teste	VP/VN	45.0%			42.5%			72.1%		61.3%	
	FP/FN	52.7%	0.5%		52.4%	0.8%		26.9%	1.0%	0.0%	38.7%
χ ² -Teste	NÃO	31	0	0	31	0	0	101	0		
	SIM	229	99	32	179	147	34	401	118		
χ ² -Teste	VP/VN	16.1%			16.6%			35.3%			
	FP/FN	83.9%	0.0%		83.4%	0.0%		64.7%	0.0%		
Param. d ₁ /d ₂	NÃO	245	40	2	193	84	10				
	SIM	15	59	30	17	63	24				
Param. d ₁ /d ₂	VP/VN	70.3%			55.5%						
	FP/FN	18.9%	0.5%		20.5%	2.6%					
Param. d ₁ d ₂	NÃO	186	20	4							
	INDEF.	69	53	25							
	SIM	5	26	3							

³ A matriz de confusão é uma tabela de contingência em que na linha está o valor previsto e na coluna o valor observado (valor verdadeiro).

VP/VN	61.9%
FP/FN	7.9% 7.4%

Fonte: Elaborado pelo autor.

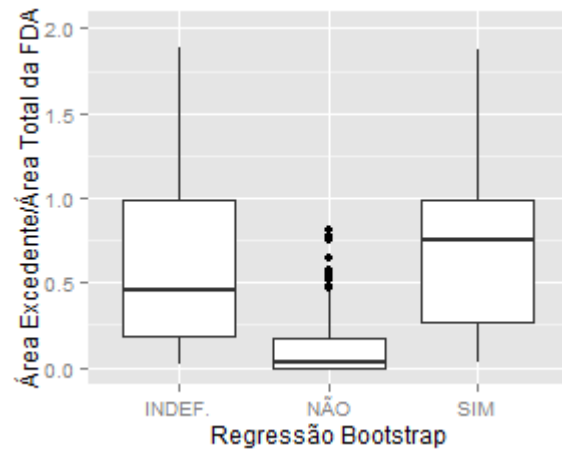
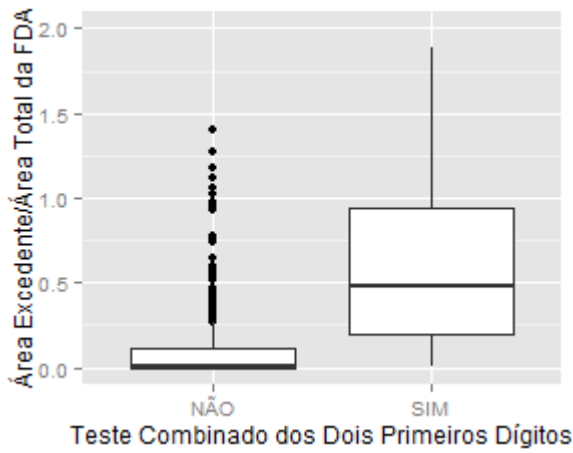
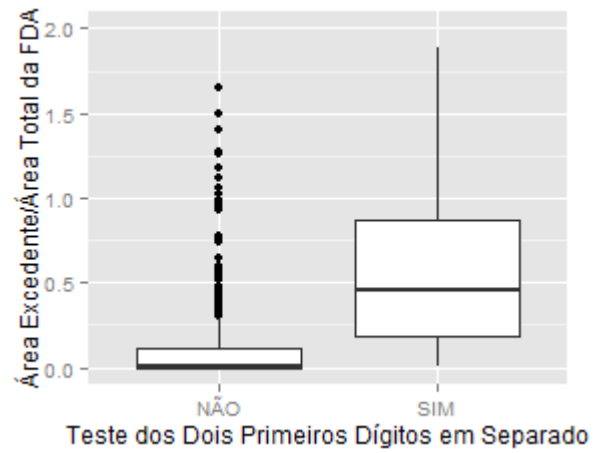
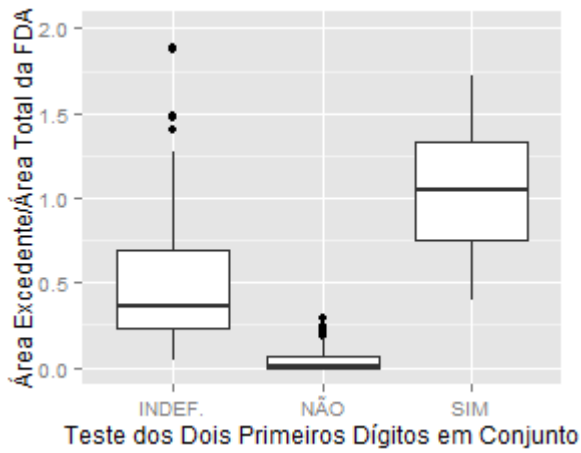
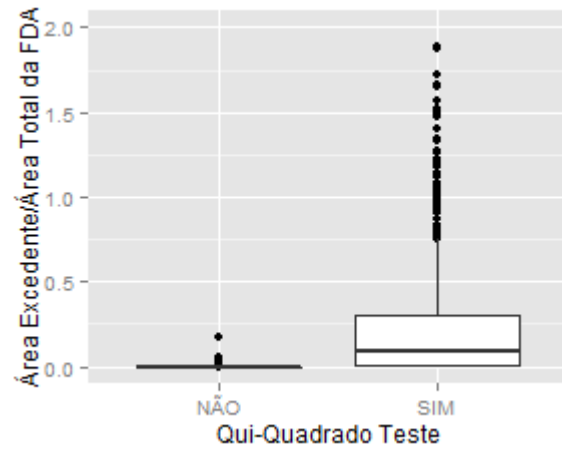
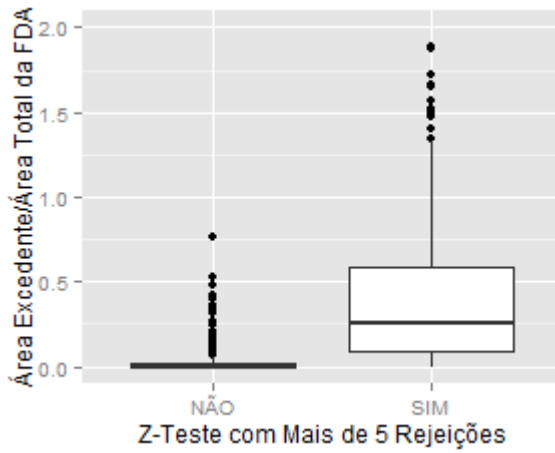
Os resultados da comparação tornam evidente o maior poder do χ^2 -Teste em relação a todos os demais e o seu baixo nível de acurácia com os resultados dos dois testes mais robustos, quais sejam, regressão *bootstrap* e teste dos parâmetros de Benford dos dois primeiros dígitos em conjunto (d_1d_2). O Z-Teste, quando considerado o mínimo de 5 rejeições, por sua vez, mostrou um nível de acurácia mais elevado, mas também se mostrou muito mais poderoso que os testes que utilizam a técnica *bootstrap*. Desses, o teste dos parâmetros de Benford dos dois primeiros dígitos em separado (d_1/d_2) se mostrou aquele com mais poder e, a regressão *bootstrap*, aquele com menos poder.

