

Érica da Silva Gomes Rabelo

**A Lei de Benford e Fraudes Eleitorais:  
O Caso das Eleições Presidenciais Brasileiras de  
2014**

Brasília

2016



Érica da Silva Gomes Rabelo

**A Lei de Benford e Fraudes Eleitorais:  
O Caso das Eleições Presidenciais Brasileiras de 2014**

Monografia apresentada como requisito para  
obtenção do grau de Bacharel no Curso de  
Ciências Econômicas do Departamento de  
Economia da Universidade de Brasília.

Universidade de Brasília – Unb  
Faculdade de Economia, Administração e Contabilidade  
Departamento de Economia

Orientador: Maurício Soares Bugarin

Brasília

2016

Érica da Silva Gomes Rabelo

# **A Lei de Benford e Fraudes Eleitorais: O Caso das Eleições Presidenciais Brasileiras de 2014**

Monografia apresentada como requisito para  
obtenção do grau de Bacharel no Curso de  
Ciências Econômicas do Departamento de  
Economia da Universidade de Brasília.

Trabalho aprovado. Brasília, 2016:

---

**Maurício Soares Bugarin**  
Orientador

---

**Flávia Ceccato Rodrigues da Cunha**  
Membro da Mesa

Brasília  
2016

*Numbers. They tell the truth?*  
(Joseph T. Wells, 2012)



# Resumo

O presente trabalho tem por objetivo analisar se há evidências de manipulação nos dados das eleições para o cargo de presidente para o Brasil no ano de 2014. Para este fim, foram realizadas avaliações quanto à conformidade com a distribuição Lei de Benford (LB). Na ausência de conformidade com a Lei de Benford, uma distribuição baseada nas eleições presidenciais passadas foi estimada e comparada com a distribuição das eleições de 2014. Para estes testes foram utilizados dados do sistema Tribunal Superior Eleitoral (TSE). Conclui-se que o teste da média dos desvios absolutos para o segundo dígito é o mais indicado para analisar dados eleitorais, o que está em conformidade com a literatura, como em Mebane (2006).

**Palavras-chaves:** Lei de Benford. Eleições Presidenciais. Fraude.

**Classificação JEL:** CO2, C10, C12, C87.





# Abstract

This study aims to analyze if there is evidence of data manipulation in the Brazilian presidential election of 2014. To do so, it was evaluated if the data follow Benford's distribution law. In the absence of agreement to Benford's law, it was estimated an historic Brazilian distribution with data taken from past presidential elections. Then, this distribution was compared to 2014's data. It was used data from Tribunal Superior Eleitoral (TSE). The mean absolute deviation test for the second digit was the most efficient test to analyze electoral data, which is in agreement with the literature, as in Mebane (2006).

**Key-words:** Benford's Law. Presidential Elections. Fraud.



# Lista de ilustrações

Figura 1 – Percentual do desempenho de cada candidato por estado - Elaboração Própria . . . . .	31
Figura 2 – Resultado das Eleições Presidenciais de 2014 - Fonte: Elaboração Própria	51
Figura 3 – Descrição percentual da quantidade de votos para o código votável 13 - Fonte: elaboração própria . . . . .	52
Figura 4 – Descrição percentual da quantidade de votos para o código votável 45 - Fonte: elaboração própria . . . . .	53
Figura 5 – Stata . . . . .	55
Figura 6 – Stata . . . . .	56
Figura 7 – Stata . . . . .	56
Figura 8 – Stata . . . . .	57
Figura 9 – Distribuição dos Primeiros Dígitos- Elaboração Própria . . . . .	64
Figura 10 – Distribuição dos Segundos Dígitos - Elaboração Própria . . . . .	65
Figura 11 – Distribuição dos dois Primeiros Dígitos - Elaboração Própria . . . . .	69
Figura 12 – Teste da Soma para código votável 13 - Elaboração Própria . . . . .	70
Figura 13 – Distribuição dos Primeiros Dígitos - Elaboração Própria . . . . .	72
Figura 14 – Distribuição dos Segundos Dígitos - Elaboração Própria . . . . .	73
Figura 15 – Distribuição dos dois Primeiros Dígitos - Elaboração Própria . . . . .	77
Figura 16 – Teste da Soma para código votável 45 - Elaboração Própria . . . . .	78
Figura 17 – Distribuição dos Primeiros Dígitos - Elaboração Própria . . . . .	82
Figura 18 – Distribuição dos Segundos Dígitos - Elaboração Própria . . . . .	83
Figura 19 – Distribuição dos dois Primeiros Dígitos - Elaboração Própria . . . . .	87
Figura 20 – Distribuição dos Primeiros Dígitos - Elaboração Própria . . . . .	88
Figura 21 – Distribuição dos Segundos Dígitos - Elaboração Própria . . . . .	89
Figura 22 – Distribuição dos dois Primeiros Dígitos - Elaboração Própria . . . . .	93



# Lista de tabelas

Tabela 1 – Frequência dos primeiros dígitos - Fonte: Adaptado Newcomb (1881) . . . . .	18
Tabela 2 – Experimentos de Benford - Fonte: Adaptado de Benford (1938) . . . . .	19
Tabela 3 – Perfil dos dados para a base de quantidade de votos nas eleições de 2014 segundo turno no Brasil - Fonte: elaboração própria . . . . .	23
Tabela 4 – Valores críticos e regiões de conformidade - Fonte: Adaptado de Drake e Nigrini (2000) . . . . .	28
Tabela 5 – Descrição percentual do desempenho de cada candidato por estado - Fonte: elaboração própria . . . . .	29
Tabela 6 – Teste Qui- Quadrado (CS) para o Segundo Dígito aplicado aos votos de cada candidato sob a distribuição de Benford - Fonte: elaboração própria . . . . .	35
Tabela 7 – Teste Qui- Quadrado (CS) para o Segundo Dígito aplicado aos votos de cada candidato sob a distribuição do Brasil - Fonte: elaboração própria . . . . .	36
Tabela 8 – Teste da Média dos Desvios Absolutos (MAD) para o Segundo Dígito aplicado aos votos de cada candidato sob a distribuição de Benford - Fonte: elaboração própria . . . . .	37
Tabela 9 – Teste da Média dos Desvios Absolutos (MAD) para o Segundo Dígito aplicado aos votos de cada candidato sob a distribuição do Brasil - Fonte: elaboração própria . . . . .	39
Tabela 10 – Descrição percentual da quantidade de votos brancos e nulos por estado - Fonte: elaboração própria . . . . .	54
Tabela 11 – Frequências dos Primeiros Dígitos da distribuição de Benford e da distribuição Histórica Brasileira - Fonte: elaboração própria . . . . .	59
Tabela 12 – Frequências dos Segundos Dígitos da distribuição de Benford e da distribuição Histórica Brasileira - Fonte: elaboração própria . . . . .	59
Tabela 13 – Frequências dos Dois Primeiros Dígitos da distribuição de Benford e da distribuição Histórica Brasileira - Fonte: elaboração própria . . . . .	61
Tabela 14 – Teste do Primeiro Dígito conforme Benford para o estado do Acre com código votável 13 - Fonte: elaboração própria . . . . .	63
Tabela 15 – Teste do Segundo Dígito conforme Benford para o estado do Acre com código votável 13 - Fonte: elaboração própria . . . . .	64
Tabela 16 – – Teste dos Dois Primeiros Dígitos conforme Benford para o estado do Acre com código votável 13 - Fonte: elaboração própria . . . . .	67
Tabela 17 – Continuação- Teste dos Dois Primeiros Dígitos conforme Benford para o estado do Acre com código votável 13 - Fonte: elaboração própria . . . . .	68
Tabela 18 – Teste da Soma para código votável 13 - Elaboração Própria . . . . .	71

Tabela 19 – Teste do Primeiro Dígito conforme Benford para o estado do Acre com código votável 45 - Fonte: elaboração própria . . . . .	72
Tabela 20 – Teste do Segundo Dígito conforme Benford para o estado do Acre com código votável 45 - Fonte: elaboração própria . . . . .	73
Tabela 21 – Teste dos Dois Primeiros Dígitos conforme Benford para o estado do Acre com código votável 45 - Fonte: elaboração própria . . . . .	75
Tabela 22 – Continuação- Teste dos Dois Primeiros Dígitos conforme Benford para o estado do Acre com código votável 45 - Fonte: elaboração própria . . . . .	76
Tabela 23 – Teste da Soma para código votável 45 . . . . .	79
Tabela 24 – Teste do Primeiro Dígito conforme a distribuição do Brasil, estado do Acre e código votável 13 - Fonte: elaboração própria . . . . .	81
Tabela 25 – Teste do Segundo Dígito conforme a distribuição do Brasil, estado do Acre e código votável 13 - Fonte: elaboração própria . . . . .	82
Tabela 26 – Teste dos Dois Primeiros Dígitos conforme a distribuição do Brasil, estado do Acre e código votável 13 - Fonte: elaboração própria . . . . .	85
Tabela 27 – Continuação- Teste dos Dois Primeiros Dígitos conforme a distribuição do Brasil, estado do Acre e código votável 13 - Fonte: elaboração própria . . . . .	86
Tabela 28 – Teste do Primeiro Dígito conforme a distribuição do Brasil, estado do Acre e código votável 45 - Fonte: elaboração própria . . . . .	87
Tabela 29 – Teste do Segundo Dígito conforme a distribuição do Brasil, estado do Acre e código votável 45 - Fonte: elaboração própria . . . . .	88
Tabela 30 – Teste dos Dois Primeiros Dígitos conforme a distribuição do Brasil, estado do Acre e código votável 45 - Fonte: elaboração própria . . . . .	91
Tabela 31 – Continuação- Teste dos Dois Primeiros Dígitos conforme a distribuição do Brasil, estado do Acre e código votável 45 - Fonte: elaboração própria . . . . .	92

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>15</b>
	Introdução	15
<b>2</b>	<b>A LEI DE BENFORD</b>	<b>17</b>
2.1	Histórico e Conceitos Básicos	17
<b>3</b>	<b>PROCEDIMENTO METODOLÓGICO</b>	<b>23</b>
3.1	Perfil dos Dados	23
3.2	Testes da Lei de Benford, Baseado nas Probabilidades dos Dígitos	24
3.2.1	Teste do Primeiro Dígito	24
3.2.2	Testes do Segundo Dígito	25
3.2.3	Teste dos Dois Primeiros Dígitos	25
3.2.4	Teste da Soma	25
3.2.5	Confronto entre o Teste da Soma e o Teste dos dois Primeiros Dígitos	26
3.3	Teste da Lei de Benford, Baseados nas Distribuições das Frequências dos Dígitos	26
3.3.1	Teste Z	27
3.3.2	Teste Qui-Quadrado	27
3.3.3	Teste da Média dos Desvios Absolutos (MAD)	28
3.4	Distribuição dos Votos na Eleição Presidencial de 2014	28
<b>4</b>	<b>ANÁLISE</b>	<b>33</b>
4.1	Distribuição Histórica Brasileira	33
4.2	Teste Qui-Quadrado (CS)	34
4.3	Teste da Média dos Desvios Absolutos	37
<b>5</b>	<b>CONCLUSÃO</b>	<b>41</b>
	Conclusão	41
	<b>REFERÊNCIAS</b>	<b>43</b>
	<b>APÊNDICES</b>	<b>49</b>
	<b>APÊNDICE A – RESULTADO DA APURAÇÃO - ELEIÇÕES 2014</b>	<b>51</b>

	<b>APÊNDICE B – CÓDIGO STATA PARA LEITURA E PROCES- SAMENTO DE DADOS . . . . .</b>	<b>55</b>
	<b>APÊNDICE C – COMPARAÇÃO ENTRE AS DISTRIBUIÇÕES . .</b>	<b>59</b>
<b>C.1</b>	<b>Primeiro Dígito . . . . .</b>	<b>59</b>
<b>C.2</b>	<b>Segundo Dígito . . . . .</b>	<b>59</b>
<b>C.3</b>	<b>Dois Primeiros Dígitos . . . . .</b>	<b>61</b>
	<b>APÊNDICE D – TABELAS DEMONSTRATIVAS DOS RESULTA- DOS DOS TESTES PARA DISTRIBUIÇÃO DE BENFORD . . . . .</b>	<b>63</b>
<b>D.1</b>	<b>Tabelas Acre para o código votável 13 . . . . .</b>	<b>63</b>
D.1.1	Primeiro Dígito . . . . .	63
D.1.2	Segundo Dígito . . . . .	64
D.1.3	Teste dos dois Primeiros Dígitos . . . . .	67
<b>D.2</b>	<b>Teste da Soma . . . . .</b>	<b>69</b>
D.2.1	Código Votável 13 . . . . .	70
<b>D.3</b>	<b>Tabelas Acre para o código votável 45 . . . . .</b>	<b>72</b>
D.3.1	Primeiro Dígito . . . . .	72
D.3.2	Segundo Dígito . . . . .	73
D.3.3	Dois Primeiros Dígitos . . . . .	75
<b>D.4</b>	<b>Teste da Soma . . . . .</b>	<b>77</b>
D.4.1	Código Votável 45 . . . . .	78
	<b>APÊNDICE E – TABELAS DEMONSTRATIVAS DOS RESULTA- DOS DOS TESTES PARA DISTRIBUIÇÃO HIS- TÓRICA BRASILEIRA . . . . .</b>	<b>81</b>
<b>E.1</b>	<b>Tabelas Acre para o código votável 13 . . . . .</b>	<b>81</b>
E.1.1	Primeiro Dígito . . . . .	81
E.1.2	Segundo Dígito . . . . .	82
E.1.3	Dois Primeiros Dígitos . . . . .	85
<b>E.2</b>	<b>Tabelas Acre para o código votável 45 . . . . .</b>	<b>87</b>
E.2.1	Primeiro Dígito . . . . .	87
E.2.2	Segundo Dígito . . . . .	88
E.2.3	Dois Primeiros Dígitos . . . . .	91



# 1 Introdução

Em 2014 foi realizada a sétima eleição presidencial direta do Brasil desde a redemocratização em 1989. O período de transição do regime militar para a democracia durou cerca de 11 anos e mais cinco anos para que houvesse a primeira eleição para presidente (Kinzo, 2001). Portanto, a conquista do direito ao voto direto para eleger seus governantes foi um processo gradual que demandou um grande esforço da sociedade civil brasileira.

No entanto, apesar das conquistas obtidas desde a redemocratização, o escrutínio eleitoral continua suscitando dúvidas. Em 2016, o Tribunal Superior Eleitoral (TSE) pediu que fossem investigados possíveis falhas em 77 mil votos referentes às eleições de 2014 (Aguiar, 2016). E apesar de o Brasil ser pioneiro no uso das urnas eletrônicas, as urnas utilizadas ainda são as mesmas desde seu primeiro uso em 1996 (Pitta, 2015).

As urnas utilizadas no Brasil pertencem à primeira geração das mesmas. Esse tipo de urna não permite que o voto seja impresso e desde que a Índia desistiu de utilizá-las, o Brasil passou a ser o único país do mundo que permanece utilizando essa tecnologia (Pitta, 2015). Outros países utilizam urnas eletrônicas, como a Bélgica e Israel, mas a geração utilizada por esses países permite que o voto seja impresso, o que assegura a possibilidade de recontagem dos votos caso necessário.

O uso das urnas eletrônicas com tecnologia antiga pode apresentar riscos à segurança do processo eleitoral, mas mesmo os países que utilizam outros meio para votar, também estão suscetíveis a manipulações de votos. Existe um método que tem sido utilizado para analisar manipulação de dados que consiste em verificar se os dados seguem a lei de Benford (Silva, 2015).

A lei de Benford é uma lei matemática que demonstra que a frequência dos dígitos de uma dada amostra de dados tem uma distribuição decrescente do dígito 1 ao dígito 9 <sup>1</sup> Mebane (2006) foi um dos primeiros a propor o uso da lei de Benford para avaliar indícios de fraude eleitoral. Segundo Pericchi e Torres (2004) e Mebane (2006), o uso da Lei de Benford para analisar segundo dígito dos dados é mais indicado para analisar dados eleitorais.

O trabalho seguirá a indicação de Mebane e utilizará a distribuição do segundo dígito para analisar os dados da eleição presidencial de 2014 <sup>2</sup>. Os dados analisados provêm do Tribunal Superior Eleitoral (TSE) <sup>3</sup>. Para a realização dos testes, foi utilizado o software estatístico Stata e o modelo Law/NigriniCycle.xlsx para produzir os resultados

---

<sup>1</sup> A lei de Benford será analisada com maiores detalhes na primeira seção desse trabalho.

<sup>2</sup> Os testes também foram feitos para o primeiro dígito e para os dois primeiros dígitos e estão contidos no apêndice do trabalho.

<sup>3</sup> <http://www.tse.jus.br/eleicoes/estatisticas/repositorio-de-dados-eleitorais>

dos gráficos.

Além da presente introdução, este trabalho está organizado da seguinte forma: o capítulo 2 apresenta o referencial teórico, isto é, a lei de Benford e algumas de suas propriedades; o capítulo 3 discute a metodologia utilizada na pesquisa e algumas técnicas estatísticas frequentemente utilizadas em análises de frequência de dígitos a partir da LB. No capítulo 4 são apresentados os resultados dos testes que utilizam a lei de Benford e a distribuição brasileira; a última seção apresenta as considerações finais.

## 2 A Lei de Benford

### 2.1 Histórico e Conceitos Básicos

A Lei de Newcomb-Benford estabelece que a frequência esperada dos dígitos "menores", como, por exemplo, os dígitos 1 e 2, são maiores do que a frequência esperada dos dígitos "maiores", como, por exemplo, os dígitos 8 e 9. A probabilidade de ocorrência dos dígitos decai continuamente do dígito 1 ao dígito 9. Sendo o maior valor esperado, o valor associado ao dígito 1, e o menor valor esperado, o valor associado ao dígito 9. A lei de Benford é verificada em um amplo conjunto de dados numéricos, desde números obtidos em páginas de jornais (Bugarin e Cunha, 2016) a números referentes a pesos atômicos (Nigrini, 2011).

Benford (1938) notou que as primeiras páginas das tábuas logarítmicas apresentavam mais manchas que as últimas páginas (Bugarin e Cunha, 2016). Esse fato indicava que as páginas que iniciavam com os primeiros dígitos eram mais acessadas do que as páginas que iniciavam com os últimos dígitos. E essa relação era decrescente do dígito 1 ao dígito 9, onde o dígito 1 era o mais acessado e o dígito 9 era o menos acessado. Tal fato contraria a intuição de que os dígitos devam apresentar uma distribuição uniforme. Mas foi Benford (1938) que derivou a frequência esperada dos dígitos (Nigrini, 2011). A distribuição dos dígitos observada por Benford é assimétrica à esquerda mas tende a ter uma distribuição uniforme quando mais dígitos são analisados

Newcomb (1881) descobriu uma propriedade muito importante, que é uma das bases para a Lei de Benford. Com isso, A probabilidade de ocorrência dos dígitos seria tal que as mantissas de seus logaritmos seriam igualmente prováveis. Com isso, o autor derivou a tabela que indica a frequência dos primeiros dígitos percebendo outra importante propriedade: a distribuição dos dígitos é assimétrica à esquerda, mas se torna mais uniforme quando analisamos dígitos posteriores. A distribuição seria muito assimétrica para o primeiro dígito e se tornaria mais uniforme para o segundo, ainda mais uniforme para o terceiro e, a partir do quarto, as diferenças de frequência seriam tão pequenas, que poderíamos considerá-las iguais. (Silva, 2015)

A tabela abaixo apresenta a frequência do primeiro e do segundo dígito. Note que a probabilidade encontrada quando o primeiro dígito corresponde ao valor "1" é cerca de 0,301, ao invés de  $\frac{1}{9}$  como se esperaria caso todos os dígitos tivessem a mesma probabilidade de ocorrência.

Tabela 1 – Frequência dos primeiros dígitos - Fonte: Adaptado Newcomb (1881)

Dígito	Primeiro Dígito	Segundo Dígito
0	-	0,119679269
1	0,301029996	0,113890103
2	0,176091259	0,108821499
3	0,124938737	0,10432956
4	0,096910013	0,100308202
5	0,079181246	0,096677236
6	0,06694679	0,093374736
7	0,057991947	0,090351989
8	0,051152522	0,087570054
9	0,045757491	0,084997352

O trabalho de Newcomb não teve impacto na época de sua publicação. Foi apenas após Frank Beford publicar o artigo intitulado "The Law of Anomalous Numbers" que a lei de Benford passou a ser utilizada para fins práticos (Silva, p. 8). Benford fez vários experimentos. Ele coletou dados sobre a área de rios, constantes físicas, tabelas com pesos atômicos, endereços, taxa de mortalidade, dentre outros. A tabela abaixo mostra que para alguns desses experimentos, os dados se ajustaram muito bem à distribuição de Benford.

Tabela 2 – Experimentos de Benford - Fonte: Adaptado de Benford (1938)

Group	Title	First Digit									Count
		1	2	3	4	5	6	7	8	9	
A	Rivers, area	31.0	16.4	10.7	11.3	7.2	8.6	5.5	4.2	5.1	335
B	Population	33.9	20.4	14.2	8.1	7.2	6.2	4.1	3.7	2.2	3259
C	Constants	41.3	14.4	4.8	8.6	10.6	5.8	1.0	2.9	10.6	104
D	Newspapers	30.0	18.0	12.0	10.0	8.0	6.0	6.0	5.0	5.0	100
E	Spec. Heat	24.0	18.4	16.2	14.6	10.6	4.1	3.2	4.8	4.1	1389
F	Pressure	29.6	18.3	12.8	9.8	8.3	6.4	5.7	4.4	4.7	703
G	H.P. Lost	30.0	18.4	11.9	10.8	8.1	7.0	5.1	5.1	3.6	690
H	Mol. Wgt.	26.7	25.2	15.4	10.8	6.7	5.1	4.1	2.8	3.2	1800
I	Drainage	27.1	23.9	13.8	12.6	8.2	5.0	5.0	2.5	1.9	159
J	Atomic Wgt.	47.2	18.7	5.5	4.4	6.6	4.4	3.3	4.4	5.5	91
K	$\frac{1}{n}, \sqrt{n} \dots$	25.7	20.3	9.7	6.8	6.6	6.8	7.2	8.0	8.9	5000
L	Design	26.8	14.8	14.3	7.5	8.3	8.4	7.0	7.3	5.6	560
M	Digest	33.4	18.5	12.4	7.5	7.1	6.5	5.5	4.9	4.2	308
N	Cost Data	32.4	18.8	10.1	10.1	9.8	5.5	4.7	5.5	3.1	741
O	X-Ray Volts	27.9	17.5	14.4	9.0	8.1	7.4	5.1	5.8	4.8	707
P	Am. League	32.7	17.6	12.6	9.8	7.4	6.4	4.9	5.6	3.0	1458
Q	Black Body	31.0	17.3	14.1	8.7	6.6	7.0	5.2	4.7	5.4	1165
R	Addresses	28.9	19.2	12.6	8.8	8.5	6.4	5.6	5.0	5.0	312
S	$n_1, n_2, \dots, n!$	25.3	16.0	12.0	10.0	8.5	8.8	6.8	7.1	5.5	900
T	Death Rate	27.0	18.6	15.7	9.4	6.7	6.5	7.2	4.8	4.1	418
Average		30.6	18.5	12.4	9.4	8.0	6.4	5.1	4.9	4.7	1011
Probable Error		$\pm 0.8$	$\pm 0.4$	$\pm 0.4$	$\pm 0.3$	$\pm 0.2$	$\pm 0.2$	$\pm 0.2$	$\pm 0.2$	$\pm 0.3$	-

Em seu trabalho, Nigrini (2011, p. 89) faz uma breve revisão sobre o desenvolvimento da literatura da Lei de Benford (LB). Outra revisão da literatura também pode ser encontrada em Café (2015). Os primeiros artigos sobre a lei de Benford, após a publicação de Benford, foram de Goudsmit e Furry (1944) e Furry e Hurwitz (1945), em que os autores discutem algumas propriedades matemáticas da LB. Hsu (1948) fez um experimento com seus estudantes, pedindo para que os mesmos escrevessem um número com 4 dígitos que não fosse relacionado a nenhuma data, evento, etc. Nesse experimento, Hsu não achou relação entre os dígitos escolhidos por seus alunos e a Lei de Benford.