5.6 Resultados dos Testes vs Razão Entre Áreas da *fda*

As características relativas ao poder dos testes elencadas na seção anterior também podem ser observadas nos gráficos *boxplot*⁴ que demonstram a relação entre os resultados dos diferentes testes e as razões entre áreas da *fda* conforme apresentado na Seção 4.3. Quanto maior o poder do teste, menor a diferença entre os ‘corpos’ associados aos diferentes resultados no gráfico e também menor o valor da métrica associado ao resultado positivo.

Figura 5.1: Gráficos *Boxplot* da Razão Entre Áreas por Resultado do Teste

⁴ Nesse tipo de gráfico, a linha horizontal representa a mediana; os lados inferior e superior do ‘corpo’ representam, respectivamente, o primeiro e terceiro quartis; os “pavios” se estendem até os valores máximo e mínimo (não considerados outliers); e os pontos representam os pontos considerados outliers.



Fonte: Elaborado pelo autor.

O Teste dos Dois Primeiros Dígitos em Conjunto (d_1d_2) é particularmente ilustrativo, dado que parece haver uma clara distinção entre as regiões referentes aos resultados 'NÃO' e 'SIM', enquanto a região intermediária é, coerentemente, associada ao resultado 'INDEF'. Os valores exatos representados nos gráficos *boxplot* são apresentados na tabela a seguir:

Tabela 5.9: Estatísticas Descritivas da Razão Entre Áreas por Teste

Z-Teste						
Resultado	Min	1Q	Mediana	Média	3Q	Max
NÃO	0,0000	0,0000	0,0004	0,0377	0,0232	0,7619
SIM	0,0000	0,0963	0,2523	0,4614	0,6561	4,7790
χ^2 -Teste						
Resultado	Min	1Q	Mediana	Média	3Q	Max
NÃO	0,0000	0,0000	0,0000	0,0037	0,0000	0,1684
SIM	0,0000	0,0049	0,0862	0,2721	0,3133	4,7790
Teste dos Dois Primeiros Dígitos em Conjunto						
Resultado	Min	1Q	Mediana	Média	3Q	Max
NÃO	0,0000	0,0000	0,0110	0,0434	0,0694	0,2890
INDEF.	0,0436	0,2362	0,3570	0,5006	0,7055	2,4400
SIM	0,3984	0,8041	1,1120	1,3630	1,5580	4,7790
Teste dos Dois Primeiros Dígitos em Separado						
Resultado	Min	1Q	Mediana	Média	3Q	Max
NÃO	0,0000	0,0000	0,0124	0,1233	0,1209	2,3270
SIM	0,0032	0,1896	0,4699	0,6753	0,9373	4,7790
Teste Combinado dos Dois Primeiros Dígitos						
Resultado	Min	1Q	Mediana	Média	3Q	Max
NÃO	0,0000	0,0000	0,0106	0,1005	0,1087	1,4020
SIM	0,0032	0,2146	0,5109	0,7198	0,9754	4,7790
Regressão Bootstrap						
Resultado	Min	1Q	Mediana	Média	3Q	Max
NÃO	0,0000	0,0004	0,0346	0,1164	0,1777	0,8151
INDEF.	0,0129	0,2241	0,5621	0,7173	1,0490	3,9740
SIM	0,0301	0,2858	0,7592	0,8670	1,0050	4,7790

Fonte: Elaborado pelo autor.

Na tabela acima, os realces em vermelho indicam os valores onde não há intersecção entre o 3º quartil de um resultado e o 1º quartil de outro, definindo o intervalo em termos da razão entre áreas que define a transição de um resultado ao outro. No Teste dos Dois Primeiros Dígitos em Conjunto isso é bastante evidente.

Com exceção feita aos dois primeiros testes, os demais, que são baseados na técnica *bootstrap*, apontam para a ausência de indício de erro ou fraude (resultado = NÃO) quando a razão entre área excedente e área total da *fda* se encontra abaixo de 0,18. Entretanto, esse valor trata-se de evidência empírica que necessita ser confirmada por uma análise extensiva do método cuja intuição apresentamos nesse trabalho.

6. TOMADAS DE CONTAS ESPECIAIS (TCEs)

6.1 Conceito

Tendo analisado a conformidade dos gastos diretos do Governo Federal com a Lei de Newcomb-Benford à procura de indícios de erros ou fraudes, introduzimos o dispositivo legal que rege a apuração de responsabilidade por danos causados ao Erário, qual seja, a Tomada de Contas Especial (TCE). Segundo o Manual de Tomada de Contas Especial, da Controladoria Geral da União (CGU), as definições para a TCE estão contidas no art. 2º da Instrução Normativa TCU nº 71/2012 e no art. 82 da Portaria Interministerial MPOG/MF/CGU nº 507/2011:

“Tomada de Contas Especial é um processo administrativo devidamente formalizado, com rito próprio, para apurar responsabilidade por ocorrência de dano à administração pública federal, com apuração de fatos, quantificação do dano, identificação dos responsáveis e obter o respectivo ressarcimento.” (art. 2º, caput, da IN/TCU n.º 71/2012).

“A Tomada de Contas Especial é um processo devidamente formalizado, dotado de rito próprio, que objetiva apurar os fatos, identificar os responsáveis e quantificar o dano causado ao Erário, visando ao seu imediato ressarcimento.” (art. 82 da Portaria Interministerial MPOG/MF/CGU n.º 507/2011).

Ainda, segundo o Manual, o processo tem por base a conduta do agente público que agiu em descumprimento à lei ou deixou de atender ao interesse público, quando da omissão no dever de prestar contas, da não comprovação da aplicação de recursos, da ocorrência de desfalque, alcance, desvio ou desaparecimento de dinheiros, bens ou valores públicos, ou de prática de ato ilegal, ilegítimo ou antieconômico de que resulte dano à administração pública federal.

Os pressupostos para instauração de TCE estão descritos no art. 5º da IN/TCU nº 71/2012, o qual transcrevemos abaixo:

“Art. 5º É pressuposto para instauração de tomada de contas especial a existência de elementos fáticos e jurídicos suficientes para:

I - comprovação da ocorrência de dano; e

II - identificação das pessoas físicas ou jurídicas que deram causa ou concorreram para a ocorrência de dano.