Um dos resultados mais importantes sobre a LB se deve ao trabalho de Pinkham (1961). Pinkham mostrou que se existe uma lei que governa as distribuições numéricas, então essa distribuição deve ser invariante em relação à escala (Nigrini, 2011, p. 90). Esse resultado implica que se um conjunto de números seguir a Lei de Benford e este conjunto for multiplicado por uma constante não nula, então o conjunto continuará seguindo a LB.

Duas demonstrações da Lei de Benford foram desenvolvidas em 1966. Uma demonstração foi feita por Feller (1966) e outra demonstração foi desenvolvida de forma independente por Flehinger (1966). Varian (1972) tabulou a frequência de alguns conjun-

tos de dados demográficos e concluiu que os dados seguiam a Lei de Benford. Tsao (1974) descobriu que os erros de arredondamento de um computador também seguem a Lei de Benford.

A primeira aplicação da Lei de Benford a auditoria foi feita por Carslaw (1988). Carslaw argumentou que gerentes de empresas têm incentivos a arredondar os números dos resultados das empresas para cima, de forma que a frequência dos dígitos maiores, como o dígito 9, seja inferior aos dígitos 1 e 2, o que condiz com a Lei de Benford. Hill (1988) mostrou que quando indivíduos criam números, esses números não seguem a Lei de Benford. Mosimann, Wiseman, and Edelman (1995) mostraram que a maioria das pessoas não conseguem gerar números aleatórios, mesmo que de forma consciente.

A primeira análise forense que utilizou a lei de Benford foi feita por Nigrini (1994). Ele deduziu que as pessoas tendem a repetir suas ações, e, portanto, caso os dados sejam fraudados, eles não seguiram a Lei de Benford.

Hill (1995) fez a maior contribuição matemática desde o trabalho de Pinkham (1961), segundo Nigrini (2011, p. 95). Hill mostrou que se distribuições são selecionadas aleatoriamente e amostras aleatórias são retiradas de cada uma dessas distribuições, então os dígitos da coleção convergirão para a distribuição logarítmica (Benford). A importância desse artigo é corroborada pela seguinte citação:

A prova mais robusta e aceita atualmente é a apresentada por Hill (1995), que deriva a Lei de Benford a partir da hipótese de invariância de base. A prova apresentada por Hill baseia-se no fato de que os números cujos primeiros dígitos seguem a distribuição de Benford são gerados a partir da combinação de outras distribuições, e ainda que uma distribuição isolada não siga perfeitamente a distribuição de Benford, a combinação das distribuições o fará. (Café, 2015, p 9.)

Os casos em que a Lei de Benford pode ser aplicada estão listadas nos trabalhos de Nigrini (2011, p.98) e em Durtschi, Hillison e Pacini (2004). A Lei de Benford é válida nos seguintes casos: I) quando os dados são aleatórios; II) quando os dados advêm de resultados micros como transações de uma empresa; III) provenham de diferentes distribuições; IV) sejam pertencentes a séries que possuem a média maior que a variância, ou seja, assimétricas; V) se o número de dados é grande, portanto, o tamanho da amostra é fundamental. E não válida: I) caso possuam máximos ou mínimos obrigatórios; II) quando os números apresentam intervenção humana, como os encontrados em endereços postais, contas bancárias, documentos, preços de produtos, números telefônicos.

Nigrini (2012) apresenta uma série de testes que possibilitam o uso da Lei de Benford para fins de auditoria. Os testes apresentados por Nigrini são <sup>1</sup>: I) primeiros dígitos, II) segundos dígitos, III) dois primeiros dígitos, IV) teste da soma, V) Teste Z, VI) Teste Qui-Quadrado (CS), V) Teste da Média dos Desvios Absolutos (MAD).

<sup>1</sup> Esses testes serão apresentados no capítulo seguinte e serão utilizados ao longo do trabalho.

No Brasil, pode-se destacar os trabalhos de Santos, Diniz e Corrar (2005), Ribeiro, Montenegro, Santos e Galvão (2005) que aplicaram a lei de Benford para auditoria contábil <sup>2</sup> Lagioia, Araújo, Alves Filho, Barros e Nascimento (2011) aplicaram a Lei de Benford à auditoria pública do ISS. E Café (2015) explica alguns trabalhos realizados recentemente:

Cunha e Bugarin (2015) aplicam os testes propostos por Nigrini (2012) à auditoria de uma obra pública brasileira, o estádio de futebol Maracanã e encontram resultados admiráveis: os dígitos selecionados pelos testes correspondem a mais de 70 volume de fraudes encontrado pelo Tribunal de Contas da União (TCU). Bugarin e Cunha (2015) aplicam os testes a mais uma obra pública, o estádio Arena Amazônia e também obtém bons resultados. Ainda, Bugarin e Cunha (2015) propõem um algoritmo para a seleção dos dígitos a serem auditados.

---

<sup>2</sup> Para maiores detalhes sobre a literatura da lei de Benford e auditorias contábeis confira Cunha (2013).





## 3 Procedimento Metodológico

Essa seção apresenta os testes da Lei de Benford propostos por Nigrini (2012) e são discutidas as condições para que um conjunto de dados seja passível de análise segundo os métodos apresentados nesse trabalho. Para que a base de dados seja utilizada é preciso que antes seja analisado o perfil dos dados. O perfil de dados é feito na seção 2.1. Em seguida são apresentados os testes sugeridos por Nigrini e discutidos por Bugarin e Cunha (2015), na seção 2.2. Na seção 2.3 são apresentados alguns métodos estatísticos mais avançados. A seção 2.4 apresenta a configuração das eleições de 2014 de acordo com o percentual de votos de cada candidato por unidade federativa. A base de dados é disponibilizada pelo Tribunal Superior Eleitoral (TSE) <sup>1</sup>, e é composta pelos votos da eleição presidencial de 2014 com base na apuração do segundo turno.

### 3.1 Perfil dos Dados

Antes de executar os testes, um exame do perfil dos dados foi realizado. O exame está detalhado na tabela a seguir:

Tabela 3 – Perfil dos dados para a base de quantidade de votos nas eleições de 2014 segundo turno no Brasil - Fonte: elaboração própria

Perfil de Dados		
Detalhamento	Contagem	% do Total
10 ou acima	110.070.804	97,68
De 1 a 9	2.613.075	2,32
Igual à zero	0	0,00
Menor que zero	0	0,00
Soma =	112.683.879	100,00
Valores baixos		
1 a 50	10.028.941	8,90
Valores muito altos		
10.000 ou acima	0	0,00

Nigrini (2011) indica que o perfil de dados é geralmente o primeiro teste a ser executado porque auxilia a entender os dados que estão sendo analisados. Além de ajudar com a compreensão, o perfil de dados pode fornecer alguns resultados iniciais interessantes, auxiliando na detecção de problemas que demonstram que não é viável prosseguir com a análise, como encontrar valores negativos em conjuntos de dados onde não deveriam conter dados deste tipo.

<sup>1</sup> <http://www.tse.jus.br/eleicoes/estatisticas/repositorio-de-dados-eleitorais>

Dependendo dos dados sob investigação, existe a possibilidade de adicionar ou subtrair categorias, ou seja, moldando o perfil de acordo com o tipo de base (NIGRINI, 2012).

Como o conjunto de dados apresentado é constituído por um número relativamente grande de dados é apropriado subdividir a base em vários estratos. São eles: a) igual ou acima de 10; b) intervalo de 1 a 9 c) os zeros; d) números negativos; e) abaixo de 50; f) acima de 100.000.

Por meio da avaliação do perfil de dados é possível concluir que os dados são passíveis de análise pela Lei de Benford, visto que apenas 2,32% das observações encontram-se abaixo do valor 10, o que não gera viés na realização do teste dos Dois Primeiros Dígitos. Não há valores muito altos, negativos ou iguais à zero, que no caso desta planilha indicariam erro. Além disso, os valores baixos encontrados representam apenas 8,9% do total, o que também não promove viés significativo para os testes.

## 3.2 Testes da Lei de Benford, Baseado nas Probabilidades dos Dígitos

Os testes apresentados nessa seção são: o Teste do Primeiro Dígito, Teste do Segundo Dígito, Teste dos dois Primeiros Dígitos, Teste z, Teste Qui Quadrado e Teste da Soma. Todos esses teste são apresentados por Nigrini (2012), e são discutidos por Bugarin e Cunha (2015).

### 3.2.1 Teste do Primeiro Dígito

Esse é o primeiro teste proposto por Nigrini (2012). Seu uso é esclarecido por dois especialistas em Lei de Benford:

Este teste verifica as frequências com que os números de 1 a 9 se repetem nos primeiros dígitos dos valores de um banco de dados. Portanto, ele divide a amostra original em apenas 9 grupos, sendo cada um dos grupos potencialmente muito grande, a depender do volume de dados, o que o caracteriza como um teste de visão “macro”. Dessa forma, esse teste pode ser útil em bancos de dados com poucos itens, talvez 300, segundo Nigrini (2012).(Bugarin e Cunha, 2015)

A frequência esperada de um dado dígito é dada pela seguinte expressão:

$$Prob(D_1 = d_1) = \log \left( 1 + \frac{1}{d_1} \right) \quad (3.1)$$

Onde,  $D_1$  =primeiro dígito e  $d_1 \in \{1, 2, \dots, 9\}$ .

P indica a probabilidade de observar o evento em parênteses, e log refere-se ao log da base 10. Por exemplo, a probabilidade esperada do primeiro dígito ser o número 2 é  $\log(1 + 1/2)$ , que é igual a 0,1761. Como pode-se observar na tabela 1.

### 3.2.2 Testes do Segundo Dígito

O teste do segundo dígito também é um dos testes propostos por Nigrini (2012). Ele é explicado pela Doutora Flávia Ceccato Rodrigues da Cunha (2013):

O Teste do Segundo Dígito testa as frequências com que os números de 0 a 9 se repetem nos segundos dígitos dos itens de um banco de dados. Ele também é um teste de visão macro, mas pode ser útil para detectar vieses nos dados. Em pagamentos, por exemplo, e outros dados em que existam preços envolvidos, o teste geralmente revela um excesso de 0's e 5's nos segundos dígitos, que são números redondos.

A frequência esperada dos segundos dígitos pela LB segue a expressão matemática abaixo.

$$Prob(D_2 = d_2) = \sum_{d_1=1}^9 \log\left(1 + \frac{1}{d_1 d_2}\right); \quad (3.2)$$

Onde  $D_2$  = segundo dígito e  $d_2 \in \{0, 1, \dots, 9\}$ .

### 3.2.3 Teste dos Dois Primeiros Dígitos

O teste dos Dois Primeiros Dígitos divide os dados em 90 grupos (10, 11 até 99) sendo portanto mais preciso que os anteriores, já que o do Primeiro Dígito divide os dados em nove grupos e o do Segundo Dígito os divide em dez grupos (Bugarin e Cunha, 2015). A frequência esperada dos dois primeiros dígitos é dada pela seguinte expressão:

$$Prob(D_1 D_2 = d_1 d_2) = \log\left(1 + \frac{1}{d_1 d_2}\right); \quad (3.3)$$

Onde  $D_1 D_2$  = Dois primeiros dígitos e  $d_1 d_2 \in \{10, 11, \dots, 99\}$

Exemplo:

$$Prob(D_1 D_2 = 11) = \log\left(1 + \frac{1}{11}\right) = \log\left(\frac{12}{11}\right) = 0.03779 \quad (3.4)$$

### 3.2.4 Teste da Soma

Nigrini (2012) desenvolveu um teste avançado após notar que ao simular uma distribuição de Benford descobriu que as somas dos números pertencentes a cada grupo de primeiros dígitos 10, 11, 12, ..., 99 resultavam em valores aproximadamente iguais, isto é, se aproximavam de  $\frac{1}{90}$  (Cunha, 2013).

A seguinte ressalva deve ser feita:

Todavia, o autor constatou que dados reais raramente se conformam a tal padrão, uma vez que eles possuem alguns valores muito altos ou muitas repetições de números de médio valor. A utilidade desse teste é exatamente alertar sobre essas situações. (Bugarin e Cunha, 2015)

Assim, o teste da soma seria como um guia para os investigadores, indicando os dados que contém números excessivamente grandes ou muitas repetições. Quando esta situação ocorre, os dados devem ser avaliados com maior cautela pelo auditor (Café, 2015).

### 3.2.5 Confronto entre o Teste da Soma e o Teste dos dois Primeiros Dígitos

Segundo Bugarin e Cunha (2015), o confronto entre o Teste dos Dois Primeiros Dígitos com o Teste da Soma equivale a confrontar o volume de dados contidos em cada grupo (frequência) e a soma de seus valores. Esta metodologia avalia a materialidade e relevância de cada grupo, portanto auxilia no processo de detecção das categorias de dígitos relevantes para auditorias.

Qualquer categoria de dois dígitos que cair na zona de desconformidade para o teste Z ou para o teste da soma é um candidato a ser analisado com maior detalhe (Bugarin e Cunha, 2015). Os autores, Bugarin e Cunha, propõem comparar a frequência de todas as categorias que foram selecionadas em ao menos um dos testes. Na palavra dos autores:

Therefore, we propose to compare the frequencies of all categories that have been selected in at least one of the two tests. If one of them shows very little frequency according to both criteria, i.e., there are few observations in that category and the value of the summation of the category's items is low, then that category should be excluded from further scrutiny. We call this comparison the "confrontation" of the two tests. Our main point in doing the confrontation is that, in the case of public works' budget, the relevance of each group should be taken into account for selecting of the digits that need further auditing. <sup>2</sup> (Bugarin e Cunha, 2015)

## 3.3 Teste da Lei de Benford, Baseados nas Distribuições das Frequências dos Dígitos

Pode-se aplicar diferentes métodos estatísticos para analisar a conformidade de uma distribuição a LB, nesta seção serão apresentados alguns testes considerando suas vantagens e desvantagens segundo Bugarin e Cunha (2015).

---

<sup>2</sup> Tradução livre: Portanto, propomos comparar as frequências de todas as categorias que foram selecionadas em ao menos um dos dois testes. Se um dos testes mostrar uma frequência pequena de acordo com os dois critérios, isto é, se existir poucas observações naquela categoria e o valor da soma dos itens da categoria for baixo, então aquela categoria deve ser excluída para uma análise mais detalhada. Nós chamamos essa comparação de "confronto" dos dois testes. Nosso ponto principal ao fazer o confronto é que, no caso de orçamentos de obras públicas, a relevância de cada grupo precisa ser levada em consideração para selecionar os dígitos que precisam de mais auditoria.

### 3.3.1 Teste Z

Este teste é utilizado para testar se a proporção real de um dígito difere significativamente ou não da proporção esperada deste dígito pela LB. Sendo definida como:

$$Z = \frac{|PR - PE| - \left(\frac{1}{2n}\right)}{\sqrt{\frac{PE(1-PE)}{n}}} \quad (3.5)$$

Onde (PR) é a proporção real, (PE) a proporção esperada, a expressão  $\frac{1}{2n}$  representa um fator de correção, utilizado apenas quando menor que o primeiro termo do numerador, ou seja que o desvio absoluto entre as proporções e  $(n)$  se refere ao número de dados. Então, a estatística Z depende do desvio absoluto entre as proporções esperadas e observadas, o tamanho da amostra e sua proporção esperada.

O nível de significância utilizado por Nigrini (2012) foi de 5%. Ressaltando que este teste realiza análise dos dígitos individualmente, e portanto não sinaliza a conformidade dos dados de forma agregada, assim o autor considera aceitável quatro ou cinco picos significativos no Teste dos Dois Primeiros Dígitos.

Bugarin e Cunha (2015) fazem a seguinte ponderação:

A posição de “n” na fórmula (3.5) faz com que, à medida que o volume de dados aumente, o resultado do Teste Z, para qualquer discrepância, torne-se mais alto. Isso significa que uma discrepância alta em um banco de dados pequeno pode não ser significativa, e uma discrepância pequena em um banco de dados grande pode ser significativa a um nível de 1 por cento.

### 3.3.2 Teste Qui-Quadrado

Estatística Qui-Quadrado de Pearsson ou Chi-Square (CS) compara a frequência absoluta e a frequência esperada dos conjuntos de dados com a finalidade de analisar se os dígitos de uma distribuição como um todo se ajustam a LB. A estatística é calculada pela fórmula:

$$\chi^2 = \sum_{i=1}^K \frac{(CR - CE)^2}{CE} \quad (3.6)$$

Onde (K) representa as categorias das séries dos dígitos ( $K = 9$ ,  $K = 10$ ,  $K = 90$ ), (CR) é contagem real dos números pertencentes a categoria, (CE) contagem esperada por Benford para a categoria. O número dos graus de liberdade é dado por  $(K - 1)$ . Se a estatística for menor que o limite tabulado, então não é possível rejeitar a hipótese de conformidade com a distribuição de Benford. Para 89 graus de liberdade e um nível de significância de 0.05, o valor crítico é 112,02. Caso contrário, rejeita-se a hipótese de conformidade. Cada dígito possui um valor qui-quadrado e a soma de todos leva ao resultado do teste, assim o resultado é comparado a um valor crítico. Quando as amostras são pequenas, esse teste é menos sensível aos desvios. Entretanto quando o tamanho da

amostra ( $n$ ) é grande, o teste se torna mais sensível a desvios (Bugarin e Cunha, 2015), (Silva, 2015).

### 3.3.3 Teste da Média dos Desvios Absolutos (MAD)

Estatística calculada independente do tamanho da amostra e portanto mais adequada para maiores bases de dados, segundo Nigrini (2012). Estimada a partir da média das frequências relativas e esperada para cada dígito descrita pela fórmula abaixo.

$$MDA = \sum_{i=1}^k \frac{|PR - PE|}{k} \quad (3.7)$$

Em que (PR) refere-se à frequência relativa encontrada para o primeiro dígito  $i$ , e (PE) frequência relativa esperada para  $i$  de acordo com a LB, (K) representa as categorias das séries dos dígitos (K=9, K=10, K=90). O sinal de módulo mostra que o desvio tem sempre valor positivo independente do resultado da comparação.

Diferentemente dos anteriores não existem valores críticos objetivos para esse teste, assim em Drake e Nigrini (2000) encontramos uma instrução baseada em um banco de valores críticos, conforme expresso a seguir:

Tabela 4 – Valores críticos e regiões de conformidade - Fonte: Adaptado de Drake e Nigrini (2000)

Dígitos	Intervalo	Conclusão
Primeiro Dígito	0,00 a 0,006	Conformidade aproximada
	0,006 a 0,012	Conformidade aceitável
	0,012 a 0,015	Conformidade marginal aceitável
	Acima de 0,015	Não conformidade
Segundo Dígito	0,000 - 0,008	Conformidade
	0,008 - 0,010	Conformidade Aceitável
	0,010- 0,012	Conformidade Marginalmente Aceitável
	Acima de 0,012	Não Conformidade
Dois Primeiros Dígitos	0,0000 de 0,0012	Conformidade aproximada
	0,0012 a 0,0018	Conformidade aceitável
	0,0018 a 0,0022	Conformidade marginal aceitável
	Acima de 0,0022	Não conformidade

## 3.4 Distribuição dos Votos na Eleição Presidencial de 2014

Uma referência mundial sobre o uso da lei de Benford para testar fraudes eleitorais é o trabalho de Mebane (2006). Mebane (2009) analisou as eleições iranianas em 2009 e descobriu que as cidades com poucos votos válidos, os números de Ahmadinejad não seguiam a distribuição de Benford. Nessas cidades o candidato eleito, Ahmadinejad, venceu com uma grande margem de votos (Bugarin e Cunha, 2015).

Para verificar se existe alguma relação entre a diferença percentual de votos entre os candidatos e a aderência da distribuição dos dados de cada candidato a LB, os estados foram divididos nas seguintes categorias: vitória com uma margem apertada, caso o candidato vencedor tenha vencido com um percentual de votos inferior a 60%; vitória com uma margem folgada, caso o candidato vencedor tenha vencido com um percentual de votos superior a 60%.

Foram realizados testes separadamente para cada código votável nas urnas, sendo eles 13, 45, 95, 96 que corresponde respectivamente ao número destinado à candidata do Partido dos Trabalhadores (PT), ao candidato do Partido da Social Democracia Brasileira (PSDB), aos Votos Brancos e aos Votos Nulos.

Tabela 5 – Descrição percentual do desempenho de cada candidato por estado - Fonte: elaboração própria

UF	DILMA 13	AÉCIO 45
AC	36,32%	63,68%
AL	62,11%	37,89%
AM	65,00%	35,00%
AP	61,44%	38,56%
BA	70,15%	29,85%
CE	76,73%	23,27%
DF	38,10%	61,90%
ES	46,13%	53,87%
GO	42,89%	57,11%
MA	78,76%	21,24%
MG	52,40%	47,60%
MS	43,65%	56,35%
MT	45,32%	54,68%
PA	57,41%	42,59%
PB	64,25%	35,75%
PE	70,19%	29,81%
PI	78,29%	21,71%
PR	39,00%	61,00%
RJ	54,92%	45,08%
RN	69,94%	30,06%
RO	45,14%	54,86%
RR	41,08%	58,92%
RS	46,46%	53,54%
SC	35,40%	64,60%
SE	66,99%	33,01%
SP	35,68%	64,32%
TO	59,47%	40,53%

A tabela acima foi classificada de acordo com os seguintes percentuais: acima de 60% definiu-se como vitória com uma margem folgada, e as células que satisfazem essa

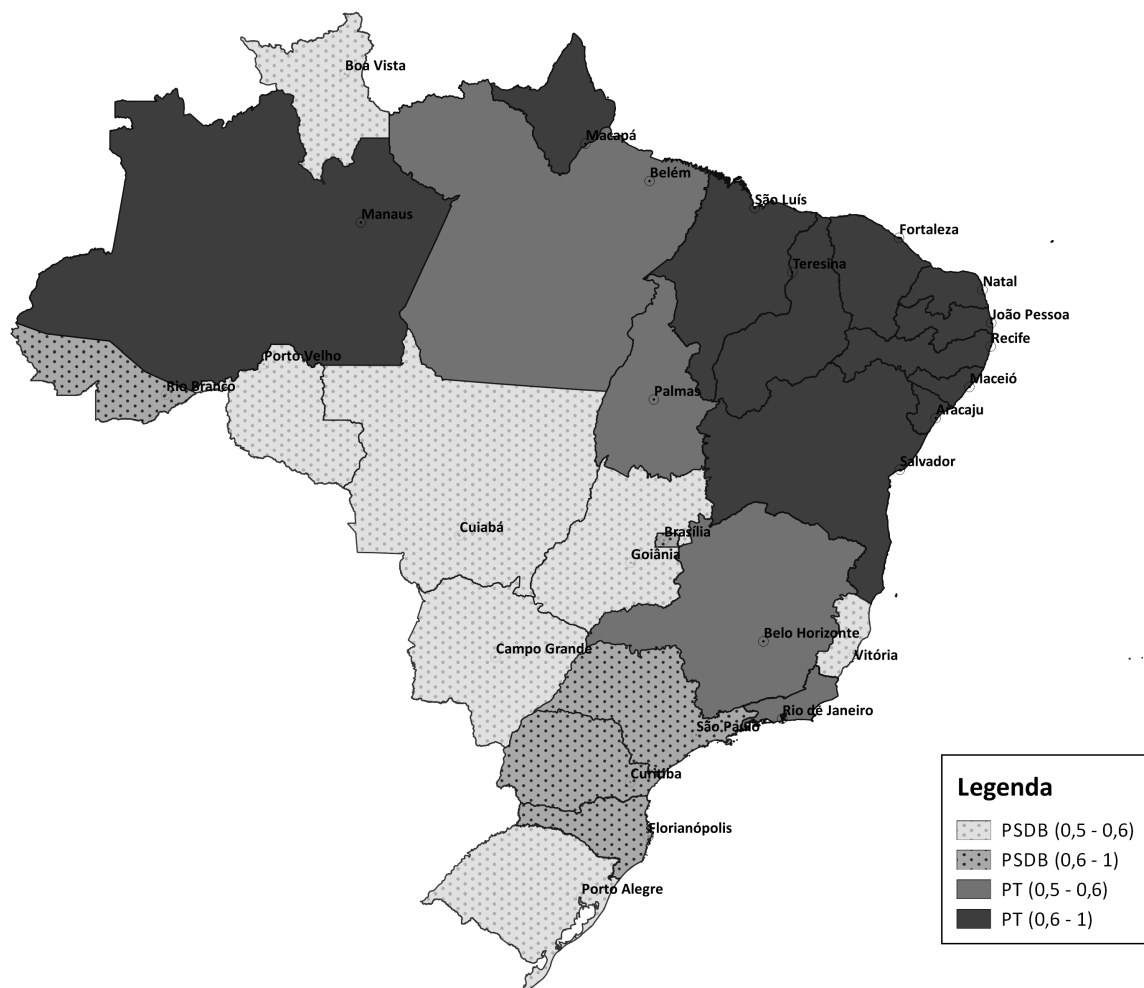
classificação estão marcadas com a cor de maior tonalidade; o intervalo de 50% a 60% foi classificado como vitória com uma margem apertada, e as células que satisfazem essa classificação estão marcadas com a cor com tonalidade mais clara. Os percentuais de 40% a 50% são classificados como derrota com uma margem apertada e abaixo de 40%, derrota com uma margem folgada. Em suma, as células foram gradualmente coloridas da tonalidade mais escura à tonalidade mais clara de acordo com a seguinte classificação: vitória com uma margem folgada, vitória com uma margem apertada, derrota com uma margem apertada e derrota com uma margem folgada.

Os estados que compõem a primeira categoria para a candidata Dilma Rousseff em ordem decrescente são: Maranhão (MA), Piauí (PI), Ceará (CE), Pernambuco (PE), Bahia (BA), Rio Grande do Norte (RN), Sergipe (SE), Amazonas (AM), Paraíba (PB), Alagoas (AL) e Amapá (AP). Para o candidato Aécio Neves: Santa Catarina (SC), São Paulo (SP), Acre (AC), Distrito Federal (DF), Paraná (PR).

O mapa abaixo sumariza o resultado das eleições considerando as unidades federativas e o percentual de votos para os candidatos segundo a classificação apresentada nessa seção.



Figura 1 – Percentual do desempenho de cada candidato por estado - Elaboração Própria





## 4 Análise

Realizaram-se os testes para todas as unidades federativas do Brasil. Devido à extensão e à quantidade de dados optou-se por inserir os resultados no apêndice. Em virtude da análise do Teste Z ser realizada por dígitos individualmente, essa avaliação foi considerada inviável de ser apresentada. O mesmo se aplica ao Teste da Soma. Portanto, os testes que serão apresentados são o Teste Qui-Quadrado, e o Teste da Média dos Desvios Absolutos.

Mebane (2006) realizou uma análise nos dados eleitorais de alguns países como os Estados Unidos, Rússia, México e descobriu que a contagem de votos tendia a seguir a Lei de Benford para o segundo dígito. Entretanto, ao avaliar as eleições iranianas em 2009, Mebane (2009, 2010) constatou que, em cidades com poucos votos nulos, os votos divergiram significativamente das previsões da distribuição Benford e o candidato vencedor, nestas situações, detinha uma grande vantagem na votação. O enfoque na distribuição do segundo dígito também é dada no trabalho de Pericchi e Torres (2011).

No entanto, no presente trabalho essa tese não se corrobora à medida que as regiões onde há baixa porcentagem de votos nulos e brancos são em sua expressiva maioria do candidato que ficou em segundo nos resultados. Com o intuito de avaliar se as dificuldades e não conformidades encontradas proviam da Lei de Benford, ou se os testes utilizados para avaliação não são adequados para este tipo de base. Ou seja, como os números observados diferiram substancialmente do que é previsto pela lei, isso poderia ser um indício de manipulação dos dados ou de algum outro fato atípico. Então foi criado um instrumento para avaliar essas conjecturas. A partir dos dados das eleições presidenciais dos anos de 2002, 2006 e 2010 foi criada uma distribuição histórica para o Brasil. A distribuição foi construída a partir dos dados disponibilizada pelo TSE sobre o segundo turno para presidente nas eleições de 2002, 2006 e 2010, segue a ideia de Benford quanto à separação dos dígitos. Os dados e planilhas foram organizados como em Silva (2015)<sup>1</sup> e Café (2015)<sup>2</sup>

### 4.1 Distribuição Histórica Brasileira

Para tentar compreender o fato que aproximadamente metade das unidades federativas do Brasil estão na zona de desconformidade com Benford quando o Teste da Média dos Desvios Absolutos é considerado, e nenhum estado apresenta conformidade para os dois candidatos, quando o Teste Qui-Quadrado é considerado, foi estimada uma distribuição histórica brasileira que pudesse explicar tais padrões. Para tanto, foram escolhidos

<sup>1</sup> <[http://bdm.unb.br/bitstream/10483/12151/1/2015\\_RafaelRichterOliveiradaSilva.pdf](http://bdm.unb.br/bitstream/10483/12151/1/2015_RafaelRichterOliveiradaSilva.pdf)>

<sup>2</sup> <[http://bdm.unb.br/bitstream/10483/11809/1/2015\\_RenataMottaCafe.pdf](http://bdm.unb.br/bitstream/10483/11809/1/2015_RenataMottaCafe.pdf)>

os anos de 2002, 2006 e 2010. Esses anos foram escolhidos considerando a disponibilidade de tempo e a limitação computacional ao manusear uma base de dados tão extensa.

Os dados foram retirados do repositório de dados eleitorais do TSE<sup>3</sup>. Para construir a distribuição histórica brasileira foram pegos os dados referentes ao candidato vencedor nas eleições de 2002, 2006 e 2010. A escolha pelos dados do candidato vencedor das eleições passadas se deve ao fato que um desvio dos dados referentes ao candidato vencedor nas eleições de 2014 pudesse ser notado com maior clareza em relação a sua distribuição histórica. A seguir os dígitos das três eleições foram contados e foi construída a frequência observada de cada dígito. A distribuição é calculada através de um processo automatizado desenvolvido no software estatístico STATA, cujo roteiro de comandos esta descrito no apêndice. Por meio desta programação foi encontrada a frequência dos grupos do Primeiro Dígito (1, 2 até 9), Segundo Dígito (0, 1 até 9), e dos Dois Primeiros Dígitos (10, 11 até 99), assim segue a ideia de Benford quanto à separação dos dígitos. Para realização dos novos testes, os valores da distribuição de Benford foram substituídos pelos valores encontrados da distribuição Histórica.<sup>4</sup>

## 4.2 Teste Qui-Quadrado (CS)

Na tabela 5, a primeira coluna se refere ao código ou número votável nas eleições, a segunda representa a unidade federativa (UF), a terceira e quarta contém resultados do Teste Qui-Quadrado segundo o código votável de cada candidato. O valor crítico para o teste do segundo dígito com nove graus de liberdade e 0,05 de significância é 16,919. A análise detalhada dos valores com respeito à rejeição, ou não da hipótese nula para os segundos dígitos de cada estado, constam no apêndice.

Os resultados deste teste estão apresentados nas tabelas mostradas a seguir.

<sup>3</sup> As informações sobre a base de dados seguem os direcionamentos a seguir: Eleições> Estatísticas eleitorais> Repositório de dados eleitorais> Página inicial> Resultados> Ano > Votação por seção eleitoral- ano > Presidente (formato ZIP). E encontram-se disponíveis em: <<http://www.tse.jus.br/eleicoes/estatisticas/repositorio-de-dados-eleitorais>>

<sup>4</sup> Um modelo de tabela para cálculo pode ser encontrado em <[www.nigrini.com/BenfordLaw/NigriniCycle.xlsx](http://www.nigrini.com/BenfordLaw/NigriniCycle.xlsx)>

Tabela 6 – Teste Qui- Quadrado (CS) para o Segundo Dígito aplicado aos votos de cada candidato sob a distribuição de Benford - Fonte: elaboração própria

Num_votável	UF	CS_13	CS_45
45	AC	110,95	63,29
13	AL	52,33	66,89
13	AM	241,94	59,82
13	AP	21,54	93,19
13	BA	198,65	121,20
13	CE	129,34	14,82
45	DF	739,61	283,19
45	ES	139,77	67,84
45	GO	956,30	177,32
13	MA	42,36	9,97
13	MG	711,13	608,53
45	MS	254,09	46,39
45	MT	66,55	179,05
13	PA	128,66	176,91
13	PB	172,81	27,66
13	PE	182,76	281,21
13	PI	33,07	14,49
45	PR	612,82	74,15
13	RJ	186,55	833,80
13	RN	81,98	46,45
45	RO	165,78	225,83
45	RR	92,54	21,48
45	RS	768,54	685,81
45	SC	635,61	90,01
13	SE	36,19	69,25
45	SP	1123,76	171,11
13	TO	158,49	51,91

Ao analisar a quarta coluna, o resultado do teste é menor que o valor crítico nos estados do CE, MA e PI, assim apenas estes apresentam conformidade com a Lei de Benford e, portanto há rejeição da hipótese nula para o restante.

Quanto a análise conjunta dos dados percebemos que apesar de muitos estados possuírem grande discrepância quanto ao valor crítico, apenas o DF e os estados de MG, RS aparecem com valores muito altos para ambos. O que não condiz com a análise em separado dos candidatos, visto que para análise dos votos da candidata Dilma indica maior índice em SP, e para o candidato Aécio, o RJ.

As tabelas a seguir demonstram os resultados do teste MDA, classificados de acordo com Drake e Nigrini (2000) segundo sua Conformidade (C), Conformidade Aceitável (CA), Conformidade Marginal Aceitável (CMA) e Não Conformidade (NC). As unidades federativas que obtiveram desconformidade tanto para o dígito 13, quanto para o 45 foram o Acre (AC), Distrito Federal (DF), Roraima (RO) e Rio Grande do Sul (RS). Os resultados

detalhados encontram-se reportados no apêndice.

Tabela 7 – Teste Qui- Quadrado (CS) para o Segundo Dígito aplicado aos votos de cada candidato sob a distribuição do Brasil - Fonte: elaboração própria

Num_votável	UF	CS_13	CS_45
45	AC	108,81	37,96
13	AL	17,46	25,71
13	AM	299,14	17,90
13	AP	38,29	61,50
13	BA	77,46	139,05
13	CE	44,44	126,94
45	DF	668,42	207,61
45	ES	74,98	18,48
45	GO	680,63	55,77
13	MA	135,12	87,54
13	MG	217,73	190,75
45	MS	170,16	17,59
45	MT	30,05	80,93
13	PA	35,30	72,80
13	PB	108,87	60,89
13	PE	469,99	332,24
13	PI	85,48	75,47
45	PR	332,14	109,47
13	RJ	326,23	408,76
13	RN	144,77	51,33
45	RO	126,99	138,25
45	RR	73,88	35,67
45	RS	362,03	292,64
45	SC	481,90	251,30
13	SE	87,38	62,60
45	SP	421,48	330,01
13	TO	88,45	47,16

Como podemos observar através da análise da tabela 6, embora a disparidade quanto ao valor crítico seja menor que em relação à distribuição de Benford, o valor calculado excedeu o valor crítico de 16,919 em todos os estados, logo se rejeitou a hipótese nula de que as frequências dos segundos dígitos dos votos (sob o ponto de vista de uma análise geral do conjunto) se conformavam com as da distribuição brasileira. O que gerou surpresa, pois se esperava que os dados convergissem nesta nova distribuição. Assim aparecem as hipóteses: I) houve fraude nos resultados das eleições, visto que eles não apresentam conformidade com relação a distribuição brasileira. II) Este tipo de teste não é adequado para esta base, já que obtivemos resultados aproximados de não conformidade, mesmo utilizando distribuições diferentes.

Quanto à análise conjunta dos dados percebemos que o código de regiões com alta

discrepância para ambos os candidatos é composta pelo DF, PE, RS, SC e SP. O que novamente não condiz com a análise em separado dos candidatos, visto que para análise dos votos da candidata Dilma indique maior índice em GO, e para o candidato Aécio o RJ.

### 4.3 Teste da Média dos Desvios Absolutos

O último teste aplicado foi o MDA. As tabelas a seguir demonstram os resultados do teste MDA, classificados de acordo com Drake e Nigrini (2000). As unidades federativas que obtiveram desconformidade tanto para o dígito 13, quanto para o 45 foram o Acre (AC), Distrito Federal (DF), Roraima (RO) e Rio Grande do Sul (RS). Os resultados detalhados encontram-se reportados no apêndice.

Tabela 8 – Teste da Média dos Desvios Absolutos (MAD) para o Segundo Dígito aplicado aos votos de cada candidato sob a distribuição de Benford - Fonte: elaboração própria

Num_votável	UF	MAD_13	Classificação	MAD_45	Classificação
45	AC	0,02072	NC	0,01821	NC
13	AL	0,00703	C	0,00847	CA
13	AM	0,01716	NC	0,00890	CA
13	AP	0,00956	CA	0,02301	NC
13	BA	0,00705	C	0,00547	C
13	CE	0,00696	C	0,00242	C
45	DF	0,03021	NC	0,01668	NC
45	ES	0,00966	CA	0,00776	C
45	GO	0,02485	NC	0,00911	CA
13	MA	0,00420	C	0,00223	C
13	MG	0,01030	CMA	0,01025	CMA
45	MS	0,01932	NC	0,00712	C
45	MT	0,00881	CA	0,01347	NC
13	PA	0,00689	C	0,00947	CA
13	PB	0,01174	CMA	0,00407	C
13	PE	0,00792	C	0,00925	CA
13	PI	0,00485	C	0,00375	C
45	PR	0,01455	NC	0,00459	C
13	RJ	0,00600	C	0,01445	NC
13	RN	0,01022	CMA	0,00681	C
45	RO	0,01941	NC	0,02074	NC
45	RR	0,02734	NC	0,01064	CMA
45	RS	0,01418	NC	0,01363	NC
45	SC	0,01871	NC	0,00702	C
13	SE	0,00780	C	0,01027	CMA
45	SP	0,01010	CMA	0,00319	C
13	TO	0,01803	NC	0,01016	CMA

Para Dilma os estados que não convergem são: TO, AM com relação aos valores na terceira coluna e AP, RJ com relação à quinta coluna. Para Aécio com relação à terceira coluna: DF, RR, GO, AC, RO, MS, SC, PR, RS. Com relação à quinta coluna: RO, AC, DF, RS, MT.

Para Dilma os estados que possuem conformidade marginalmente são: PB, MG, RN na terceira coluna e SE, MG, TO na quinta coluna. Para Aécio, os estados que possuem conformidade marginalmente aceitáveis são: SP na terceira e RR na quinta coluna.

Para Dilma os estados que possuem conformidade aceitável são: AP com relação à terceira coluna e PA, PE, AM, AL com relação à quinta coluna. Para Aécio com relação à terceira coluna: ES, MT. Com relação à quinta coluna: GO.

Para Dilma os estados que convergem são: PE, SE, BA, AL, CE, PA, RJ, PI, MA com relação à terceira coluna e RN, BA, PB, PI, CE, MA com relação à quinta coluna. Para o Aécio, na quinta coluna, os estados são ES, MS, SC, PR, SP.

Entre os estados que não estão na zona de conformidade e que pertencem à primeira categoria quanto à porcentagem, acima de 60% dos votos, está o AM para o candidato incumbente e SC, AC, DF e PR para o candidato desafiante quando os testes são realizados sob o código votável 13 representados pelos valores da terceira coluna. Seguindo a mesma ordem dos candidatos os resultados para os testes com o código votável 45 na quinta coluna são: AP para o primeiro, e AC, DF e PR para o segundo.

Não se visualiza recorrência entre os estados, o que dificulta a análise. Deste modo, as grandes discrepâncias entre a lei de Benford e o número de votos em alguns estados parecem decorrer, em grande medida, do próprio desvio já presente nas distribuições da população e do eleitorado.

Buscando ampliar nossa visão sobre a pesquisa, verificaremos a seguir os resultados decorrentes do teste utilizando a distribuição brasileira com o intuito de observar se haverá contraste com os resultados obtidos sob a distribuição de Benford.

Para Dilma os estados que não convergem são: AM, RN, AP, TO, PE, SE com relação à interpretação dos valores da terceira coluna e AP com relação à quinta coluna. Para Aécio com relação à terceira coluna: DF, RR, GO, AC, RO, SC, MS. Com relação à quinta coluna: RO, DF, RR, AC.



Tabela 9 – Teste da Média dos Desvios Absolutos (MAD) para o Segundo Dígito aplicado aos votos de cada candidato sob a distribuição do Brasil - Fonte: elaboração própria

Num_votável	UF	MAD_13	Classificação	MAD_45	Classificação
45	AC	0,02037	NC	0,01368	NC
13	AL	0,00445	C	0,00540	C
13	AM	0,01897	NC	0,00435	C
13	AP	0,01347	NC	0,01904	NC
13	BA	0,00388	C	0,00548	C
13	CE	0,00406	C	0,00727	C
45	DF	0,02793	NC	0,01473	NC
45	ES	0,00616	C	0,00381	C
45	GO	0,02072	NC	0,00513	C
13	MA	0,00826	CA	0,00611	C
13	MG	0,00549	C	0,00557	C
45	MS	0,01570	NC	0,00426	C
45	MT	0,00533	C	0,00891	CA
13	PA	0,00324	C	0,00549	C
13	PB	0,00912	CA	0,00652	C
13	PE	<b>0,01238</b>	NC	0,01079	CMA
13	PI	0,00874	CA	0,00796	C
45	PR	0,01080	CMA	0,00608	C
13	RJ	0,00845	CA	0,00996	CA
13	RN	0,01378	NC	0,00792	C
45	RO	0,01713	NC	0,01586	NC
45	RR	0,02506	NC	0,01411	NC
45	RS	0,00937	CA	0,00882	CA
45	SC	0,01603	NC	0,01158	CMA
13	SE	<b>0,01236</b>	NC	0,00966	CA
45	SP	0,00555	C	0,00476	C
13	TO	0,01316	NC	0,00972	CA

Para Dilma o único estado que esta na zona de conformidade marginalmente aceitável é PE, conforme valores da quinta coluna. Para Aécio são os estados do PR e SC, na terceira e quinta coluna respectivamente.

Para Dilma os estados que possuem conformidade aceitável são: PB, PI, RJ, MA com relação à 3ª coluna e RJ, TO, CE com relação à quinta coluna. Para Aécio com relação à terceira coluna: RS. Com relação à quinta coluna: MT e RS.

Para Dilma os estados que convergem são: MG, AL, CE, BA, PA com relação à quarta coluna e PI, RN, CE, PB, MA, MG, PA, BA, AL, AM com relação à quinta coluna. Para o Aécio com relação à terceira coluna: ES, SP, MT. E PR, GO, SP, MS, ES Com relação à quinta coluna.

Os estados que estão na zona de desconformidade e que pertencem à primeira

categoria quanto à porcentagem, acima de 60% dos votos, estão os estados PE, RN, SE, AM e AP para o código votável 13, e os estados de SC, AC e DF para o código votável 45. Sendo que PE, SE, estão bem próximos da linha de conformidade marginal aceitável. Percebemos que a distribuição entre os candidatos esta mais uniforme, visto que cada um fica com aproximadamente metade desses estados como se pode observar na quarta coluna. Já na sexta coluna, os estados que estão em desconformidade e pertencem à primeira categoria quanto à porcentagem são: o AP para o código votável 13 e os estados de SC, AC e o DF para o código votável 45, portanto observa-se assimetria quanto à quantidade de regiões para cada candidato.

## 5 Conclusão

Esse trabalho testou a aplicação da Lei de Benford (LB) para avaliação dos resultados das eleições presidenciais de 2014. Como ferramentas empregadas foram utilizados alguns testes, dentre eles o Teste do Segundo Dígito e para mensurar a conformidade com a LB foi aplicado o Teste Qui-Quadrado (CS), e a Média dos Desvios Absolutos (MDA).

Mediante avaliação do perfil de dados não foram encontradas evidências contrárias ao uso da base de dados, porém nenhum dos critérios empregados se mostrou adequado, uma vez que ao se observarem os resultados, é notável a recorrência de desconformidade, o que não proporciona um diagnóstico conciso. A desconformidade dos dados com a lei de benford é um fenômeno encontrado com frequência em dados eleitorais, como salienta Mebane: “the evidence is strong that departures from 2BL, which also occur frequently, are related both to normal political phenomena and to serious election anomalies <sup>1</sup>” (Mebane, 2010 ).