§ 1º A demonstração de que tratam os incisos I e II deste artigo abrange, obrigatoriamente:

I - descrição detalhada da situação que deu origem ao dano, lastreada em documentos, narrativas e outros elementos probatórios que deem suporte à comprovação de sua ocorrência;

II - exame da suficiência e da adequação das informações, contidas em pareceres de agentes públicos, quanto à identificação e quantificação do dano;

III - evidenciação da relação entre a situação que deu origem ao dano e a conduta ilegal, ilegítima ou antieconômica da pessoa física ou jurídica a quem se imputa a obrigação de ressarcir os cofres públicos, por ter causado ou concorrido para a ocorrência de dano.”

Além disso, de acordo com o art. 3 da IN/TCU nº 56/2007, a tomada de contas especial só deve ser instaurada pela autoridade administrativa federal após esgotadas as providências administrativas internas sem obtenção do ressarcimento pretendido.

O objetivo da Tomada de Contas Especial, ainda segundo o Manual, é apurar responsabilidade por ocorrência de dano à administração pública federal - com

levantamento de fatos, quantificação do dano, identificação dos responsáveis - e obter o respectivo ressarcimento.

Finalmente, dentre os diversos motivos para a instauração de uma Tomada de Contas Especial, nos detemos apenas naquele ao qual se aplica a análise de conformidade com a Lei de Newcomb-Benford, qual seja, a impugnação de despesas que, segundo o Manual, ocorre:

*“...quando são verificadas irregularidades na comprovação da execução de despesas do convênio, tais como **documentos fiscais inidôneos, pagamento irregular de despesas, superfaturamento na contratação de obras e serviços, entre outros.** Nestas situações, o débito original deverá ser quantificado conforme as irregularidades constatadas.” (grifo nosso)*

6.2 Base de Dados

A partir dos dados relativos à instauração de Tomadas de Contas Especiais disponíveis no sítio da CGU (no caminho Página Inicial > Assuntos > Auditoria e Fiscalização > Avaliação da Gestão dos Administradores > Tomada de Contas Especial > Processos Encaminhados ao TCU), elaboramos a tabela abaixo relativa à impugnação de despesas, por órgão, para o período analisado na seção anterior, de 2011 a 2014:

Tabela 6.1: Tomadas de Contas Especiais

Órgão	2011	2012	2013	2014	Total
20000	1	0	0	0	1
20102	0	nd	nd	nd	nd
20113	0	0	0	0	0
22000	5	0	2	0	7
24000	0	1	3	0	4
25000	0	0	0	0	0
26000	43	24	74	0	141
28000	1	2	2	0	5
30000	0	0	0	0	0
32000	1	1	0	0	2
33000	0	0	0	0	0

35000	0	0	0	0	0
36000	83	17	39	0	139
38000	6	4	1	0	11
39000	3	1	1	0	5
41000	14	0	1	0	15
42000	2	6	20	0	28
44000	4	2	3	0	9
49000	5	3	5	0	13
51000	6	5	12	1	24
52000	0	0	0	0	0
53000	49	13	14	1	77
54000	33	10	67	1	111
55000	14	16	33	2	65
56000	11	4	4	0	19
58000	0	1	1	0	2
Total	281	110	282	5	678

Fonte: Elaborado pelo autor.

Dado que a instauração de uma TCE é um processo que requer tempo, é natural que a quantidade total referente a despesas realizadas em 2014 ainda seja muito pequena. Além disso, podemos verificar que os Ministérios da Educação, Saúde e Turismo concentram a maior parte das TCEs instauradas no período.

6.3 Análise de Desempenho das Predições

Então, comparamos os resultados dos testes demonstrados no Quadro 5.1 com os da tabela acima, a fim de determinar o nível de acurácia da não-conformidade com a Lei de Newcomb-Benford em relação à instauração de Tomadas de Contas Especiais. Considerando que a sinalização de não-conformidade pode englobar diversas TCEs, desde que instauradas sobre despesas classificadas sobre o mesmo elemento, dado o órgão e ano, efetuamos a comparação em termos binários, isto é, comparando a conformidade (ou não) com a existência de TCEs instauradas (ou não). Assim, a partir de uma matriz de confusão onde o Falso Positivo (FP) ocorre quando foi sinalizada não-conformidade e não foram instauradas TCEs; o Falso Negativo (FN), quando

existem TCEs instauradas e foi sinalizada conformidade; o Verdadeiro Positivo (VP) nos casos em que foram instauradas TCEs e foi sinalizada não-conformidade; e o Verdadeiro Negativo (VN), no caso restante, elaboramos a tabela abaixo:

Tabela 6.2: Comparação das Predições com as TCEs

Órgão	2011	2012	2013	2014
20000	VP	FP	VN	VN
20102	VN	nd	nd	nd
20113	FP	FP	FP	FP
22000	VP	FP	VP	FP
24000	FP	VP	VP	VN
25000	FP	VN	FP	FP
26000	VP	VP	VP	FP
28000	FN	FN	VP	FP
30000	VN	FP	FP	FP
32000	VP	VP	FP	FP
33000	FP	FP	FP	FP
35000	VN	VN	VN	FP
36000	VP	VP	VP	FP
38000	VP	VP	FN	FP
39000	VP	VP	VP	FP
41000	FN	FP	FN	VN
42000	VP	VP	VP	FP
44000	VP	VP	VP	FP
49000	FN	FN	FN	FP
51000	FN	FN	FN	FN
52000	FP	FP	FP	FP
53000	FN	VP	VP	VP
54000	VP	FN	FN	FN
55000	VP	VP	FN	VP
56000	VP	VP	VP	FP
58000	FP	FN	FN	VN
VP	46%	44%	40%	8%
VN	12%	8%	8%	16%
FP	23%	28%	24%	68%
FN	19%	20%	28%	8%
Acurácia = VP+VN	58%	52%	48%	24%
Sensibilidade = VP/(VP+FN)	71%	69%	59%	50%
Especificidade = VN/(VN+FP)	33%	22%	25%	19%

Fonte: Elaborado pelo autor.

Nas três últimas linhas, apresentamos algumas métricas de desempenho de predição obtidas a partir da matriz de confusão. A acurácia é a proporção de predições

corretas, isto é, o acerto total. A sensibilidade, por sua vez, avalia a capacidade do modelo de classificar corretamente um indivíduo como evento, no caso, de apontar corretamente a existência de erros ou fraudes. Já a especificidade, se refere à classificação correta do indivíduo como não-evento, isto é, negar corretamente a existência de erros ou fraudes.

Podemos observar na tabela acima que o nível de acurácia é tão mais elevado quanto maior o tempo decorrido desde a realização da despesa. Esse resultado é encorajador, no sentido que pode significar a correta sinalização antecipada de indícios de erro ou fraude pela análise de conformidade com a distribuição de Newcomb-Benford e que, ao longo do tempo, é confirmada pela instauração de Tomadas de Contas Especiais.

Além disso, o elevado nível de sensibilidade calculado, cujo comportamento ao longo do tempo é semelhante àquele da acurácia, sugere que a análise de conformidade é um bom preditor da instauração de TCEs. Sob essa ótica, é razoável supor que aqueles valores atualmente classificados como FP, principalmente em 2014, podem ser futuramente reclassificados como VP, quando da instauração das TCEs correspondentes.

Por outro lado, a existência de Falsos Negativos, em vermelho na tabela, concentrados principalmente nos Ministérios do Desenvolvimento Agrário (49000) e do Esporte (51000), e o conseqüente baixo nível de especificidade calculado, sugere a necessidade de aprofundamento da análise apresentada nesse trabalho. Uma hipótese para esse fenômeno é que a baixa granularidade aqui considerada, isto é, o

agrupamento de diversas despesas por órgão e elemento de despesa ao longo de um ano, seja insuficiente para detectar algumas não-conformidades. Portanto, trabalhos futuros poderiam se focar no aumento dessa granularidade: seja temporalmente, pelo agrupamento mensal das despesas; seja hierarquicamente, pelo agrupamento pela unidade gestora da despesa subordinada ao órgão.

7. CONSIDERAÇÕES FINAIS

A implantação da auditoria contínua no âmbito da auditoria externa do setor público é um desafio que precisa ser enfrentado. O modelo contabilométrico baseado na Lei de Newcomb-Benford, nos moldes propostos por Nigrini (1996) e introduzido no Brasil por Santos et al. (2003) sofre de sérias limitações relacionadas ao poder excessivo dos testes empregados, implicando em um grande número de falsos positivos, cujo custo é o de uma investigação desnecessária.

Novos testes, baseados na técnica de *bootstrap* (Suh e Headrick, 2010; Su, Headrick e Minaburo, 2011), permitem contornar essas limitações, sendo mais robustos que os testes tradicionais. Tais testes baseiam-se na estimação, através de simulação computacional, de parâmetros populacionais e coeficientes de regressão associados à verdadeira distribuição de Newcomb-Benford e na sua comparação com os respectivos parâmetros e coeficientes associados à uma dada distribuição empiricamente observada.