Como pode ser comprovado acerca do desfecho do teste da MDA, pois os resultados não possuem padrão de convergência quando analisados em conjunto, e quando analisados individualmente apresentam reincidência de desconformidade nas unidades regionais onde há baixo percentual de votos brancos e nulos, o que poderia ser interpretado como um maior grau de manipulação de dados na campanha eleitoral. Segundo Mebane (2009), tais regiões deveriam apresentar uma quantidade expressiva de votos favoráveis ao candidato vencedor, mas nessas regiões, o candidato vencedor das eleições tem uma quantidade de votos inferior à média que recebeu nas demais unidades federativas.

Com a distribuição histórica para o Brasil buscou-se responder a crítica aos resultados. Havia a percepção de que se fossem encontrados seguimentos discrepantes da distribuição de Benford, essa seria uma prova de que o problema está com a distribuição e não com os testes relacionados a ela. De modo que foram encontradas as mesmas disfunções como em Benford, e, portanto, não é possível chegar a uma conclusão clara quanto à distribuição brasileira.

Sabe-se que a aplicação da lei de benford deve ser interpretada com cautela (Mebane, 2006). As fraudes eleitorais não são independentes da localidade e não ocorrem sempre da mesma maneira, portanto, seus indicadores requerem uma análise cuidadosa, e o resultado dos testes devem ser considerados em seu conjunto (Levin, Cohn, Ordeshook e Alvarez, 2009), o que novamente impossibilita chegar-se a um resultado peremptório.

Outro indício da permanência da desconformidade está relacionado ao desvio que os dados de 2002 possuem para o candidato do PT, visto que neste ano a eleição foi ganha no segundo turno com uma diferença expressiva no percentual para este candidato,

---

<sup>1</sup> Há uma forte evidência que um desvio da 2LB, que também ocorre frequentemente, está relacionada tanto com fenômenos políticos normais quanto com anomalias eleitorais graves.

e devido à pequena quantidade de anos utilizados de 2002, 2006 e 2010, uma observação com valores extremos pode ter afetado a inferência da real distribuição brasileira.

Aparentemente, o tamanho da base por estado não se adequa aos testes. Buscando extrair padrões na quantidade de zonas e seções eleitorais, uma proposta seria realizar os mesmos testes para os municípios acreditando que regiões com maior dispersão tenderiam a seguir a lei para o eleitorado (e para o número de votos). Tal observação pode ser inferida dos dados, onde a distribuição de alguns dos estados mais populosos do Brasil como São Paulo, Minas Gerais, Bahia, conformam com a distribuição da Lei de Benford.

Quanto aos testes, conclui-se que o teste da média dos desvios absolutos para o segundo dígito é o teste ao qual a distribuição dos dígitos observada melhor se adequa à lei de Benford e à distribuição histórica brasileira. Já o teste qui-quadrado deve ser analisado com cautela, já que o mesmo é mais sensível observações extremas, o que está de acordo com a literatura, como em Mebane (2006, 2010), Percchi e Torres (2004, 2011).

## Referências

- Aguiar, G. TSE pede à PF e Procuradoria investigação sobre falha em 77 mil votos de 2014. Disponível em:< <http://politica.estadao.com.br/noticias/geral,tse-pede-a-pf-e-procuradoria-investigacao-sobre-falha-em-40-mil-votos-de-2014,10000056311>>. Acesso em: 2 de julho. 2016.
- Benford, F. (1938) “The law of anomalous numbers”. *Proceedings of the American Philosophical Society* 78 (4), 551-572.
- Berger, A., e Hill, T. (2011). A basic theory of Benford’s Law. *Probability Surveys*, 8, 1–126.
- Bugarin, M.; Cunha, F.C.R “Benford’s Law for audit of public works: an analysis of overpricing in Maracanã soccer arena’s renovation”. *Economics Bulletin*, v.35, p. 1168-1176, 2015.
- Bugarin, M.; Cunha, F.C.R. “A didactic note on the use of Benford’s Law in public works auditing, with an application to the construction of Brazilian Amazon Arena 2014 World Cup soccer stadium”. *Economia (Yokohama)*,2016, no prelo.
- Cunha, F.C.R.; Bugarin, M. S. Lei de Benford para a auditoria de obras públicas: análise de sobrepreço na construção da Arena da Amazônia. *BLC – Boletim de Licitações e Contratos*, São Paulo, NDJ, ano 28, n. 8, p. 841-857, ago. 2015.
- Cunha, F.C.R. e Bugarin, M. S. (2015), “Benford Law for audit of public works: An analysis of overpricing in Maracanã soccer arena’s renovation”. *Economics Bulletin*. 35, p.120-129.
- Cunha, F.C.R. (2013), “Aplicações da Lei Newcomb-Benford à Auditoria de Obras Públicas”. *Dissertação (Mestrado em Regulação e Gestão de Negócios) – Universidade de Brasília*, Brasília, 486p.
- Cho, W. T. e B. Gaines (2007) “Breaking the (Benford) Law: Statistical Fraud Detection in Campaign Finance”. *The American Statistician* 61(3): 218-223.
- Café, R. M. O uso da lei de Benford na auditoria de obras públicas: o caso do VLP. 2015. Trabalho de Conclusão de Curso (monografia bacharelado em economia), p 54.

- Carslaw, C. (1988). Anomalies in income numbers: Evidence of goal oriented behavior. *The Accounting Review*, 63 (April), 321–327.
- Deckert, J., Myagkov, M. e Ordersook, P. C., (2011) Benford's Law and the Detection of Election Fraud. *Political Analysis*, 19(3): 245-268.
- Drake, P. e Nigrini, M. (2000) "Computer assisted analytical procedures using Benford's Law". *Journal of Accounting Education*, v. 18, p. 127-146.
- Durtschi, C., Hillison, W., e Pacini, C. (2004). The effective use of Benford's Law to assist in detecting fraud in accounting data. *Journal of Forensic Accounting*, 5(1), 17–34.
- Feller, W. (1966). *An introduction to probability theory and its applications*. New York, NY: John Wiley & Sons.
- Flehinger, B. (1966). On the probability that a random integer has initial digit "A." *The American Mathematical Monthly*, 73 (10), 1056–1061.
- Furry, W., e Hurwitz, H. (1945). Distribution of numbers and distribution of significant figures. *Nature*, 155, 52–53.
- Goudsmit, S., e Furry, W. (1944). Significant figures of numbers in statistical tables. *Nature*, 154, 800–801.
- Hamilton, L.C. (2012). *Statistics with Stata*. University of New Hampshire.
- Hill, T. (1988). Random number guessing and the first digit phenomenon. *Psychological Reports*, 62 (3), 967–971.
- Hill, T. P. (1995). A statistical derivation of the significant-digit law. *Statistical Science*, 10(4), 354–363.
- Hill, T. P. (1995). Base-invariance Implies Benford's law. *Proceedings of the American Mathematical Society*, 123 (3), 887-895.
- Hsu, E. (1948). An experimental study on "mental numbers" and a new application. *The Journal of General Psychology*, 38, 57–67.
- Kinzo, M. D. G. 2001. " A democratização brasileira: um balanço do processo político

desde a transição”. São Paulo em Perspectiva, v.15 n.4, p. 3-12, out.\dez.

Lagioia, U. C. T.; Araujo, I. J. C.; Alves Filho, B. F.; Barros, M. A. B., e Nascimento, S. G. O. de A. S. do. Aplicabilidade da Lei de Newcomb-Benford nas fiscalizações do Imposto sobre Serviços – ISS. Revista de Contabilidade e Finanças – USP, Sao Paulo, 22 (56) 203-224. Maio/junho/julho/agosto de 2011.

Levin, I., Gabe A. C., Peter C. O. e Alvarez. R. M. (2009). “Detecting Voter Fraud in an Electronic Voting Context: An Analysis of the Unlimited Reelection Vote in Venezuela.” 2009 Electronic Voting Technology Workshop/Workshop on Trustworthy Elections, Montreal, August 10–11.

Mebane, W. R. (2006) “Election Forensics: vote counts and Benford’s Law”. Summer Meeting of the Political Methodology Society. Papers, Posters and Syllabi. No 620.

Mebane, W.R. (2010) "Election Fraud or Strategic Voting? Can Seconddigit Tests Tell the Difference?" Summer Meeting of the Political Methodology Society, University of Iowa.

Mebane, W. R. (2010) “Fraud in the 2009 presidential election in Iran?” Chance 23(1), 6-15.

Mebane, W. R. (2009) “Note on the presidential election In Iran”. Disponível em <http://www.benfordonline.net/fullreference/756>. Acessado em 30 de março de 2014.

Mosimann, J., Wiseman, C., e Edelman, R. (1995). Data fabrication: Can people generate random digits? Accountability in Research, 4 (1), 31–55.

Newcomb, S. (1881) “Note on the frequency of the different digits in natural numbers”. The American Journal of Mathematics 4(1), 39-40.

Nigrini, M. J. (1992) The Detection of Income Tax Evasion Through an Analysis of Digital Frequencies. Ph.D. thesis. Cincinnati, OH: University of Cincinnati.

Nigrini, M. J.(2000) Digital analysis using Benford’s Law: Tests Statistics for Auditors. Global Audit Publications: Vancouver.

Nigrini, M.J. (2011) Forensic Analytics: Methods and Techniques for Forensic Accounting Investigations. John Wiley & Sons, Inc.: Hoboken, New Jersey.

Nigrini, M. J. (2012) Benford's Law. Applications for Forensic Accounting Auditing, and Fraud Detection. John Wiley & Sons, Inc.: Hoboken, New Jersey.

Nigrini, M., e Mittermaier, L. (1997). "The use of Benford's Law as an aid in analytical procedures". Auditing: A Journal of Practice and Theory, 16(2), 52-67.

Nigrinicycle. Disponível em: <<http://www.nigrini.com/BenfordsLaw/NigriniCycle.xlsx>>. Acesso em: 19.mai.2016.

Pericchi, L.R. e Torres, D. "La Ley de Newcomb-Benford y sus aplicaciones al Referendum Revocatorio en Venezuela." Reporte Técnico no-definitivo 2a. versión: Octubre 01,2004.

Pericchi, L. e Torres, D. (2011) "Quick Anomaly Detection by the Newcomb-Benford Law, with Applications to Electoral Processes Data form the USA, Puerto Rico and Venezuela". Stastical Science , 26 (4), 502-516.

Pinkham, Roger S. (1961) On the distribution of first significant digits. The Annals of Mathematical Statistics. Vol. 32, n.4, pp.1223-1230.

Pitta, I. Veja como funciona a urna eletrônica no Brasil e em outros países. Disponível em:< <http://politica.estadao.com.br/noticias/geral,veja-como-funciona-a-urna-eletronica-no-brasil-e-em-outros-paises,10000002427>>. Acesso em: 2 de julho. 2016.

Santos, J.; Diniz, J. A.; Corrar, L. O Foco é a Teoria Amostral nos Campos da Auditoria Contábil Tradicional e da Auditoria Digital: testando a Lei de Newcomb-Benford para o primeiro dígito nas contas públicas. Brazilian Business Review 2 (1), 71-89. 2005.

Silva, R. R.O. Em busca de transparência: a lei de Benford aplicada às despesas eleitorais. 2015. 85 f. Trabalho de Conclusão de Curso (monografia em bacharelado em economia), pg 6.

Tsao, N. (1974). On the distribution of significant digits and roundoff errors. Communications of the ACM, 17, 269-271.

Varian, H. (1972). Benford's Law. The American Statistician, 23 (June), 65-66.



Link Dados

Tribunal superior eleitoral (TSE). Disponível em: <<http://www.tse.jus.br/eleicoes/estatisticas/repositorio-de-dados-eleitorais>>. Acesso em: 25.mar.2016.

Google Drive

<<https://drive.google.com/drive/folders/0B5N1NvqnAD-xT2hRZlk5TkxzRnM>>



# Apêndices



# APÊNDICE A – Resultado da Apuração - Eleições 2014

Figura 2 – Resultado das Eleições Presidenciais de 2014 - Fonte: Elaboração Própria



Figura 3 – Descrição percentual da quantidade de votos para o código votável 13 - Fonte: elaboração própria

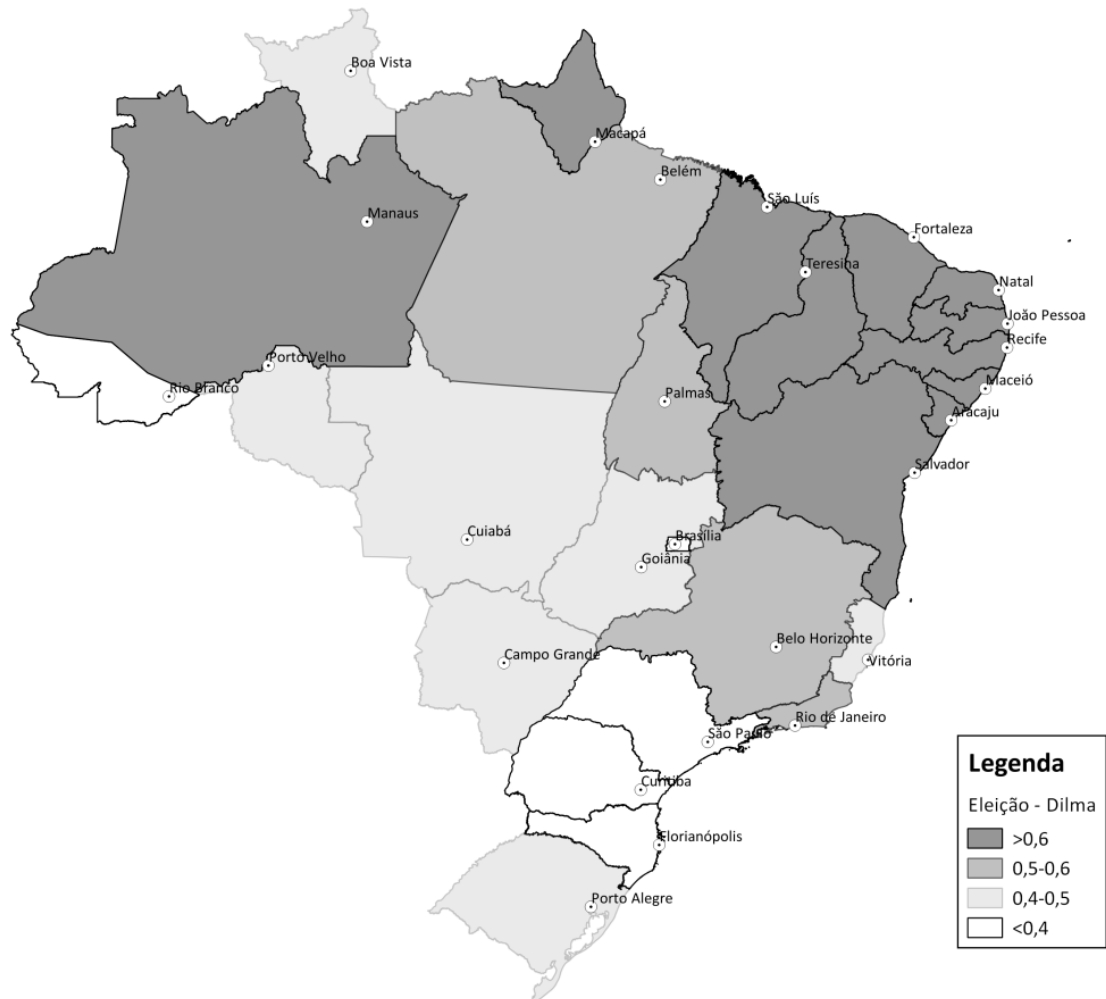
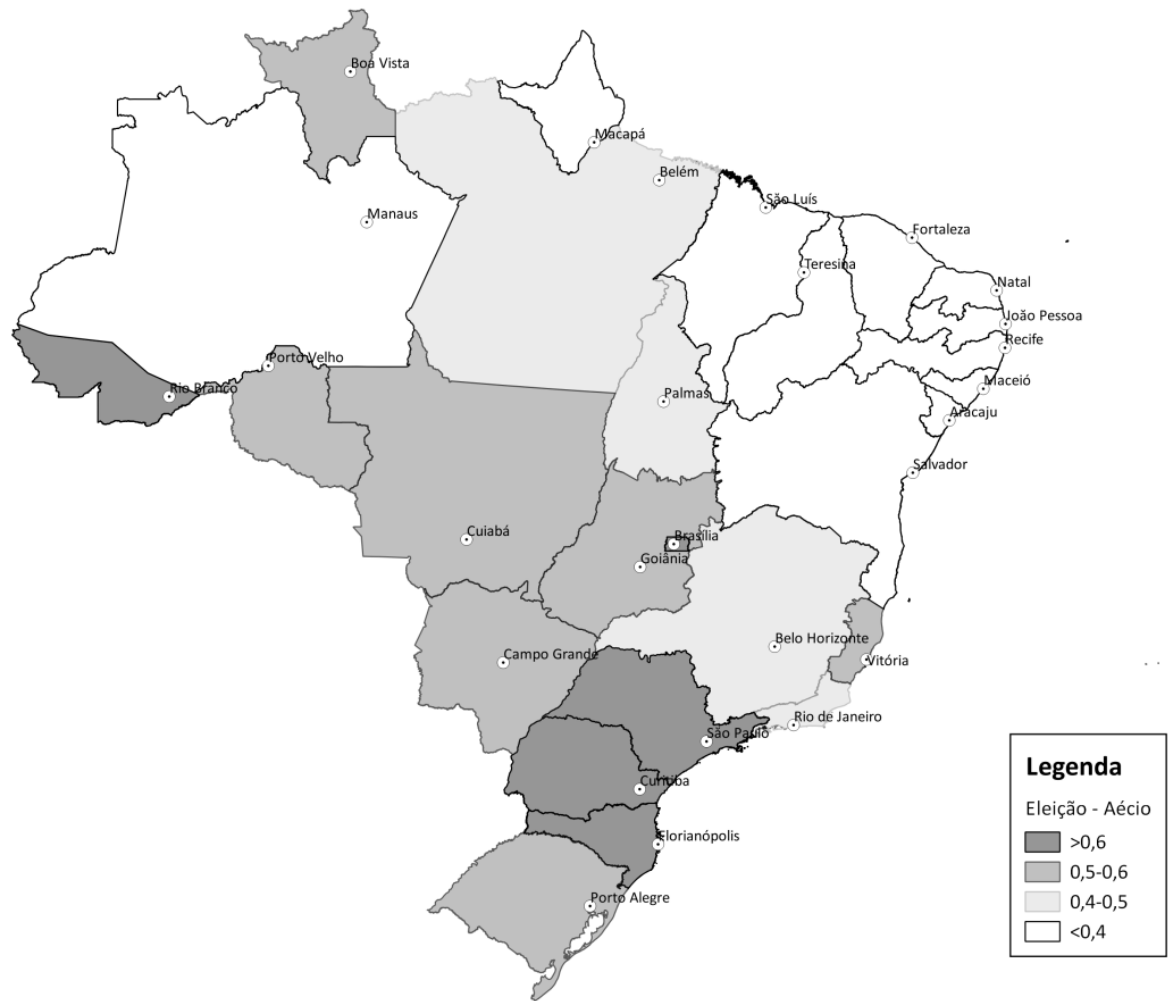


Figura 4 – Descrição percentual da quantidade de votos para o código votável 45 - Fonte: elaboração própria



A base de dados é disponibilizada pelo Tribunal Superior Eleitoral (TSE) <sup>1</sup> e tabela a seguir mostra o resultado da apuração da quantidade de votos destinados ao código votável 95 e 96.

Tabela 10 – Descrição percentual da quantidade de votos brancos e nulos por estado -  
Fonte: elaboração própria

<b>SIGLA_UF</b>	<b>PCT_BRANCOS</b>	<b>PCT_NULOS</b>
AC	0,91%	2,15%
AL	1,65%	4,18%
AM	1,71%	6,68%
AP	0,86%	4,28%
BA	1,51%	4,63%
CE	1,51%	5,38%
DF	1,91%	6,77%
ES	1,84%	3,21%
GO	1,76%	4,95%
MA	1,10%	2,79%
MG	1,54%	3,75%
MS	1,15%	2,10%
MT	0,93%	2,20%
PA	1,16%	4,67%
PB	2,03%	6,14%
PE	2,08%	4,61%
PI	0,90%	3,47%
PR	1,32%	2,55%
RJ	3,20%	12,12%
RN	2,03%	9,51%
RO	1,24%	4,44%
RR	0,96%	4,71%
RS	2,27%	4,09%
SC	1,33%	3,01%
SE	1,66%	4,41%
SP	2,10%	4,78%
TO	0,95%	3,14%

<sup>1</sup> <http://www.tse.jus.br/eleicoes/estatisticas/repositorio-de-dados-eleitorais>



# APÊNDICE B – Código STATA para leitura e processamento de dados

O software STATA versão 14 foi utilizado para realizar as análises.

## Preparação dos Dados no STATA

Como tamanho das bases é extenso o processo de preparação é fundamental, pois simplifica a quantidade de informações. Disponibilizaremos o exemplo para o ano de 2010, mas os outros anos (2002, 2006) seguem este padrão. Ou seja, é necessário realizar estes passos para cada um dos anos.

Figura 5 – Stata

```

2  *Preparação da base de dados das eleições presidenciais segundo turno para o ano 2010
3
4  global dados "Z:\Equipe\Erica\MONO- Dados\TSE_votos\votacao_secao_2010_BR"
5  local file "${dados}\votacao_secao_2010_BR.txt"
6
7  * Lê o arquivo de dados
8  import delimited using "`file'", delimiters(";") varnames(nonames) clear
9
10 * Renomeia as variáveis
11 rename v1 DATA_GERACAO
12 rename v2 HORA_GERACAO
13 rename v3 ANO_ELEICAO
14 rename v4 NUM_TURNO
15 rename v5 DESCRICAO_ELEICAO
16 rename v6 SIGLA_UF
17 rename v7 SIGLA_UE
18 rename v8 CODIGO_MUNICIPIO
19 rename v9 NOME_MUNICIPIO
20 rename v10 NUM_ZONA
21 rename v11 NUM_SECAO
22 rename v12 CODIGO_CARGO
23 rename v13 DESCRICAO_CARGO
24 rename v14 NUM_VOTAVEL
25 rename v15 QTDE_VOTOS
26
27 * Muda o formato do número, para melhor visualização
28 format QTDE_VOTOS %20.0fc
29
30 * Cargo
31 tab DESCRICAO_CARGO
32
33 * Mantém apenas votos válidos e segundo turno
34 keep if NUM_TURNO==2
35 keep if NUM_VOTAVEL==13
36 save "Z:\Equipe\Erica\MONO- Dados\STATA\dados\votacao_geral_secao_2010.dta", replace
37
38

```

Figura 6 – Stata

```

39  *---- Calcula para Brasil ----*
40
41  * Calcula o total de votos de cada candidato
42  use "Z:\Equipe\Erica\MONO- Dados\STATA\dados\votacao_geral_secao_2010.dta", clear
43  collapse (sum) QTDE_VOTOS, by(NUM_VOTAVEL)
44  save "Z:\Equipe\Erica\MONO- Dados\STATA\dados\QTDE_VOTOS_2010_BRASIL.dta", replace
45
46
47  *---- Calcula por UF ----*
48
49  * Calcula o total de votos por UF
50  use "Z:\Equipe\Erica\MONO- Dados\STATA\dados\votacao_geral_secao_2010.dta", clear
51  collapse (sum) QTDE_VOTOS, by(SIGLA_UF NUM_VOTAVEL)
52  reshape wide QTDE_VOTOS, i(SIGLA_UF) j(NUM_VOTAVEL)
53  save "Z:\Equipe\Erica\MONO- Dados\STATA\dados\QTDE_VOTOS_2010_UFS.dta", replace
54
55
56  *---- Calcula os dígitos ----*
57
58  use "Z:\Equipe\Erica\MONO- Dados\STATA\dados\votacao_geral_secao_2010.dta", clear
59  collapse (sum) QTDE_VOTOS, by(SIGLA_UF CODIGO_MUNICIPIO NUM_ZONA NUM_SECAO NUM_VOTAVEL)
60
61  * Muda o formato do arquivo (tabela dinamica)
62  gen SECAO = CODIGO_MUNICIPIO*1000000 + NUM_ZONA*10000 + NUM_SECAO
63  format SECAO %20.0f
64  reshape wide QTDE_VOTOS, i(SECAO) j(NUM_VOTAVEL)
65
66  * Calcula os dígitos
67  foreach sigla in 13 {
68      gen QTDE_VOTOS`sigla'_str = string(QTDE_VOTOS`sigla') if QTDE_VOTOS`sigla'!=0 & ~missing(QTDE_VOTOS`sigla')
69      gen PRIMEIRO_DIGITO`sigla' = substr(QTDE_VOTOS`sigla'_str, 1, 1) if length(QTDE_VOTOS`sigla'_str) > 0
70      gen SEGUNDO_DIGITO`sigla' = substr(QTDE_VOTOS`sigla'_str, 2, 1) if length(QTDE_VOTOS`sigla'_str) > 1
71      gen DOIS_PRIMEIROS_DIGITOS`sigla' = substr(QTDE_VOTOS`sigla'_str, 1, 2) if length(QTDE_VOTOS`sigla'_str) > 1
72  }
73
74  keep QTDE_VOTOS13 QTDE_VOTOS13_str PRIMEIRO_DIGITO13 SEGUNDO_DIGITO13 DOIS_PRIMEIROS_DIGITOS13
75  save "Z:\Equipe\Erica\MONO- Dados\STATA\dados\votacao_digitos_2010.dta", replace

```

## Junção dos dados

Após realizar o passo anterior para todos os anos, deve-se concatenar os dados destes anos em apenas um arquivo

Figura 7 – Stata

```

2  /*---- Monta o arquivo de dados ----*/
3
4  * Local dos arquivos
5  global dados "Z:\Equipe\Erica\MONO- Dados\STATA\dados"
6  global resultados "Z:\Equipe\Erica\MONO- Dados\STATA\documentos"
7
8  * Junta os arquivos
9
10 use "Z:\Equipe\Erica\MONO- Dados\STATA\dados\votacao_digitos_2002.dta", clear
11 append using "Z:\Equipe\Erica\MONO- Dados\STATA\dados\votacao_digitos_2006.dta"
12 append using "Z:\Equipe\Erica\MONO- Dados\STATA\dados\votacao_digitos_2010.dta"
13 save "\MONO- Dados\STATA\dados\votacao_digitos_append.dta", replace
14

```

Encontrando a distribuição Histórica Brasileira das eleições presidenciais segundo turno para os anos de 2002 a 2010.

Figura 8 – Stata

```
16 /*---- Calcula as frequencias históricas do Brasil ----*/
17
18 use "Z:\Equipe\Erica\MONO- Dados\STATA\dados\votacao_digitos_append.dta", clear
19 contract PRIMEIRO_DIGITO13, percent(FREQ_PRIMEIRO_DIGITO)
20 replace FREQ_PRIMEIRO_DIGITO = FREQ_PRIMEIRO_DIGITO/100
21 keep FREQ_PRIMEIRO_DIGITO
22 export excel "Z:\Equipe\Erica\MONO- Dados\Excels\BRASIL_HISTORICO.xlsx", cell(L3) sheetmodify
23
24 use "Z:\Equipe\Erica\MONO- Dados\STATA\dados\votacao_digitos_append.dta" if ~missing(SEGUNDO_DIGITO13), clear
25 contract SEGUNDO_DIGITO13, percent(FREQ_SEGUNDO_DIGITO)
26 replace FREQ_SEGUNDO_DIGITO = FREQ_SEGUNDO_DIGITO/100
27 keep FREQ_SEGUNDO_DIGITO
28 export excel "Z:\Equipe\Erica\MONO- Dados\Excels\BRASIL_HISTORICO.xlsx", cell(X3) sheetmodify
29
30 use "Z:\Equipe\Erica\MONO- Dados\STATA\dados\votacao_digitos_append.dta" if ~missing(DOIS_PRIMEIROS_DIGITOS13), clear
31 contract DOIS_PRIMEIROS_DIGITOS13, percent(FREQ_DOIS_PRIMEIROS_DIGITOS)
32 replace FREQ_DOIS_PRIMEIROS_DIGITOS = FREQ_DOIS_PRIMEIROS_DIGITOS/100
33 keep FREQ_DOIS_PRIMEIROS_DIGITOS
34 export excel "Z:\Equipe\Erica\MONO- Dados\Excels\BRASIL_HISTORICO.xlsx", cell(AJ3) sheetmodify
```



# APÊNDICE C – Comparação entre as distribuições

As tabelas a seguir apresentam as frequências dos dígitos da distribuição de Benford e da distribuição brasileira em relação aos testes do primeiro dígito, do segundo dígito e dos dois primeiros dígitos.

## C.1 Primeiro Dígito

Tabela 11 – Frequências dos Primeiros Dígitos da distribuição de Benford e da distribuição Histórica Brasileira - Fonte: elaboração própria

Díg.	Benford	Brasil
<b>1</b>	0,301	0,5501
<b>2</b>	0,176	0,2237
<b>3</b>	0,125	0,0341
<b>4</b>	0,097	0,0154
<b>5</b>	0,079	0,0212
<b>6</b>	0,067	0,0280
<b>7</b>	0,058	0,0353
<b>8</b>	0,051	0,0429
<b>9</b>	0,046	0,0493

## C.2 Segundo Dígito

Tabela 12 – Frequências dos Segundos Dígitos da distribuição de Benford e da distribuição Histórica Brasileira - Fonte: elaboração própria

Díg.	Benford	Brasil
<b>0</b>	0,120	0,1197
<b>1</b>	0,114	0,1190
<b>2</b>	0,109	0,1157
<b>3</b>	0,104	0,1107
<b>4</b>	0,100	0,1047
<b>5</b>	0,097	0,0983
<b>6</b>	0,093	0,0913
<b>7</b>	0,090	0,0854
<b>8</b>	0,088	0,0803
<b>9</b>	0,085	0,0749



## C.3 Dois Primeiros Dígitos

Tabela 13 – Frequências dos Dois Primeiros Dígitos da distribuição de Benford e da distribuição Histórica Brasileira - Fonte: elaboração própria

Díg.	Benford	Brasil	Díg.	Benford	Brasil
<b>10</b>	0,0414	0,0554	<b>55</b>	0,0078	0,0021
<b>11</b>	0,0378	0,0590	<b>56</b>	0,0077	0,0022
<b>12</b>	0,0348	0,0609	<b>57</b>	0,0076	0,0022
<b>13</b>	0,0322	0,0607	<b>58</b>	0,0074	0,0024
<b>14</b>	0,0300	0,0594	<b>59</b>	0,0073	0,0024
<b>15</b>	0,0280	0,0570	<b>60</b>	0,0072	0,0024
<b>16</b>	0,0263	0,0543	<b>61</b>	0,0071	0,0026
<b>17</b>	0,0248	0,0515	<b>62</b>	0,0069	0,0026
<b>18</b>	0,0235	0,0480	<b>63</b>	0,0068	0,0026
<b>19</b>	0,0223	0,0443	<b>64</b>	0,0067	0,0027
<b>20</b>	0,0212	0,0403	<b>65</b>	0,0066	0,0028
<b>21</b>	0,0202	0,0362	<b>66</b>	0,0065	0,0028
<b>22</b>	0,0193	0,0317	<b>67</b>	0,0064	0,0030
<b>23</b>	0,0185	0,0273	<b>68</b>	0,0063	0,0031
<b>24</b>	0,0177	0,0233	<b>69</b>	0,0062	0,0032
<b>25</b>	0,0170	0,0192	<b>70</b>	0,0062	0,0031
<b>26</b>	0,0164	0,0151	<b>71</b>	0,0061	0,0034
<b>27</b>	0,0158	0,0123	<b>72</b>	0,0060	0,0033
<b>28</b>	0,0152	0,0101	<b>73</b>	0,0059	0,0034
<b>29</b>	0,0147	0,0082	<b>74</b>	0,0058	0,0034
<b>30</b>	0,0142	0,0067	<b>75</b>	0,0058	0,0035
<b>31</b>	0,0138	0,0057	<b>76</b>	0,0057	0,0037
<b>32</b>	0,0134	0,0048	<b>77</b>	0,0056	0,0037
<b>33</b>	0,0130	0,0040	<b>78</b>	0,0055	0,0038
<b>34</b>	0,0126	0,0033	<b>79</b>	0,0055	0,0039
<b>35</b>	0,0122	0,0027	<b>80</b>	0,0054	0,0039
<b>36</b>	0,0119	0,0022	<b>81</b>	0,0053	0,0040
<b>37</b>	0,0116	0,0017	<b>82</b>	0,0053	0,0040
<b>38</b>	0,0113	0,0015	<b>83</b>	0,0052	0,0043
<b>39</b>	0,0110	0,0014	<b>84</b>	0,0051	0,0043
<b>40</b>	0,0107	0,0014	<b>85</b>	0,0051	0,0044
<b>41</b>	0,0105	0,0014	<b>86</b>	0,0050	0,0043
<b>42</b>	0,0102	0,0015	<b>87</b>	0,0050	0,0044
<b>43</b>	0,0100	0,0015	<b>88</b>	0,0049	0,0046
<b>44</b>	0,0098	0,0015	<b>89</b>	0,0049	0,0046
<b>45</b>	0,0095	0,0015	<b>90</b>	0,0048	0,0046
<b>46</b>	0,0093	0,0016	<b>91</b>	0,0047	0,0047
<b>47</b>	0,0091	0,0016	<b>92</b>	0,0047	0,0048
<b>48</b>	0,0090	0,0017	<b>93</b>	0,0046	0,0048
<b>49</b>	0,0088	0,0017	<b>94</b>	0,0046	0,0049
<b>50</b>	0,0086	0,0018	<b>95</b>	0,0045	0,0049
<b>51</b>	0,0084	0,0019	<b>96</b>	0,0045	0,0050
<b>52</b>	0,0083	0,0020	<b>97</b>	0,0045	0,0052
<b>53</b>	0,0081	0,0020	<b>98</b>	0,0044	0,0051
<b>54</b>	0,0080	0,0021	<b>99</b>	0,0044	0,0052





# APÊNDICE D – Tabelas Demonstrativas dos Resultados dos Testes para Distribuição de Benford

Devido a quantidade de tabelas e gráficos por estado, o estado <sup>1</sup> do Acre (AC) foi escolhido como estado representativo, visto que possui um dos menores percentuais de dígitos brancos apenas 0,91% e 2,15% de nulos. A escolha do Acre se deve a indicação de Mebane (2009), em que Mebane afirma que as regiões em que o presidente iraniano eleito, Mahmoud Ahmadinejad, obteve votações expressivas tinham um percentual baixo de votos inválidos. Nessas regiões, os dados referentes aos votos de Mahmoud Ahmadinejad divergiam significativamente da distribuição de Benford.

## D.1 Tabelas Acre para o código votável 13

### D.1.1 Primeiro Dígito

Tabela 14 – Teste do Primeiro Dígito conforme Benford para o estado do Acre com código votável 13 - Fonte: elaboração própria

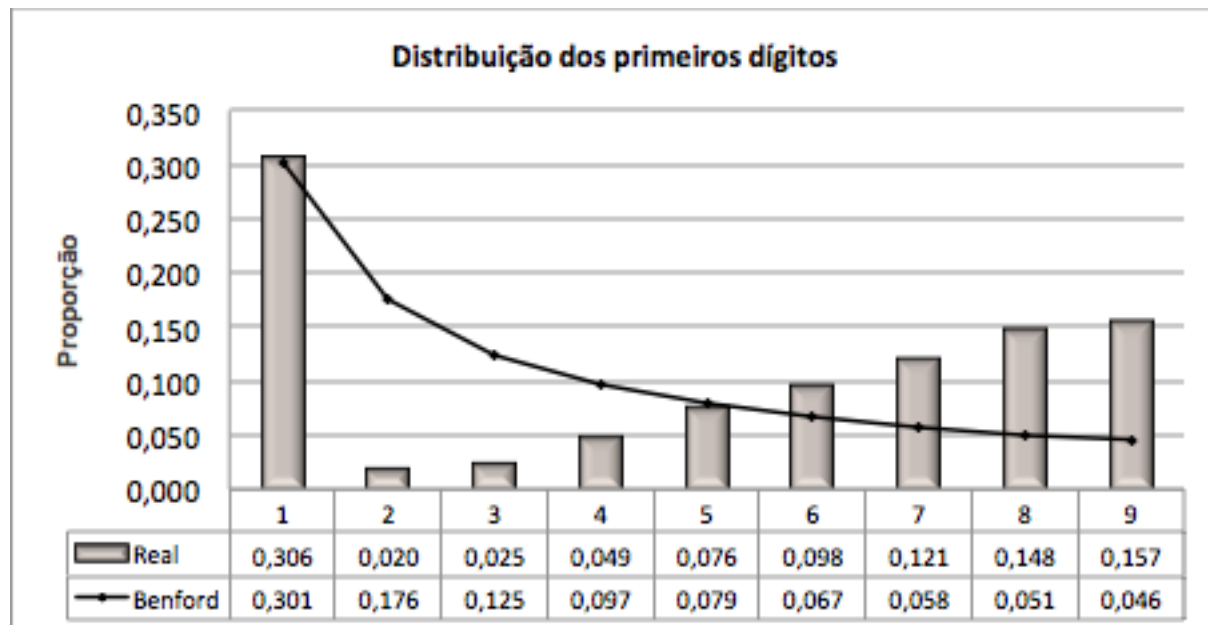
Dígito	Cont.	Real	Benford	Diferença	Teste Z	CS	MAD
1	501	0,306	0,301	0,005	0,418	0,139	0,005
2	32	0,020	0,176	-0,156	16,564	227,666	0,156
3	41	0,025	0,125	-0,100	12,172	130,840	0,100
4	80	0,049	0,097	-0,048	6,528	39,094	0,048
5	125	0,076	0,079	-0,003	0,350	0,145	0,003
6	160	0,098	0,067	0,031	4,917	23,087	0,031
7	198	0,121	0,058	0,063	10,827	111,854	0,063
8	243	0,148	0,051	0,097	17,837	304,771	0,097
9	257	0,157	0,046	0,111	21,346	438,423	0,111
N	<b>1637</b>	<b>1,0</b>			<b>CS Real</b>	<b>1276,019</b>	<b>0,614</b>
					<b>CS Limite</b>	<b>15,507</b>	<b>0,0683</b>

Na Tabela 13, a primeira coluna se refere aos primeiros dígitos dos votos; a segunda corresponde às frequências absolutas desses dígitos no banco de dados; a terceira se refere às frequências da 2ª coluna em termos relativos; a quarta coluna corresponde às frequências relativas padrão da Lei de Benford; a quinta é a diferença entre as frequências

<sup>1</sup> Em consequência da quantidade de gráficos e tabelas resultantes, há apenas a demonstração de todos os resultados para um estado. Desta forma, para visualizar outros resultados faz-se necessário acessar uma página no Google Drive. Disponível em: <<https://drive.google.com/drive/folders/0B5N1NvqnAD-xT2hRZlk5TkxzRnM>>

da 3ª e da 4ª colunas; a sexta traz os resultados do Teste Z; a sétima apresenta os resultados do Teste Qui-Quadrado; e a oitava mostra os valores do MDA.

Figura 9 – Distribuição dos Primeiros Dígitos- Elaboração Própria



### D.1.2 Segundo Dígito

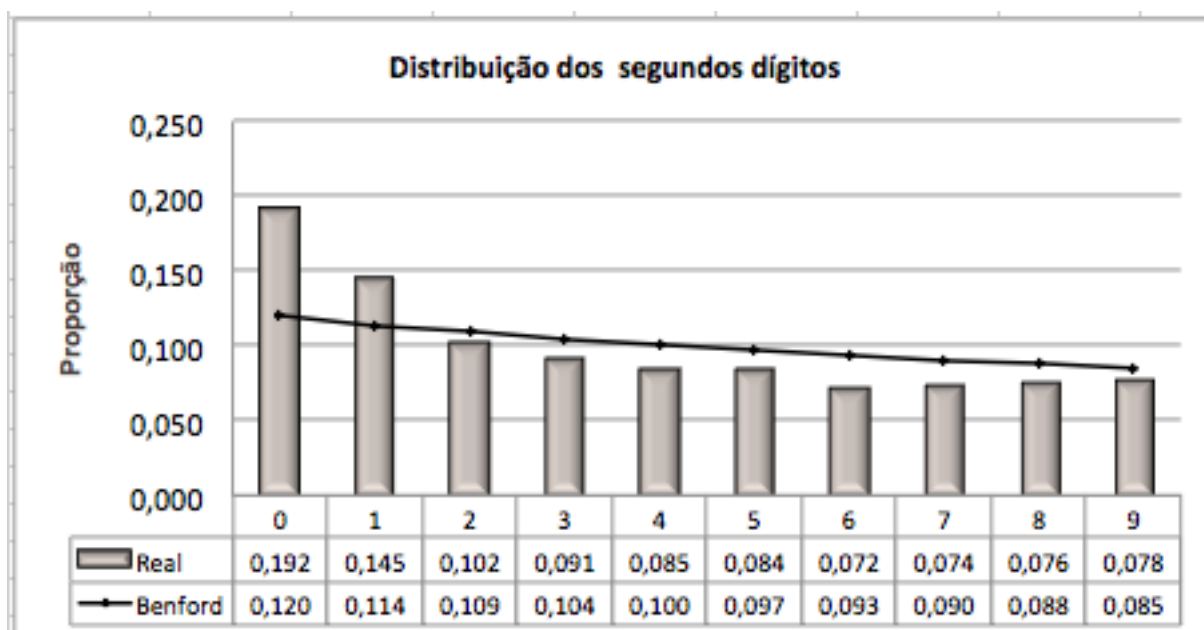
Tabela 15 – Teste do Segundo Dígito conforme Benford para o estado do Acre com código votável 13 - Fonte: elaboração própria

Dígito	Cont.	Real	Benford	Diferença	Teste Z	CS	MAD
0	314	0,192	0,120	0,072	8,963	71,292	0,072
1	237	0,145	0,114	0,031	3,930	13,954	0,031
2	167	0,102	0,109	-0,007	0,825	0,666	0,007
3	149	0,091	0,104	-0,013	1,640	2,531	0,013
4	138	0,085	0,100	-0,015	2,038	3,891	0,015
5	137	0,084	0,097	-0,013	1,740	2,867	0,013
6	118	0,072	0,093	-0,021	2,836	7,516	0,021
7	121	0,074	0,090	-0,016	2,195	4,560	0,016
8	124	0,076	0,088	-0,012	1,670	2,679	0,012
9	127	0,078	0,085	-0,007	0,996	0,990	0,007
<b>N</b>	<b>1632</b>	<b>1,0</b>			<b>CS Real</b>	<b>110,948</b>	<b>0,207</b>
					<b>CS Limite</b>	<b>16,919</b>	<b>0,02072</b>

Na Tabela 14, a primeira coluna se refere aos segundos dígitos dos votos; a segunda corresponde às frequências absolutas desses dígitos no banco de dados; a terceira se refere às frequências da 2ª coluna em termos relativos; a quarta coluna corresponde às frequências relativas padrão da Lei de Benford; a quinta é a diferença entre as frequências da 3ª e

da 4ª coluna; a sexta traz os resultados do Teste Z; a sétima apresenta os resultados do Teste Qui-Quadrado; e a oitava mostra os valores do MDA.