Nesse contexto, o presente trabalho teve três objetivos principais: i) aplicar testes de conformidade com a Lei de Newcomb-Benford, tradicionais e novos, sobre os gastos diretos do Governo Federal; ii) realizar uma análise de desempenho das predições obtidas, a partir dos dados de instauração de Tomadas de Contas Especiais; e iii) propor uma abordagem experimental para um teste de conformidade baseado na estimação de intervalos de confiança para a função de distribuição acumulada dos dois primeiros dígitos.

Iniciamos o trabalho apresentando uma breve revisão da literatura, incluindo os conceitos de despesa pública, auditoria e contabilometria, com destaque para a aplicação da Lei de Newcomb-Benford como método contabilométrico de auditoria contábil. Então, introduzimos a distribuição de Newcomb-Benford propriamente dita, descrevendo os tradicionais testes de hipótese à ela associados, assim como as críticas à tais testes e os novos testes que se propõe a superá-las. Passamos, então, à técnica *bootstrap* e à sua aplicação no contexto da Lei de Newcomb-Benford, estimando intervalos de confiança para os parâmetros populacionais da distribuição e para os coeficientes de regressão entre os dígitos. Também elencamos os critérios para análise de conformidade e detalhamos a abordagem experimental que se constitui em um dos objetivos desse trabalho, baseada na aplicação da técnica *bootstrap* para a estimação de intervalos de confiança para a função de distribuição acumulada.

Prosseguimos, então, para a análise propriamente dita dos gastos diretos do Governo Federal referentes à 26 órgãos e 8 elementos de despesa entre 2011 e 2014, constituindo um universo de mais de 17 milhões de despesas. A maior parte da quantidade de despesas analisadas concentrou-se nos Ministérios da Educação, Defesa, Saúde e Fazenda. Em relação aos elementos de despesa, Outros Serviços de Terceiros – PJ e Material de Consumo respondem por mais de 75% da quantidade total.

Em seguida, aplicamos os diversos testes e critérios de conformidade apresentados sobre esses dados e apresentamos os resultados individuais para indícios de erro ou fraude em tabelas anuais. Em seguida, consolidamos esses resultados em um quadro esquemático, a fim de visualizar mais facilmente a distribuição das sinalizações desses

indícios ao longo do tempo, dos órgãos, e dos elementos de despesa. De um total de 808 combinações analisadas, 110 (13,61%) apresentaram indícios de erro ou fraude segundo a metodologia utilizada. Os Ministérios da Agricultura, Pesca e Abastecimento e de Minas e Energia apresentaram, respectivamente, 11 e 10 sinalizações; enquanto os elementos de despesa Locação de Mão de Obra e Outros Serviços de Terceiros – PJ, apresentaram, respectivamente, 46 e 32 sinalizações.

Nosso próximo passo foi comparar o desempenho dos diferentes testes. Para tal, elaboramos uma matriz de confusão comparando os testes dois a dois e, assumindo o resultado de um dos testes como verdade, medimos a acurácia do outro. Além disso, medimos também a proporção de Falsos Positivos e Falsos Negativos, cujos resultados indicam o excesso de poder de um teste em relação ao outro. Encontramos evidências do excesso de poder do χ^2 -Teste e do Z-Teste em relação aos testes baseados na técnica *bootstrap*. Entre esses últimos, o teste dos parâmetros de Benford, considerando os dígitos em separado se mostrou aquele com mais poder, enquanto a regressão *bootstrap*, aquele com menos.

Em relação à nossa abordagem experimental, realizamos um comparativo entre os valores calculados para a métrica proposta, qual seja, a razão entre área excedente e área total da *fda*, e os resultados obtidos a partir dos testes aplicados. Nesse caso, tornou-se evidente que, quanto maior o poder do teste considerado, tanto menor a diferença dada pela métrica entre dois resultados distintos e também menor o valor dela associado ao resultado positivo. O teste dos dois primeiros dígitos em conjunto, cujo resultado pode ser indefinido, foi particularmente ilustrativo, dado que os intervalos da métrica associados aos diferentes resultados parecem claramente distintos.

Empiricamente, os testes baseados na técnica *bootstrap* apontam para ausência de indício de erro ou fraude quando a razão entre área excedente e área total da *fda* se encontra abaixo de 0,18. Esse resultado, contudo, necessita ser confirmado através de um estudo mais aprofundado.