Figura 10 – Distribuição dos Segundos Dígitos - Elaboração Própria





## D.1.3 Teste dos dois Primeiros Dígitos

Tabela 16 – Teste dos Dois Primeiros Dígitos conforme Benford para o estado do Acre com código votável 13 - Fonte: elaboração própria

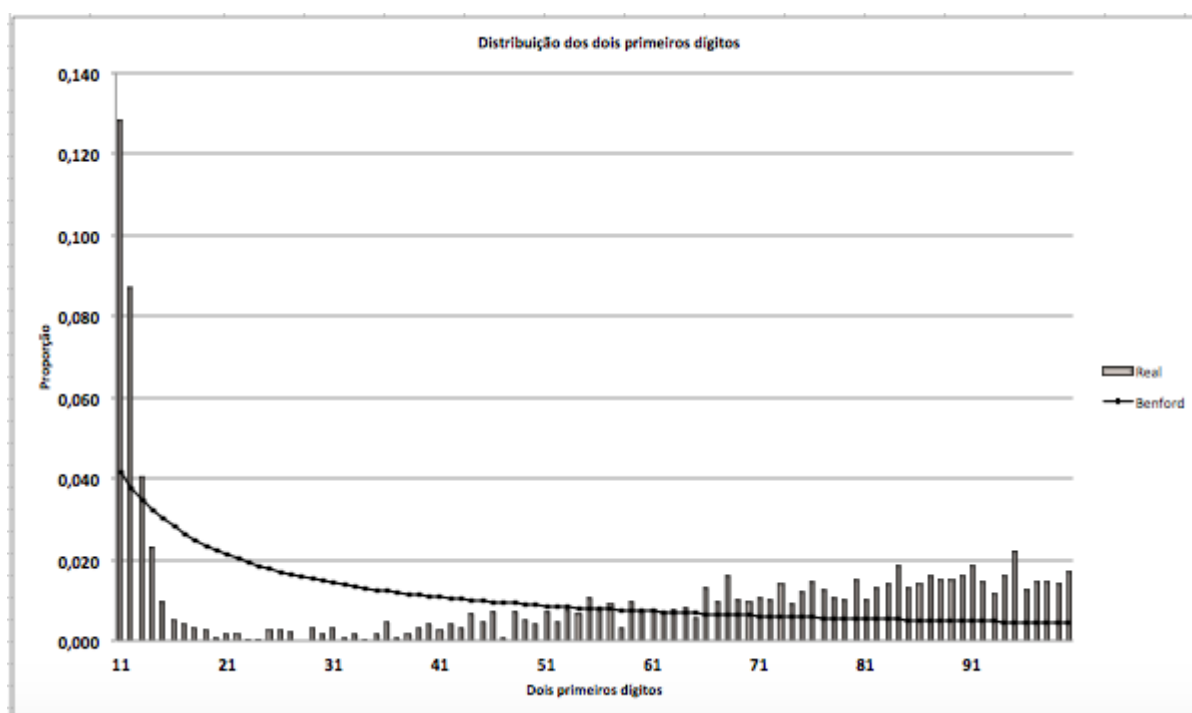
Díg.	Cont.	Real	LB	Dif.	Z	CS	MAD
10	209	0,128	0,041	0,087	12,433	296,172	0,087
11	142	0,087	0,038	0,049	7,337	104,632	0,049
12	66	0,040	0,035	0,006	0,797	1,514	0,006
13	38	0,023	0,032	-0,009	1,353	4,017	0,009
14	16	0,010	0,030	-0,020	3,301	22,135	0,020
15	9	0,006	0,028	-0,023	3,820	29,514	0,023
16	7	0,004	0,026	-0,022	3,852	30,109	0,022
17	6	0,004	0,025	-0,021	3,799	29,401	0,021
18	5	0,003	0,023	-0,020	3,765	28,974	0,020
19	2	0,001	0,022	-0,021	3,987	32,465	0,021
20	3	0,002	0,021	-0,019	3,866	28,841	0,019
21	3	0,002	0,020	-0,018	3,633	27,245	0,018
22	1	0,001	0,019	-0,019	3,783	29,538	0,019
23	1	0,001	0,018	-0,018	3,689	28,198	0,018
24	5	0,003	0,018	-0,015	3,066	19,797	0,015
25	5	0,003	0,017	-0,014	2,972	18,698	0,014
26	4	0,002	0,016	-0,014	3,022	19,347	0,014
27	0	0,000	0,016	-0,016	3,506	25,776	0,016
28	6	0,004	0,015	-0,012	2,574	14,319	0,012
29	3	0,002	0,015	-0,013	2,934	18,403	0,013
30	6	0,004	0,014	-0,011	2,419	12,789	0,011
31	2	0,001	0,014	-0,013	2,951	18,680	0,013
32	3	0,002	0,013	-0,012	2,737	16,223	0,012
33	1	0,001	0,013	-0,012	2,988	19,206	0,012
34	3	0,002	0,013	-0,011	2,619	14,984	0,011
35	8	0,005	0,012	-0,007	1,761	7,172	0,007
36	2	0,001	0,012	-0,011	2,672	15,626	0,011
37	3	0,002	0,012	-0,010	2,620	13,378	0,010
38	6	0,004	0,011	-0,008	2,072	8,366	0,008
39	7	0,004	0,011	-0,007	1,684	6,675	0,007
40	5	0,003	0,011	-0,008	1,971	8,930	0,008
41	7	0,004	0,010	-0,006	1,576	5,949	0,006
42	6	0,004	0,010	-0,007	1,699	6,836	0,007
43	11	0,007	0,010	-0,003	0,764	1,720	0,003
44	8	0,005	0,010	-0,005	1,422	3,946	0,005
45	12	0,007	0,010	-0,002	0,470	0,822	0,002
46	2	0,001	0,009	-0,008	2,247	11,505	0,008
47	12	0,007	0,009	-0,002	0,359	0,572	0,002
48	9	0,006	0,009	-0,003	0,866	2,157	0,003
49	7	0,004	0,009	-0,004	1,197	3,741	0,004
50	12	0,007	0,009	-0,001	0,200	0,295	0,001
51	8	0,005	0,008	-0,004	0,921	2,413	0,004
52	13	0,008	0,008	0,000	-0,094	0,019	0,000
53	11	0,007	0,008	-0,001	0,248	0,382	0,001
54	18	0,011	0,008	0,003	0,795	1,918	0,003

Tabela 17 – Continuação- Teste dos Dois Primeiros Dígitos conforme Benford para o estado do Acre com código votável 13 - Fonte: elaboração própria

Díg.	Cont.	Real	LB	Dif.	Z	CS	MAD
55	13	0,008	0,008	0,000	-0,151	0,004	0,000
56	15	0,009	0,008	0,002	0,297	0,480	0,002
57	6	0,004	0,008	-0,004	1,088	3,247	0,004
58	16	0,010	0,007	0,002	0,595	1,245	0,002
59	12	0,007	0,007	0,000	-0,186	0,001	0,000
60	13	0,008	0,007	0,001	0,062	0,141	0,001
61	12	0,007	0,007	0,000	0,100	0,020	0,000
62	13	0,008	0,007	0,001	0,143	0,243	0,001
63	14	0,009	0,007	0,002	0,396	0,722	0,002
64	10	0,006	0,007	-0,001	0,213	0,089	0,001
65	22	0,013	0,007	0,007	2,215	11,549	0,007
66	16	0,010	0,007	0,003	1,169	2,677	0,003
67	27	0,017	0,006	0,010	3,421	25,926	0,010
68	17	0,010	0,006	0,004	1,259	4,278	0,004
69	16	0,010	0,006	0,004	1,078	3,301	0,004
70	18	0,011	0,006	0,005	1,569	6,281	0,005
71	17	0,010	0,006	0,004	1,385	5,067	0,004
72	23	0,014	0,006	0,008	3,022	17,887	0,008
73	15	0,009	0,006	0,003	1,006	2,976	0,003
74	20	0,012	0,006	0,006	2,200	11,558	0,006
75	24	0,015	0,006	0,009	3,177	22,744	0,009
76	21	0,013	0,006	0,007	2,523	14,863	0,007
77	18	0,011	0,006	0,005	1,859	8,573	0,005
78	17	0,010	0,006	0,005	1,660	7,037	0,005
79	25	0,015	0,005	0,010	3,848	29,019	0,010
80	17	0,010	0,005	0,005	1,735	7,628	0,005
81	22	0,013	0,005	0,008	2,983	20,350	0,008
82	23	0,014	0,005	0,009	3,271	24,166	0,009
83	31	0,019	0,005	0,014	5,276	59,702	0,014
84	22	0,013	0,005	0,008	3,356	22,090	0,008
85	23	0,014	0,005	0,009	3,404	26,103	0,009
86	27	0,017	0,005	0,012	4,446	43,162	0,012
87	25	0,015	0,005	0,010	3,993	35,258	0,010
88	25	0,015	0,005	0,010	4,287	36,048	0,010
89	27	0,017	0,005	0,012	4,591	45,973	0,012
90	31	0,019	0,005	0,014	5,660	68,537	0,014
91	24	0,015	0,005	0,010	3,917	34,105	0,010
92	19	0,012	0,005	0,007	2,670	16,775	0,007
93	27	0,017	0,005	0,012	5,036	49,748	0,012
94	36	0,022	0,005	0,017	7,173	108,294	0,017
95	21	0,013	0,005	0,008	3,558	24,842	0,008
96	24	0,015	0,005	0,010	4,128	37,767	0,010
97	24	0,015	0,004	0,010	4,169	38,505	0,010
98	23	0,014	0,004	0,010	3,944	34,712	0,010
99	28	0,017	0,004	0,013	5,320	61,184	0,013
<b>N</b>	<b>1632</b>			<b>CS</b>	<b>Real</b>	<b>1968,27</b>	<b>0,905</b>
				<b>CS</b>	<b>Limite</b>	<b>112,02</b>	<b>0,010058</b>

Na tabela 15, “Díg.” se refere aos dois primeiros dígitos dos valores; “Cont.” são as frequências absolutas com que os dígitos se repetem na planilha; “Real” são as frequências relativas com que os dígitos aparecem na planilha; “LB” são as frequências padrão da Lei de Benford; “Dif.” é a diferença entre “Real” e “LB”; “Z” se refere aos resultados do Teste Z para mensurar a conformidade com a Lei de Benford; “CS” mostra os resultados do Teste Qui-Quadrado; e “MAD” traz o resultado da Média dos Desvios Absolutos.

Figura 11 – Distribuição dos dois Primeiros Dígitos - Elaboração Própria

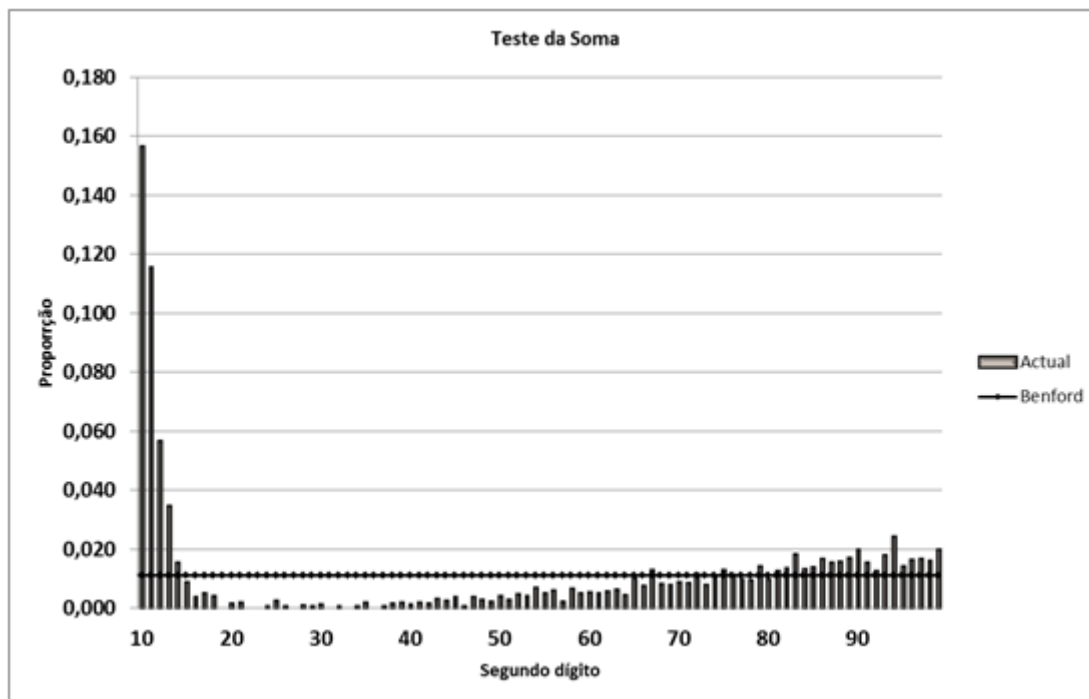


## D.2 Teste da Soma

O teste da soma é utilizado para identificar dígitos cujo somatório difere consideravelmente de  $\frac{1}{90}$  ou 1,1% do valor total da amostra, isto é, aplicado para identificar os números que ocorrem com maior frequência. Nas Tabelas a seguir, a 1ª e 6ª colunas se referem aos dois primeiros dígitos da quantidade de votos; a 2ª e 7ª colunas correspondem à soma dos itens que possuem os dois primeiros dígitos apontados na 1ª e 6ª colunas; a 3ª e 8ª colunas mostram as proporções das somas calculadas na 2ª e 7ª colunas em relação ao somatório de todos os votos da planilha deste estado; a 4ª e 9ª colunas apresentam as frequências padrão da Lei de Benford; e a 5ª e 10ª colunas trazem a diferença entre as proporções das somas e as frequências da Lei de Benford.

## D.2.1 Código Votável 13

Figura 12 – Teste da Soma para código votável 13 - Elaboração Própria



Segundo a distribuição de Benford, cada dígito dos votos eleitorais, que são em seu total 138.962, deveria somar aproximadamente 2.667,38. Nos dígitos onde houve picos constata-se, para esse teste, que os dados não se aproximaram de forma satisfatória de uma sequência de Benford.





### D.3 Tabelas Acre para o código votável 45

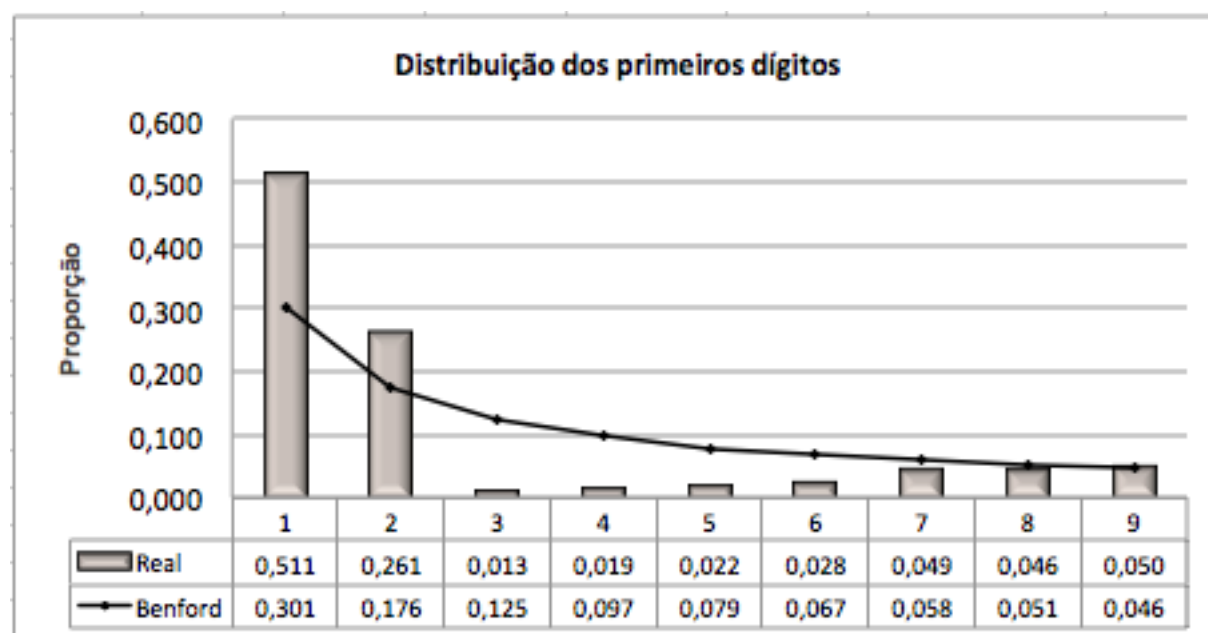
#### D.3.1 Primeiro Dígito

Tabela 19 – Teste do Primeiro Dígito conforme Benford para o estado do Acre com código votável 45 - Fonte: elaboração própria

Dígito	Cont.	Real	Benford	Diferença	Teste Z	CS	MAD
1	835	0,511	0,301	0,210	18,469	240,003	0,210
2	427	0,261	0,176	0,085	9,022	67,799	0,085
3	21	0,013	0,125	-0,112	13,640	164,285	0,112
4	31	0,019	0,097	-0,078	10,591	102,468	0,078
5	36	0,022	0,079	-0,057	8,471	67,053	0,057
6	46	0,028	0,067	-0,039	6,215	36,751	0,039
7	80	0,049	0,058	-0,009	1,502	2,286	0,009
8	75	0,046	0,051	-0,005	0,874	0,824	0,005
9	82	0,050	0,046	0,004	0,752	0,631	0,004
N	<b>1633</b>	<b>1,0</b>			<b>CS Real</b>	<b>682,0996</b>	<b>0,600</b>
					<b>CS Limite</b>	<b>15,507</b>	<b>0,06667</b>

Tabela 16, a primeira coluna se refere aos primeiros dígitos dos votos; a segunda corresponde às frequências absolutas desses dígitos no banco de dados; a terceira se refere às frequências da 2a coluna em termos relativos; a quarta coluna corresponde às frequências relativas padrão da Lei de Benford; a quinta é a diferença entre as frequências da 3a e da 4a colunas; a sexta traz os resultados do Teste Z; a sétima apresenta os resultados do Teste Qui-Quadrado; e a oitava mostra os valores do MDA

Figura 13 – Distribuição dos Primeiros Dígitos - Elaboração Própria



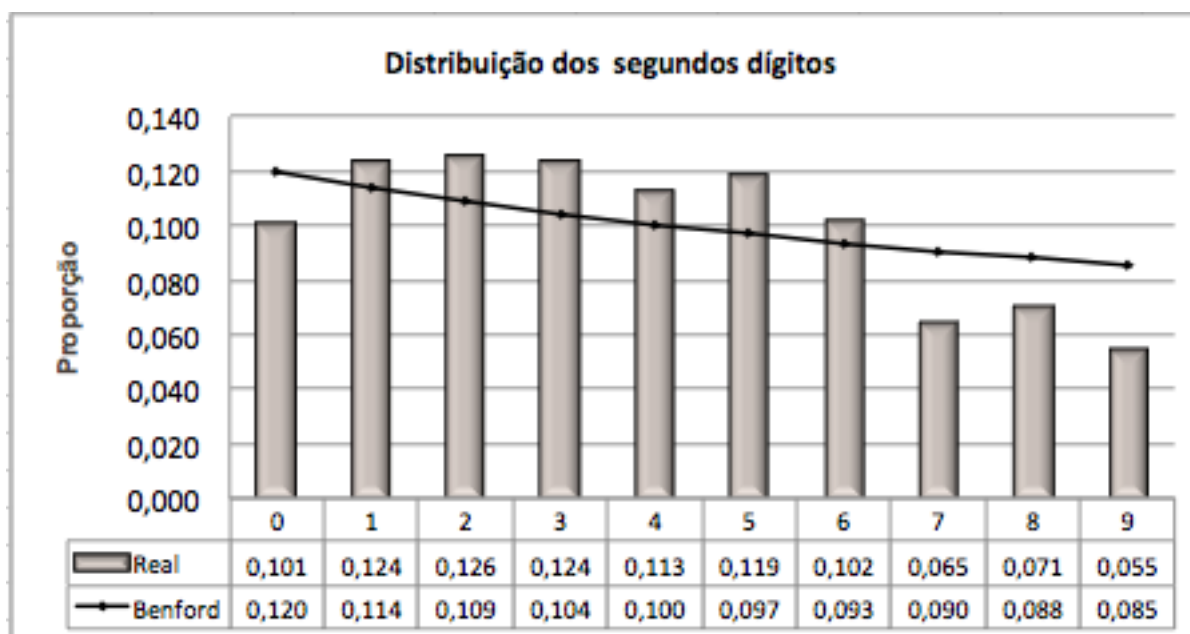
## D.3.2 Segundo Dígito

Tabela 20 – Teste do Segundo Dígito conforme Benford para o estado do Acre com código votável 45 - Fonte: elaboração própria

Dígito	Cont.	Real	Benford	Diferença	Teste Z	CS	MAD
0	165	0,101	0,120	-0,019	2,269	4,684	0,019
1	201	0,124	0,114	0,010	1,172	1,299	0,010
2	205	0,126	0,109	0,017	2,160	4,313	0,017
3	202	0,124	0,104	0,020	2,623	6,355	0,020
4	184	0,113	0,100	0,013	1,719	2,789	0,013
5	194	0,119	0,097	0,022	2,989	8,295	0,022
6	166	0,102	0,093	0,009	1,211	1,426	0,009
7	105	0,065	0,090	-0,025	3,546	11,722	0,025
8	115	0,071	0,088	-0,017	2,422	5,545	0,017
9	90	0,055	0,085	-0,030	4,249	16,865	0,030
N	1627						
					CS Real	63,292	0,1821
					CS Limite	16,919	0,01821

Na Tabela 17, a primeira coluna se refere aos segundos dígitos dos votos; a segunda corresponde às frequências absolutas desses dígitos no banco de dados; a terceira se refere às frequências da 2a coluna em termos relativos; a quarta coluna corresponde às frequências relativas padrão da Lei de Benford; a quinta é a diferença entre as frequências da 3a e da 4a colunas; a sexta traz os resultados do Teste Z; a sétima apresenta os resultados do Teste Qui-Quadrado; e a oitava mostra os valores do MDA.

Figura 14 – Distribuição dos Segundos Dígitos - Elaboração Própria





## D.3.3 Dois Primeiros Dígitos

Tabela 21 – Teste dos Dois Primeiros Dígitos conforme Benford para o estado do Acre com código votável 45 - Fonte: elaboração própria

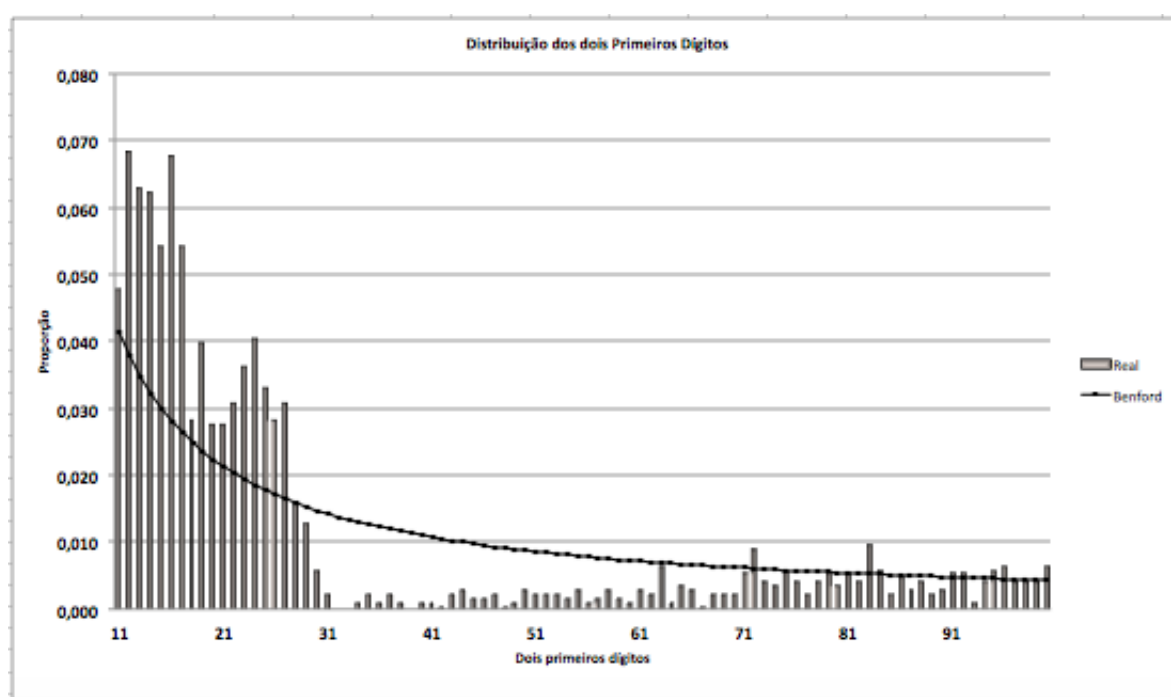
Díg.	Cont.	Real	Benford	Diferença	Teste Z	CS	MAD
10	78	0,048	0,041	0,007	0,859	1,685	0,007
11	111	0,068	0,038	0,030	4,502	39,882	0,030
12	102	0,063	0,035	0,028	4,293	36,511	0,028
13	101	0,062	0,032	0,030	4,775	45,172	0,030
14	88	0,054	0,030	0,024	3,970	31,601	0,024
15	110	0,068	0,028	0,040	6,795	90,938	0,040
16	88	0,054	0,026	0,028	4,880	47,615	0,028
17	46	0,028	0,025	0,003	0,526	0,780	0,003
18	65	0,040	0,023	0,016	3,015	18,795	0,016
19	45	0,028	0,022	0,005	0,932	2,115	0,005
20	45	0,028	0,021	0,006	1,293	3,213	0,006
21	50	0,031	0,020	0,011	2,030	8,926	0,011
22	59	0,036	0,019	0,017	3,420	24,236	0,017
23	66	0,041	0,018	0,022	4,589	42,922	0,022
24	54	0,033	0,018	0,015	3,240	21,938	0,015
25	46	0,028	0,017	0,011	2,365	12,067	0,011
26	50	0,031	0,016	0,014	3,113	20,415	0,014
27	26	0,016	0,016	0,000	-0,096	0,004	0,000
28	21	0,013	0,015	-0,002	0,406	0,581	0,002
29	10	0,006	0,015	-0,009	1,905	8,129	0,009
30	4	0,002	0,014	-0,012	2,715	15,860	0,012
31	0	0,000	0,014	-0,014	3,253	22,434	0,014
32	0	0,000	0,013	-0,013	3,198	21,743	0,013
33	2	0,001	0,013	-0,012	2,832	17,284	0,012
34	4	0,002	0,013	-0,010	2,459	13,264	0,010
35	2	0,001	0,012	-0,011	2,723	16,106	0,011
36	4	0,002	0,012	-0,009	2,345	12,186	0,009
37	2	0,001	0,012	-0,010	2,784	15,056	0,010
38	0	0,000	0,011	-0,011	3,074	18,354	0,011
39	2	0,001	0,011	-0,010	2,528	14,113	0,010
40	2	0,001	0,011	-0,009	2,484	13,677	0,009
41	1	0,001	0,010	-0,010	2,615	15,086	0,010
42	4	0,002	0,010	-0,008	2,048	9,589	0,008
43	5	0,003	0,010	-0,007	1,825	7,783	0,007
44	3	0,002	0,010	-0,008	2,317	10,446	0,008
45	3	0,002	0,010	-0,008	2,100	10,110	0,008
46	4	0,002	0,009	-0,007	1,878	8,249	0,007
47	1	0,001	0,009	-0,009	2,396	12,943	0,009
48	2	0,001	0,009	-0,008	2,175	10,844	0,008
49	5	0,003	0,009	-0,006	1,573	6,026	0,006
50	4	0,002	0,009	-0,006	1,726	7,136	0,006
51	4	0,002	0,008	-0,006	1,690	6,887	0,006
52	4	0,002	0,008	-0,006	1,655	6,648	0,006
53	3	0,002	0,008	-0,006	1,818	7,889	0,006
54	5	0,003	0,008	-0,005	1,389	4,894	0,005

Tabela 22 – Continuação- Teste dos Dois Primeiros Dígitos conforme Benford para o estado do Acre com código votável 45 - Fonte: elaboração própria

Díg.	Cont.	Real	Benford	Diferença	Teste Z	CS	MAD
55	2	0,001	0,008	-0,007	1,957	9,046	0,007
56	3	0,002	0,008	-0,006	1,726	7,226	0,006
57	5	0,003	0,008	-0,004	1,288	4,323	0,004
58	3	0,002	0,007	-0,006	1,668	6,824	0,006
59	2	0,001	0,007	-0,006	1,848	8,213	0,006
60	5	0,003	0,007	-0,004	1,194	3,820	0,004
61	4	0,002	0,007	-0,005	1,582	4,882	0,005
62	11	0,007	0,007	0,000	-0,144	0,008	0,000
63	2	0,001	0,007	-0,006	1,748	7,487	0,006
64	6	0,004	0,007	-0,003	1,072	2,241	0,003
65	5	0,003	0,007	-0,004	1,047	3,105	0,004
66	1	0,001	0,007	-0,006	2,113	8,720	0,006
67	4	0,002	0,006	-0,004	1,213	3,997	0,004
68	4	0,002	0,006	-0,004	1,188	3,867	0,004
69	4	0,002	0,006	-0,004	1,164	3,741	0,004
70	9	0,006	0,006	-0,001	0,009	0,104	0,001
71	15	0,009	0,006	0,003	0,941	2,650	0,003
72	7	0,004	0,006	-0,002	0,629	0,774	0,002
73	6	0,004	0,006	-0,002	0,607	1,358	0,002
74	10	0,006	0,006	0,000	-0,109	0,028	0,000
75	7	0,004	0,006	-0,001	0,322	0,595	0,001
76	4	0,002	0,006	-0,003	1,001	2,969	0,003
77	7	0,004	0,006	-0,001	0,269	0,492	0,001
78	9	0,006	0,006	0,000	-0,234	0,000	0,000
79	6	0,004	0,005	-0,002	0,693	0,938	0,002
80	9	0,006	0,005	0,000	-0,184	0,006	0,000
81	7	0,004	0,005	-0,001	0,167	0,322	0,001
82	16	0,010	0,005	0,005	1,577	6,454	0,005
83	10	0,006	0,005	0,001	0,137	0,279	0,001
84	4	0,002	0,005	-0,003	1,079	2,276	0,003
85	8	0,005	0,005	0,000	-0,179	0,008	0,000
86	5	0,003	0,005	-0,002	0,547	1,229	0,002
87	7	0,004	0,005	-0,001	0,023	0,143	0,001
88	4	0,002	0,005	-0,002	1,008	1,988	0,002
89	5	0,003	0,005	-0,002	0,487	1,062	0,002
90	9	0,006	0,005	0,001	0,054	0,182	0,001
91	9	0,006	0,005	0,001	0,076	0,211	0,001
92	2	0,001	0,005	-0,003	1,205	4,163	0,003
93	7	0,004	0,005	0,000	0,145	0,041	0,000
94	10	0,006	0,005	0,002	0,403	0,851	0,002
95	11	0,007	0,005	0,002	0,947	1,753	0,002
96	7	0,004	0,005	0,000	-0,174	0,014	0,000
97	7	0,004	0,004	0,000	-0,195	0,008	0,000
98	7	0,004	0,004	0,000	-0,216	0,004	0,000
99	11	0,007	0,004	0,002	0,782	2,140	0,002
<b>N</b>	<b>1627</b>				<b>CS Real</b>	<b>872,677</b>	<b>0,651</b>
					<b>CS Limite</b>	<b>112,02</b>	<b>0,00724</b>

Nas tabelas a seguir, “Díg.” se refere aos dois primeiros dígitos dos valores; “Cont.” são as frequências absolutas com que os dígitos se repetem na planilha; “Real” são as frequências relativas com que os dígitos aparecem na planilha; “LB” são as frequências padrão da Lei de Benford; “Dif.” é a diferença entre “Real” e “LB”; “Z” se refere aos resultados do Teste Z para mensurar a conformidade com a Lei de Benford; “CS” mostra os resultados do Teste Qui-Quadrado; e “MAD” traz o resultado da Média dos Desvios Absolutos.

Figura 15 – Distribuição dos dois Primeiros Dígitos - Elaboração Própria

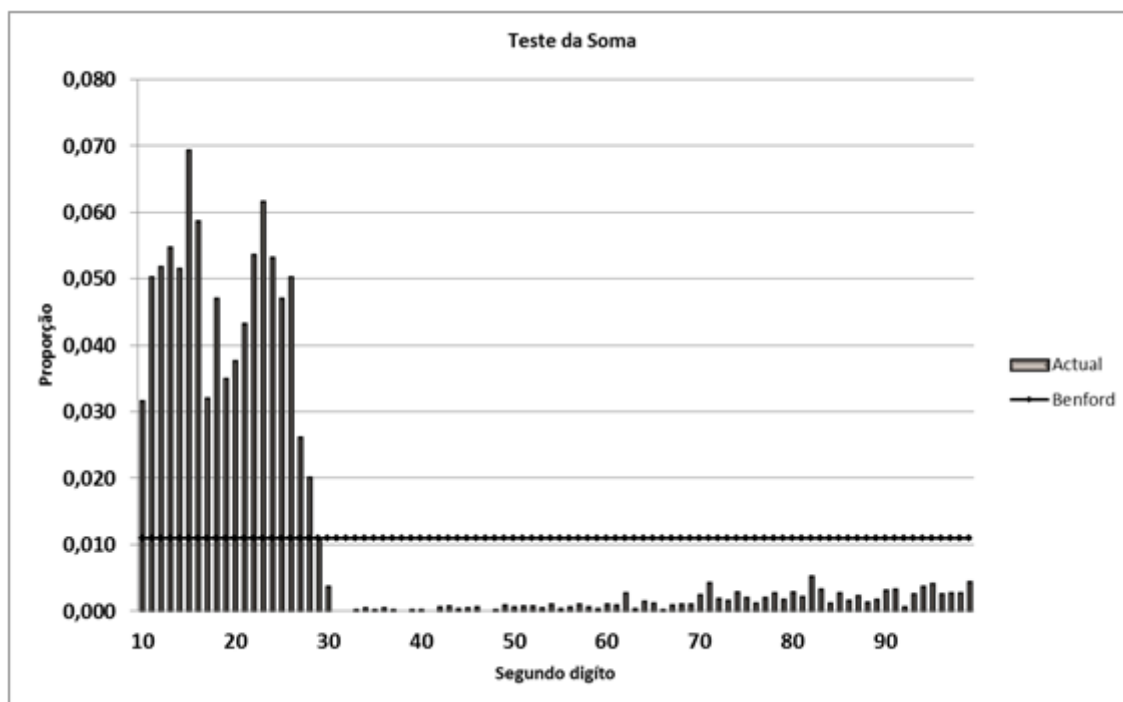


## D.4 Teste da Soma

O teste da soma é utilizado para identificar dígitos cujo somatório difere consideravelmente de  $\frac{1}{90}$  ou 1,1% do valor total da amostra, isto é, aplicado para identificar os números que ocorrem com maior frequência. Nas Tabelas a seguir, a 1ª e 6ª colunas se referem aos dois primeiros dígitos da quantidade de votos; a 2ª e 7ª colunas correspondem à soma dos itens que possuem os dois primeiros dígitos apontados na 1ª e 6ª colunas; a 3ª e 8ª colunas mostram as proporções das somas calculadas na 2ª e 7ª colunas em relação ao somatório de todos os votos da planilha deste estado; a 4ª e 9ª colunas apresentam as frequências padrão da Lei de Benford; e a 5ª e 10ª colunas trazem a diferença entre as proporções das somas e as frequências da Lei de Benford.

## D.4.1 Código Votável 45

Figura 16 – Teste da Soma para código votável 45 - Elaboração Própria



Segundo a distribuição de Benford, cada dígito dos votos eleitorais, que são em seu total 243.646 deveria somar aproximadamente 5.192,38. Nos dígitos onde houve picos constata-se, para esse teste, que os dados não se aproximaram de uma sequência de Benford.



Tabela 23 – Teste da Soma para código votável 45

Díg.	Soma	Real	LB	Dif.	Díg.	Soma	Real	LB	Dif.	
10	7.714	0,032	0,011	0,021	55	110	0,000	0,011	-0,011	
11	12.264	0,050	0,011	0,039	56	168	0,001	0,011	-0,010	
12	12.648	0,052	0,011	0,041	57	285	0,001	0,011	-0,010	
13	13.362	0,055	0,011	0,044	58	174	0,001	0,011	-0,010	
14	12.558	0,052	0,011	0,041	59	118	0,000	0,011	-0,011	
15	16.893	0,069	0,011	0,058	60	300	0,001	0,011	-0,010	
16	14.318	0,059	0,011	0,048	61	244	0,001	0,011	-0,010	
17	7.816	0,032	0,011	0,021	62	682	0,003	0,011	-0,008	
18	11.473	0,047	0,011	0,036	63	126	0,001	0,011	-0,010	
19	8.551	0,035	0,011	0,024	64	384	0,002	0,011	-0,009	
20	9.204	0,038	0,011	0,027	65	325	0,001	0,011	-0,010	
21	10.558	0,043	0,011	0,032	66	66	0,000	0,011	-0,011	
22	13.091	0,054	0,011	0,043	67	268	0,001	0,011	-0,010	
23	15.040	0,062	0,011	0,051	68	272	0,001	0,011	-0,010	
24	12.972	0,053	0,011	0,042	69	276	0,001	0,011	-0,010	
25	11.478	0,047	0,011	0,036	70	630	0,003	0,011	-0,008	
26	12.245	0,050	0,011	0,039	71	1.065	0,004	0,011	-0,007	
27	6.380	0,026	0,011	0,015	72	504	0,002	0,011	-0,009	
28	4.933	0,020	0,011	0,009	73	438	0,002	0,011	-0,009	
29	2.671	0,011	0,011	0,000	74	740	0,003	0,011	-0,008	
30	950	0,004	0,011	-0,007	75	525	0,002	0,011	-0,009	
31	0	0,000	0,011	-0,011	76	304	0,001	0,011	-0,010	
32	0	0,000	0,011	-0,011	77	539	0,002	0,011	-0,009	
33	66	0,000	0,011	-0,011	78	702	0,003	0,011	-0,008	
34	136	0,001	0,011	-0,010	79	474	0,002	0,011	-0,009	
35	70	0,000	0,011	-0,011	80	720	0,003	0,011	-0,008	
36	144	0,001	0,011	-0,010	81	567	0,002	0,011	-0,009	
37	74	0,000	0,011	-0,011	82	1.312	0,005	0,011	-0,006	
38	0	0,000	0,011	-0,011	83	830	0,003	0,011	-0,008	
39	78	0,000	0,011	-0,011	84	336	0,001	0,011	-0,010	
40	80	0,000	0,011	-0,011	85	680	0,003	0,011	-0,008	
41	41	0,000	0,011	-0,011	86	430	0,002	0,011	-0,009	
42	168	0,001	0,011	-0,010	87	609	0,002	0,011	-0,009	
43	215	0,001	0,011	-0,010	88	352	0,001	0,011	-0,010	
44	132	0,001	0,011	-0,010	89	445	0,002	0,011	-0,009	
45	135	0,001	0,011	-0,010	90	810	0,003	0,011	-0,008	
46	184	0,001	0,011	-0,010	91	819	0,003	0,011	-0,008	
47	47	0,000	0,011	-0,011	92	184	0,001	0,011	-0,010	
48	96	0,000	0,011	-0,011	93	651	0,003	0,011	-0,008	
49	245	0,001	0,011	-0,010	94	940	0,004	0,011	-0,007	
50	200	0,001	0,011	-0,010	95	1.045	0,004	0,011	-0,007	
51	204	0,001	0,011	-0,010	96	672	0,003	0,011	-0,008	
52	208	0,001	0,011	-0,010	97	679	0,003	0,011	-0,008	
53	159	0,001	0,011	-0,010	98	686	0,003	0,011	-0,008	
54	270	0,001	0,011	-0,010	99	1.089	0,004	0,011	-0,007	
<b>Soma</b>	<b>dos valores</b>	<b>unit</b>								<b>243.646</b>



# APÊNDICE E – Tabelas Demonstrativas dos Resultados dos Testes para Distribuição Histórica Brasileira

## E.1 Tabelas Acre para o código votável 13

Devido a grande quantidade de tabelas e gráficos, dado que existem as tabelas referentes aos testes apresentados no capítulo 2 para cada uma das unidades federativas, optou-se por escolher apenas um estado para ser incluído nesse apêndice <sup>1</sup>. Assim optou-se por escolher um estado representativo, o Acre (AC), visto que esse possui um dos menores percentuais de dígitos brancos apenas 0,91% e 2,15% de nulos.

A escolha do Acre se deve a indicação de Mebane (2009), em que Mebane afirma que as regiões em que o presidente iraniano eleito, Mahmoud Ahmadinejad, obteve votações expressivas tinham um percentual baixo de votos inválidos. Nessas regiões, os dados referentes aos votos de Mahmoud Ahmadinejad divergiam significativamente da distribuição de Benford.

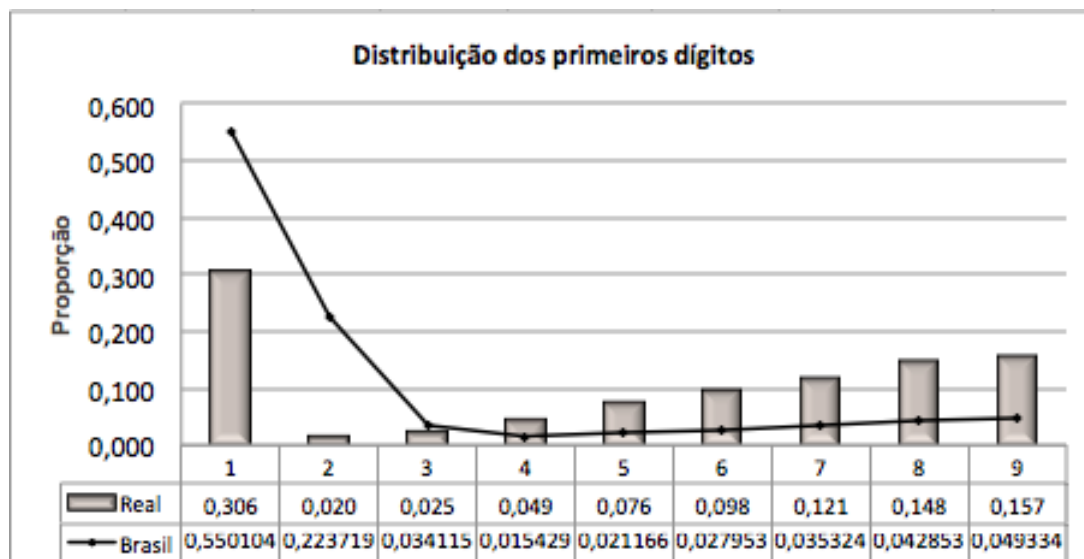
### E.1.1 Primeiro Dígito

Tabela 24 – Teste do Primeiro Dígito conforme a distribuição do Brasil, estado do Acre e código votável 13 - Fonte: elaboração própria

Dígito	Cont.	Real	Brasil	Diferença	Teste Z	CS	MAD
1	501	0,306	0,550	-0,244	19,794	177,249	0,244
2	32	0,020	0,224	-0,204	19,763	305,025	0,204
3	41	0,025	0,034	-0,009	1,950	3,947	0,009
4	80	0,049	0,015	0,033	10,860	118,647	0,033
5	125	0,076	0,021	0,055	15,404	235,592	0,055
6	160	0,098	0,028	0,070	17,028	285,208	0,070
7	198	0,121	0,035	0,086	18,672	339,789	0,086
8	243	0,148	0,043	0,106	21,001	425,897	0,106
9	257	0,157	0,049	0,108	20,026	384,599	0,108
N	1637	1,0				<b>CS Real</b>	<b>0,915</b>
						<b>CS Limite</b>	<b>0,1016</b>

<sup>1</sup> Para visualizar os resultados para as demais unidades federativas, basta acessar a seguinte página no Google Drive: <<https://drive.google.com/drive/folders/0B5N1NvqnAD-xT2hRZlk5TkxzRnM>> .

Figura 17 – Distribuição dos Primeiros Dígitos - Elaboração Própria

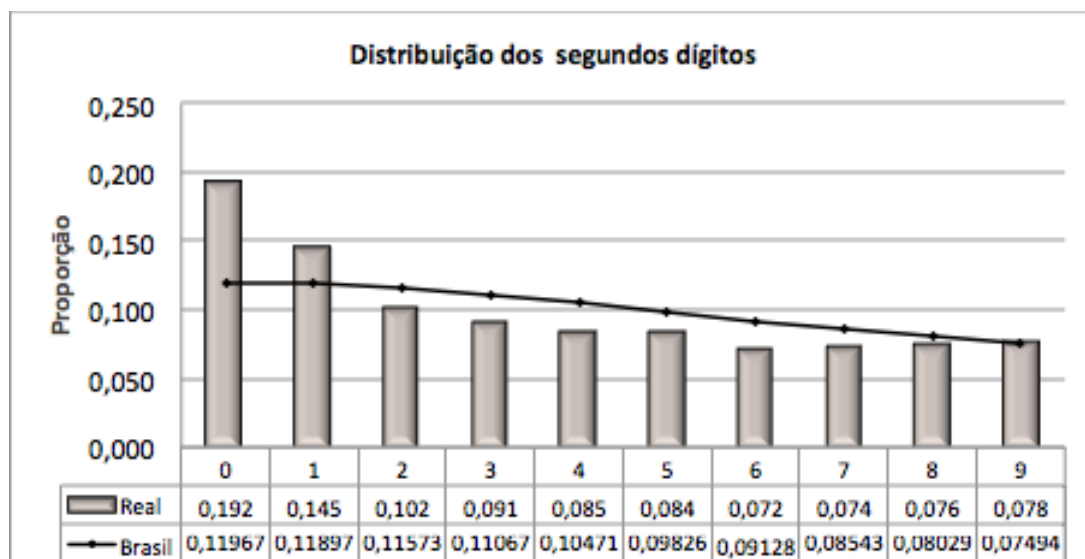


### E.1.2 Segundo Dígito

Tabela 25 – Teste do Segundo Dígito conforme a distribuição do Brasil, estado do Acre e código votável 13 - Fonte: elaboração própria

Dígito	Cont.	Real	Brasil	Diferença	Teste Z	CS	MAD
0	314	0,192	0,1197	0,073	9,013	72,121	0,073
1	237	0,145	0,1190	0,026	3,236	9,447	0,026
2	167	0,102	0,1157	-0,013	1,654	2,535	0,013
3	149	0,091	0,1107	-0,019	2,455	5,535	0,019
4	138	0,085	0,1047	-0,020	2,619	6,331	0,020
5	137	0,084	0,0983	-0,014	1,901	3,405	0,014
6	118	0,072	0,0913	-0,019	2,619	6,438	0,019
7	121	0,074	0,0854	-0,011	1,588	2,438	0,011
8	124	0,076	0,0803	-0,004	0,596	0,378	0,004
9	127	0,078	0,0749	0,003	0,394	0,180	0,003
N	1632	1,0				CS Real 108,807	0,204
						CS Limite 16,919	0,0204

Figura 18 – Distribuição dos Segundos Dígitos - Elaboração Própria





## E.1.3 Dois Primeiros Dígitos

Tabela 26 – Teste dos Dois Primeiros Dígitos conforme a distribuição do Brasil, estado do Acre e código votável 13 - Fonte: elaboração própria