Então, introduzimos o dispositivo legal que rege a apuração de responsabilidade por danos causados ao Erário, qual seja, a Tomada de Contas Especial (TCE), cuja instauração pela autoridade administrativa só deve ser realizada após esgotadas todas as providências administrativas internas sem obtenção do ressarcimento pretendido. Nosso foco, contudo, limitou-se às TCEs cujo motivo de instauração foi a impugnação das despesas decorrida de irregularidades em sua comprovação, tais como documentos fiscais inidôneos, pagamento irregular de despesas e superfaturamento na contratação de obras e serviços; dado que tais irregularidades são passíveis de sinalização pela análise de conformidade com a Lei de Newcomb-Benford. No período entre 2011 e 2014, encontramos 678 TCEs instauradas, sua maioria nos Ministérios da Educação, Saúde e Turismo, e nos anos de 2011 e 2013. No ano de 2014, por outro lado, encontramos apenas 5 TCEs de impugnação de despesas, dado que o processo de instauração requer tempo.

Em seguida, comparamos os resultados dos testes às TCEs instauradas a fim de determinar o nível de acurácia da análise de conformidade com a Lei de Newcomb-Benford. Como a sinalização de não-conformidade pode englobar diversas TCEs, desde que instauradas sobre despesas classificadas sobre o mesmo elemento, dado o órgão e ano, escolhemos realizar uma comparação em termos binários. Assim, para os casos em que (não) existe sinalização e (não) existe TCE instaurada, classificamos

como Verdadeiro Positivo (Negativo). Entretanto, se há sinalização, mas não encontramos uma TCE correspondente, trata-se de um Falso Positivo. Caso contrário, estamos lidando com um Falso Negativo. A partir dessas classificações, apresentamos algumas métricas de desempenho: acurácia, sensibilidade e especificidade. Tanto os níveis de acurácia quanto de sensibilidade são estritamente crescentes em relação ao tempo decorrido desde a realização da despesa. Esse resultado é encorajador, no sentido que a análise de conformidade com a Lei de Newcomb-Benford pode estar sinalizando antecipadamente irregularidades que, ao longo do tempo, serão confirmadas pela instauração de Tomadas de Contas Especiais. Mais ainda, o elevado nível de sensibilidade, isto é, a capacidade de apontar corretamente a existência de erros ou fraudes, sugere que, de fato, a análise de conformidade é um bom preditor da instauração de TCEs. Logo, podemos supor que valores atualmente classificamos como Falso Positivo sejam reclassificados como Verdadeiro Positivo no futuro, de modo que o nível de acurácia deve aumentar ao longo do tempo, o que já observamos na prática.

Finalmente, o baixo nível de especificidade observado, decorrente dos Falsos Negativos que se concentram, principalmente, nos Ministérios do Desenvolvimento Agrário e do Esporte, sugere uma necessidade de aprofundamento da análise aqui apresentada. Uma hipótese para esse fenômeno é a de que o nível de granularidade considerado é insuficiente para a detecção de algumas não-conformidades. Assim, trabalhos futuros poderiam considerar aumentar o nível de granularidade, considerando o agrupamento mensal de despesas ou, ainda, o agrupamento pela unidade gestora da despesa subordinada ao órgão.

Em suma, o presente trabalho buscou apresentar e aplicar novos testes de conformidade com a Lei de Newcomb-Benford, baseados na técnica *bootstrap*, na realidade da auditoria contábil do setor público brasileiro, relacionando os resultados encontrados a partir desses testes com a instauração de Tomadas de Contas Especiais a fim de demonstrar o valor preditivo dos primeiros em relação às últimas.

8. REFERÊNCIAS

BENFORD, F. 1938. The law of anomalous numbers. Proceedings of the American Philosophical Society, Vol. 78 n. 4 p. 551-572, 1938.

CARSLAW, C.A.P.N. Anomalies in income numbers: Evidence of goal oriented behavior. The Accounting Review Vol. LXIII n. 2, p. 321-327, 1988.

CHO, W.K.; GAINES, B.J. Breaking the (Benford) law: Statistical fraud detection in campaign finance. The American Statistician, Vol. 61 n. 3, p. 218-223, 2007.

CUNHA, F.C.R. Aplicações da Lei Newcomb-Benford à Auditoria de Obras Públicas. Dissertação (Mestrado em Regulação e Gestão de Negócios) – Universidade de Brasília, Brasília, 2013. 486p.

DINIZ J. A.; CORRAR L. J.; SLOMSKI V. Análise digital: uma abordagem cognitiva na detecção de não conformidade em prestações de contas municipais. In: 10º CONGRESSO USP DE CONTROLADORIA E CONTABILIDADE, FEA/USP, 2010.

EFRON, B; TIBSHIRANI, R. J. An introduction to the bootstrap. Monographs on Statistics and Applied Probability 57, Chapman and Hall/CRC, Boca Raton, Florida, 1998.

FRANCO, H.; MARRA, E. Auditoria contábil. 2ª ed. São Paulo, Ed. Atlas, 1991.

EFRON, B. Bootstrap methods: Another look at the Jackknife. The Annals of Statistics, Vol. 7 n. 1, p. 1-26, 1979.

GEYER, C. L.; WILLIAMSON, P. P. Detecting fraud in data sets using Benford's Law. Communications in Statistics: Simulation and Computation Vol. 33 n. 1, p. 229-246, 2004.

HILL, T.P: A Statistical Derivation of the Significant-Digit Law. Statistical Science Vol. 4,

p. 354-363, 1996.

HOGG, R. V.; McKEAN, J. W.; CRAIG, A. T. Introduction to Mathematical Statistics, 6th ed., Pearson Prentice Hall, Upper Saddle River, New Jersey, 2005

IUDÍCIBUS, S. Existirá a Contabilometria? Revista Brasileira de Contabilidade, n.41, p. 44-60, 1982.

MARION, J.C.; SILVA, L.B. Contabilometria: Novo Campo de Estudos para a Contabilidade. Revista Brasileira de Contabilidade, ano 16, n. 59, p. 34-41, 1986.

NEWCOMB, S. Note on the Frequency of Use of the Different Digits in Natural Numbers. American Journal of Mathematics, Vol. 4, p. 39-40, 1881.

NIGRINI, M.J. A taxpayer compliance application of Benford's Law. The Journal of the American Taxation Association Vol. 18, p. 72-91, 1996.

NIGRINI, M.J. Digital Analysis Using Benford's Law: Tests & Statistics for Auditors, Global Audit Publications, Vancouver, 2000.

NIGRINI, M.J. Benford's law: applications for forensic accounting, auditing, and fraud detection. Fraudulent Financial Statements, Hoboken, NJ: Wiley, 2012.

NIGRINI, M.J.; MITTERMAIER, L. J. The Use of Benford's Law as an Aid in Analytical Procedures. Auditing: A Journal of Practice & Theory, p. 52-67, 1997.

PEDHAZUR, E. Multiple Regression and Behavioral Research, 3rd Ed., Victoria, Australia: Wadsworth, 1997.

PINKHAM, R.S. On the Distribution of First Significant Digits. Annals of Mathematical Statistics, Vol. 32, p. 1223-1230, 1961.

SANTOS, J.; DINIZ, J.A.; RIBEIRO FILHO, J.F. A Lei de Newcomb-Benford: uma aplicação para determinar o DNA-equivalente das despesas no setor público. In: 3º SEMINÁRIO DE USP DE CONTABILIDADE E CONTROLADORIA, São Paulo, 2003.

SANTOS, J.; DINIZ, J. A.; CORRAR, L.J. O foco é a teoria amostral nos campos de auditoria contábil tradicional e da auditoria digital: testando a lei Newcomb-Benford para o primeiro dígito nas contas públicas. Brazilian Business Review – BBR, n. 1, p. 71-89, 2005.

SANTOS, J.; RIBEIRO FILHO, J.F.; LAGIOIA, U.; ALVES FILHO, B.F.; ARAÚJO, I.J.C. Aplicações da lei de Newcomb-Benford na auditoria tributária do imposto sobre serviços de qualquer natureza (ISS). Revista Contabilidade & Finanças, Vol. 20, Janeiro/Abril, p. 79-94, 2009.

SANTOS, J.; COSTA, J.I.F.; TRAVASSOS, S.K.M. Análise de Conformidade nos Gastos Públicos dos Entes Federativos: Estudo de Caso de uma Aplicação da Lei de Newcomb-Benford para o Primeiro e Segundo Dígito em Dois Estados Brasileiros. Revista Contabilidade & Finanças, Vol.23, n.60, 2012.

SUH, I.; HEADRICK, T.C. A comparative analysis of the bootstrap versus traditional statistical procedures applied to digital analysis based on Benford's Law. Journal of Forensic and Investigative Accounting, Vol. 2, n. 2, p. 144-175, 2010.

SUH, I., HEADRICK, T.C.; MINABURO, S. An Effective and Efficient Analytic Technique: A Bootstrap Regression Procedure and Benford's Law. Journal of Forensic & Investigative Accounting, Vol. 3, n. 3, p.25-44, 2011.

VASARHELYI, M. A.; HALPER, F.B. The continuous audit of online systems. Auditing: A Journal of Practice and Theory, Vol. 10 n. 1, p. 110-125, 1991.