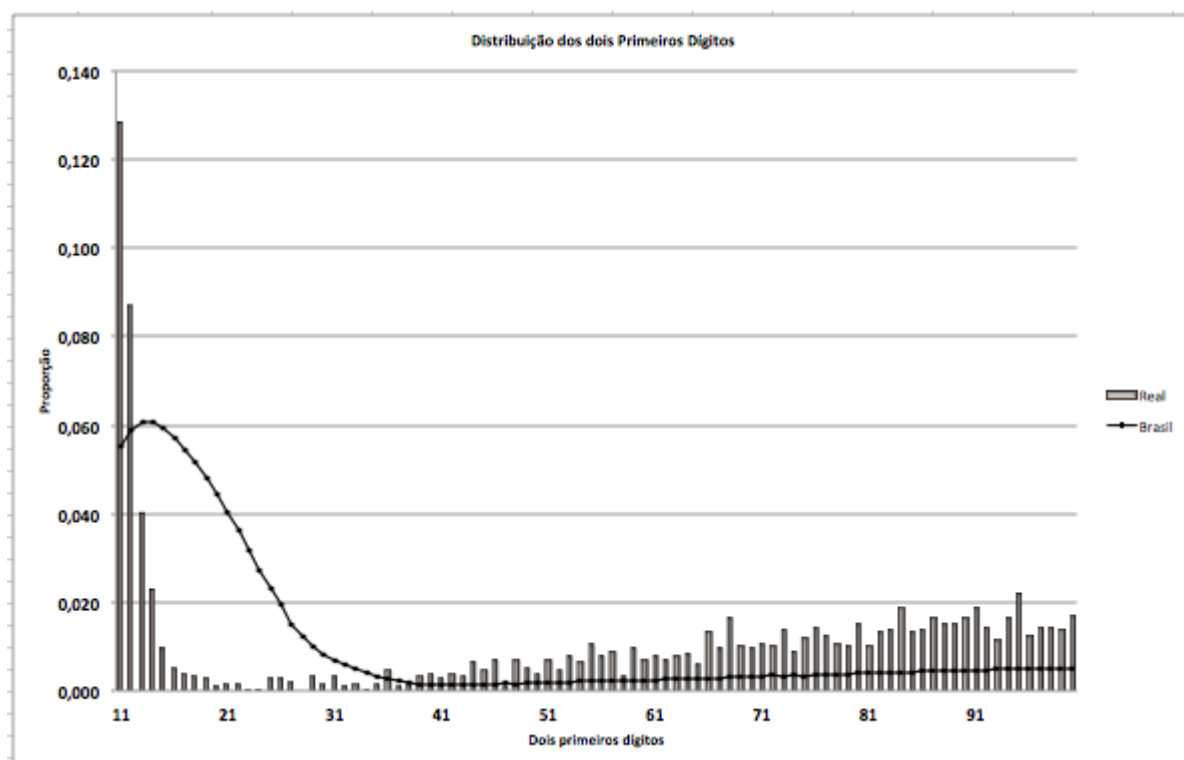
Díg.	Cont.	Real	Brasil	Dif.	Teste Z	CS	MAD
10	209	0,128	0,0554	0,073	9,058	155,32	0,073
11	142	0,087	0,0590	0,028	3,342	21,645	0,028
12	66	0,040	0,0609	-0,020	2,388	11,204	0,020
13	38	0,023	0,0607	-0,037	4,440	37,696	0,037
14	16	0,010	0,0594	-0,050	5,963	67,555	0,050
15	9	0,006	0,0570	-0,051	6,315	75,888	0,051
16	7	0,004	0,0543	-0,050	6,276	75,231	0,050
17	6	0,004	0,0515	-0,048	6,145	72,407	0,048
18	5	0,003	0,0480	-0,045	5,967	68,649	0,045
19	2	0,001	0,0443	-0,043	5,940	68,366	0,043
20	3	0,002	0,0403	-0,039	5,631	59,971	0,039
21	3	0,002	0,0362	-0,034	5,203	53,279	0,034
22	1	0,001	0,0317	-0,031	5,010	49,809	0,031
23	1	0,001	0,0273	-0,027	4,610	42,651	0,027
24	5	0,003	0,0233	-0,020	3,746	28,698	0,020
25	5	0,003	0,0192	-0,016	3,263	22,208	0,016
26	4	0,002	0,0151	-0,013	2,846	17,332	0,013
27	0	0,000	0,0123	-0,012	3,049	20,025	0,012
28	6	0,004	0,0101	-0,006	1,666	6,608	0,006
29	3	0,002	0,0082	-0,006	1,841	8,081	0,006
30	6	0,004	0,0067	-0,003	0,861	2,258	0,003
31	2	0,001	0,0057	-0,004	1,485	5,764	0,004
32	3	0,002	0,0048	-0,003	0,990	3,021	0,003
33	1	0,001	0,0040	-0,003	1,261	4,636	0,003
34	3	0,002	0,0033	-0,001	0,422	1,034	0,001
35	8	0,005	0,0027	0,002	0,893	2,968	0,002
36	2	0,001	0,0022	-0,001	0,246	0,742	0,001
37	3	0,002	0,0017	0,000	0,074	0,011	0,000
38	6	0,004	0,0015	0,002	1,624	5,190	0,002
39	7	0,004	0,0014	0,003	1,744	9,590	0,003
40	5	0,003	0,0014	0,002	0,823	3,271	0,002
41	7	0,004	0,0014	0,003	1,783	9,972	0,003
42	6	0,004	0,0015	0,002	1,159	5,077	0,002
43	11	0,007	0,0015	0,005	3,511	30,938	0,005
44	8	0,005	0,0015	0,003	2,547	12,765	0,003
45	12	0,007	0,0015	0,006	3,880	36,848	0,006
46	2	0,001	0,0016	0,000	-0,179	0,132	0,000
47	12	0,007	0,0016	0,006	3,756	34,616	0,006
48	9	0,006	0,0017	0,004	2,205	13,542	0,004
49	7	0,004	0,0017	0,003	1,356	6,195	0,003
50	12	0,007	0,0018	0,006	3,374	28,188	0,006
51	8	0,005	0,0019	0,003	1,590	7,789	0,003
52	13	0,008	0,0020	0,006	3,468	29,282	0,006
53	11	0,007	0,0020	0,005	2,611	17,656	0,005
54	18	0,011	0,0021	0,009	5,275	62,926	0,009

Tabela 27 – Continuação- Teste dos Dois Primeiros Dígitos conforme a distribuição do Brasil, estado do Acre e código votável 13 - Fonte: elaboração própria

Díg.	Cont.	Real	BRA	Dif.	Teste Z	CS	MAD
55	13	0,008	0,0021	0,006	3,289	26,450	0,006
56	15	0,009	0,0022	0,007	3,928	36,348	0,007
57	6	0,004	0,0022	0,002	0,562	1,724	0,002
58	16	0,010	0,0024	0,007	4,035	37,938	0,007
59	12	0,007	0,0024	0,005	2,513	16,142	0,005
60	13	0,008	0,0024	0,006	2,917	21,058	0,006
61	12	0,007	0,0026	0,005	2,686	14,186	0,005
62	13	0,008	0,0026	0,005	2,647	17,518	0,005
63	14	0,009	0,0026	0,006	2,999	21,898	0,006
64	10	0,006	0,0027	0,003	1,888	7,004	0,003
65	22	0,013	0,0028	0,011	5,438	65,312	0,011
66	16	0,010	0,0028	0,007	3,782	28,112	0,007
67	27	0,017	0,0030	0,014	6,887	102,074	0,014
68	17	0,010	0,0031	0,007	3,485	28,346	0,007
69	16	0,010	0,0032	0,007	3,092	22,736	0,007
70	18	0,011	0,0031	0,008	3,791	33,092	0,008
71	17	0,010	0,0034	0,007	3,168	23,600	0,007
72	23	0,014	0,0033	0,011	5,415	57,596	0,011
73	15	0,009	0,0034	0,006	2,561	16,064	0,006
74	20	0,012	0,0034	0,009	4,129	38,546	0,009
75	24	0,015	0,0035	0,011	5,112	57,370	0,011
76	21	0,013	0,0037	0,009	4,095	37,721	0,009
77	18	0,011	0,0037	0,007	3,151	23,181	0,007
78	17	0,010	0,0038	0,007	2,817	18,865	0,007
79	25	0,015	0,0039	0,011	5,243	53,962	0,011
80	17	0,010	0,0039	0,007	2,737	17,858	0,007
81	22	0,013	0,0040	0,010	4,066	37,006	0,010
82	23	0,014	0,0040	0,010	4,294	40,971	0,010
83	31	0,019	0,0043	0,015	6,234	82,926	0,015
84	22	0,013	0,0043	0,009	4,079	32,652	0,009
85	23	0,014	0,0044	0,010	3,963	35,050	0,010
86	27	0,017	0,0043	0,012	5,072	55,874	0,012
87	25	0,015	0,0044	0,011	4,493	44,391	0,011
88	25	0,015	0,0046	0,011	4,537	40,385	0,011
89	27	0,017	0,0046	0,012	4,852	51,218	0,012
90	31	0,019	0,0046	0,014	5,860	73,402	0,014
91	24	0,015	0,0047	0,010	3,923	34,214	0,010
92	19	0,012	0,0048	0,007	2,587	15,794	0,007
93	27	0,017	0,0048	0,012	4,889	46,881	0,012
94	36	0,022	0,0049	0,017	6,877	99,627	0,017
95	21	0,013	0,0049	0,008	3,253	20,748	0,008
96	24	0,015	0,0050	0,010	3,737	31,134	0,010
97	24	0,015	0,0052	0,010	3,593	28,841	0,010
98	23	0,014	0,0051	0,009	3,365	25,524	0,009
99	28	0,017	0,0052	0,012	4,556	45,141	0,012
<b>N</b>	<b>1627</b>				<b>CS Real</b>	<b>2963,516</b>	<b>1,133</b>
					<b>CS Limite</b>	<b>112,02</b>	<b>0,01259</b>



Figura 19 – Distribuição dos dois Primeiros Dígitos - Elaboração Própria



## E.2 Tabelas Acre para o código votável 45

### E.2.1 Primeiro Dígito

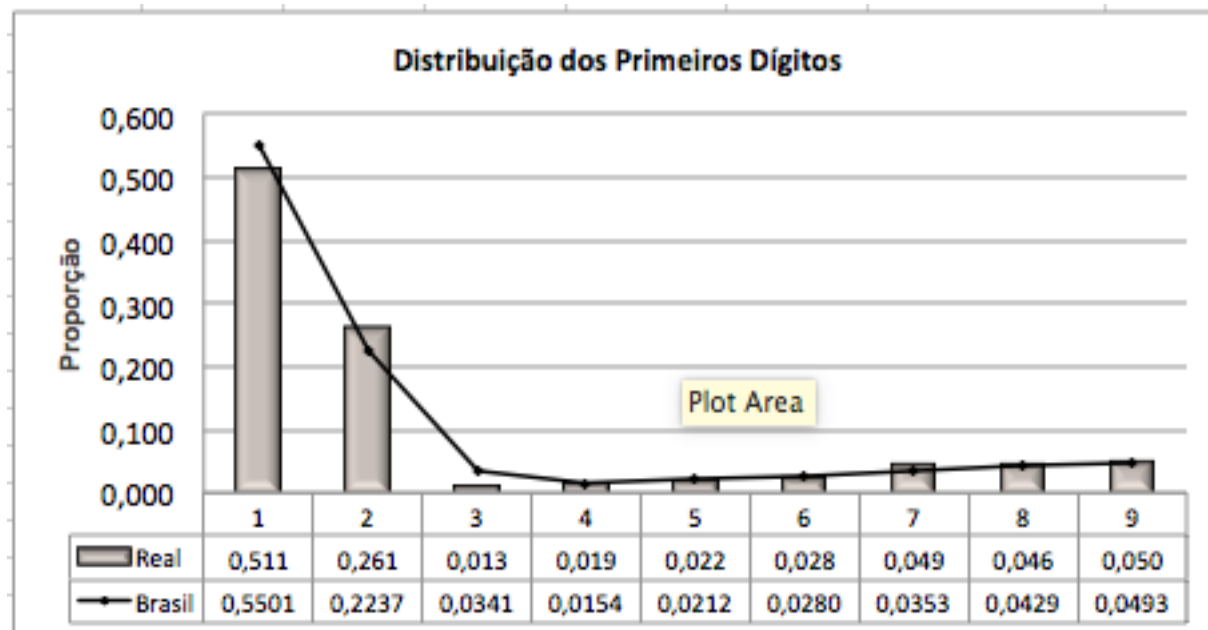
Tabela 28 – Teste do Primeiro Dígito conforme a distribuição do Brasil, estado do Acre e código votável 45 - Fonte: elaboração própria

Dígito	Cont.	Real	Brasil	Diferença	Teste Z	CS	MAD
1	835	0,511	0,550	-0,039	3,119	4,463	0,039
2	427	0,261	0,224	0,038	3,625	10,409	0,038
3	21	0,013	0,034	-0,021	4,655	21,626	0,021
4	31	0,019	0,015	0,004	1,063	1,337	0,004
5	36	0,022	0,021	0,001	0,160	0,060	0,001
6	46	0,028	0,028	0,000	-0,022	0,003	0,000
7	80	0,049	0,035	0,014	2,919	8,632	0,014
8	75	0,046	0,043	0,003	0,551	0,360	0,003
9	82	0,050	0,049	0,001	0,107	0,026	0,001
N	<b>1633</b>	<b>1,0</b>				<b>CS Real 46,916</b>	<b>0,120</b>
						<b>CS Limite 15,507</b>	<b>0,01334</b>

A primeira coluna se refere aos primeiros dígitos dos votos; a segunda corresponde às frequências absolutas desses dígitos no banco de dados; a terceira se refere às frequências da 2a coluna em termos relativos; a quarta coluna corresponde às frequências relativas padrão da Lei de Benford; a quinta é a diferença entre as frequências da 3a e da 4a

colunas; a sexta traz os resultados do Teste Z; a sétima apresenta os resultados do Teste Qui-Quadrado; e a oitava mostra os valores do MDA.

Figura 20 – Distribuição dos Primeiros Dígitos - Elaboração Própria



### E.2.2 Segundo Dígito

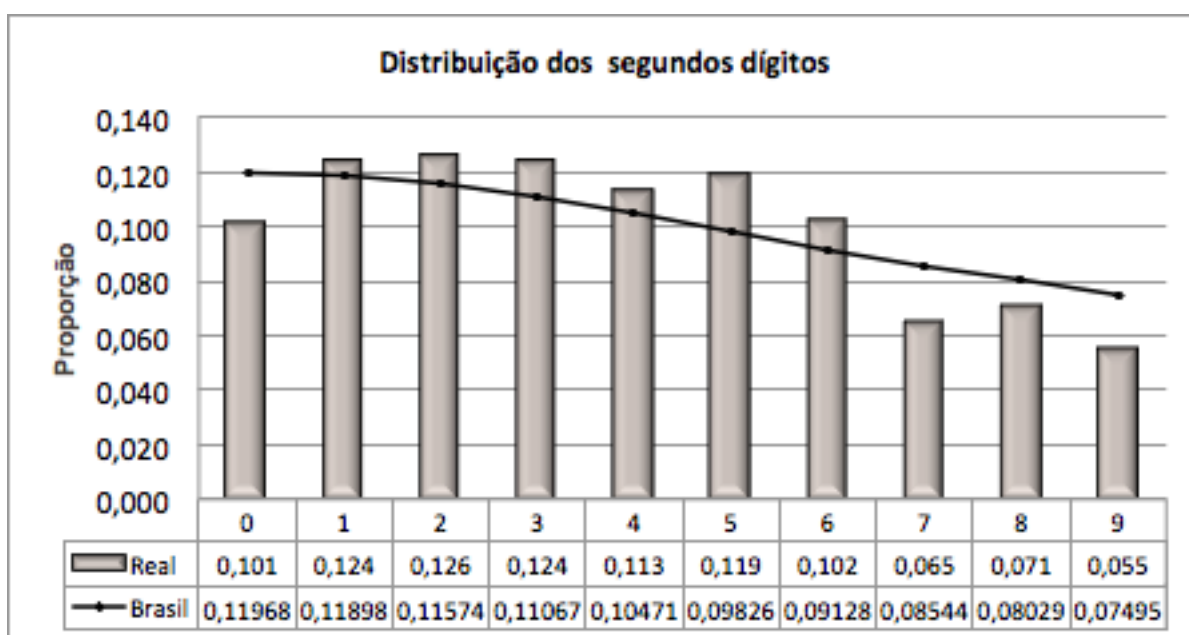
Tabela 29 – Teste do Segundo Dígito conforme a distribuição do Brasil, estado do Acre e código votável 45 - Fonte: elaboração própria

Dígito	Cont.	Real	Brasil	Diferença	Teste Z	CS	MAD
0	165	0,101	0,1197	-0,018	2,232	4,535	0,018
1	201	0,124	0,1190	0,005	0,530	0,285	0,005
2	205	0,126	0,1157	0,010	1,255	1,480	0,010
3	202	0,124	0,1107	0,013	1,694	2,672	0,013
4	184	0,113	0,1047	0,008	1,063	1,091	0,008
5	194	0,119	0,0983	0,021	2,801	7,284	0,021
6	166	0,102	0,0913	0,011	1,462	2,059	0,011
7	105	0,065	0,0854	-0,021	2,972	8,321	0,021
8	115	0,071	0,0803	-0,010	1,381	1,872	0,010
9	90	0,055	0,0749	-0,020	2,960	8,365	0,020
N	1627	1,0			CS Real	37,963	0,137
					CS Limite	16,919	0,01368

Nesta tabela, a primeira coluna se refere aos segundos dígitos dos votos; a segunda corresponde às frequências absolutas desses dígitos no banco de dados; a terceira se refere às frequências da 2a coluna em termos relativos; a quarta coluna corresponde às frequências relativas padrão da Lei de Benford; a quinta é a diferença entre as frequên-

cias da 3a e da 4a colunas; a sexta traz os resultados do Teste Z; a sétima apresenta os resultados do Teste Qui-Quadrado; e a oitava mostra os valores do MDA.

Figura 21 – Distribuição dos Segundos Dígitos - Elaboração Própria





## E.2.3 Dois Primeiros Dígitos

Tabela 30 – Teste dos Dois Primeiros Dígitos conforme a distribuição do Brasil, estado do Acre e código votável 45 - Fonte: elaboração própria

Díg.	Cont.	Real	Brasil	Dif.	Teste Z	CS	MAD
10	78	0,048	0,0554	-0,007	0,866	155,32	1,647
11	111	0,068	0,0590	0,009	1,049	21,645	2,330
12	102	0,063	0,0609	0,002	0,145	11,204	0,087
13	101	0,062	0,0607	0,001	0,088	37,696	0,048
14	88	0,054	0,0594	-0,005	0,571	67,555	0,768
15	110	0,068	0,0570	0,011	1,242	75,888	3,215
16	88	0,054	0,0543	0,000	-0,045	75,231	0,002
17	46	0,028	0,0515	-0,023	2,941	72,407	16,995
18	65	0,040	0,0480	-0,008	1,002	68,649	2,194
19	45	0,028	0,0443	-0,017	2,244	68,366	10,179
20	45	0,028	0,0403	-0,013	1,855	59,971	6,486
21	50	0,031	0,0362	-0,005	0,754	53,279	1,358
22	59	0,036	0,0317	0,005	0,644	49,809	1,052
23	66	0,041	0,0273	0,013	2,226	42,651	10,397
24	54	0,033	0,0233	0,010	1,769	28,698	6,815
25	46	0,028	0,0192	0,009	1,764	22,208	6,886
26	50	0,031	0,0151	0,016	3,537	17,332	26,202
27	26	0,016	0,0123	0,004	0,812	20,025	1,825
28	21	0,013	0,0101	0,003	0,647	6,608	1,313
29	10	0,006	0,0082	-0,002	0,468	8,081	0,849
30	4	0,002	0,0067	-0,004	1,290	2,258	4,407
31	0	0,000	0,0057	-0,006	1,952	5,764	9,307
32	0	0,000	0,0048	-0,005	1,753	3,021	7,855
33	2	0,001	0,0040	-0,003	0,978	4,636	3,081
34	4	0,002	0,0033	-0,001	0,110	1,034	0,335
35	2	0,001	0,0027	-0,001	0,476	2,968	1,291
36	4	0,002	0,0022	0,000	-0,231	0,742	0,037
37	2	0,001	0,0017	-0,001	0,347	0,011	0,236
38	0	0,000	0,0015	-0,001	1,114	5,190	2,433
39	2	0,001	0,0014	0,000	-0,324	9,590	0,038
40	2	0,001	0,0014	0,000	-0,340	3,271	0,031
41	1	0,001	0,0014	-0,001	0,127	9,972	0,694
42	4	0,002	0,0015	0,001	0,256	5,077	0,971
43	5	0,003	0,0015	0,002	0,754	30,938	2,862
44	3	0,002	0,0015	0,000	0,265	12,765	0,138
45	3	0,002	0,0015	0,000	-0,201	36,848	0,119
46	4	0,002	0,0016	0,001	0,197	0,132	0,789
47	1	0,001	0,0016	-0,001	0,258	34,616	0,952
48	2	0,001	0,0017	0,000	-0,073	13,542	0,234
49	5	0,003	0,0017	0,001	0,514	6,195	1,703
50	4	0,002	0,0018	0,001	0,043	28,188	0,405
51	4	0,002	0,0019	0,001	-0,027	7,789	0,273
52	4	0,002	0,0020	0,000	-0,087	29,282	0,179
53	3	0,002	0,0020	0,000	-0,259	17,656	0,031
54	5	0,003	0,0021	0,001	0,247	62,926	0,775

Tabela 31 – Continuação- Teste dos Dois Primeiros Dígitos conforme a distribuição do Brasil, estado do Acre e código votável 45 - Fonte: elaboração própria

Díg.	Cont.	Real	BRA	Dif.	Teste Z	CS	MAD
55	2	0,001	0,0021	-0,001	0,176	0,602	0,001
56	3	0,002	0,0022	0,000	-0,154	0,092	0,000
57	5	0,003	0,0022	0,001	0,189	0,621	0,001
58	3	0,002	0,0024	-0,001	-0,043	0,193	0,001
59	2	0,001	0,0024	-0,001	0,354	0,978	0,001
60	5	0,003	0,0024	0,001	0,041	0,307	0,001
61	4	0,002	0,0026	0,000	0,080	0,012	0,000
62	11	0,007	0,0026	0,004	1,971	10,455	0,004
63	2	0,001	0,0026	-0,001	0,449	1,219	0,001
64	6	0,004	0,0027	0,001	0,538	0,568	0,001
65	5	0,003	0,0028	0,000	-0,197	0,033	0,000
66	1	0,001	0,0028	-0,002	1,199	2,816	0,002
67	4	0,002	0,0030	0,000	-0,058	0,135	0,000
68	4	0,002	0,0031	-0,001	0,014	0,210	0,001
69	4	0,002	0,0032	-0,001	0,052	0,256	0,001
70	9	0,006	0,0031	0,002	0,946	3,102	0,002
71	15	0,009	0,0034	0,006	2,576	16,181	0,006
72	7	0,004	0,0033	0,001	0,503	0,495	0,001
73	6	0,004	0,0034	0,000	-0,158	0,039	0,000
74	10	0,006	0,0034	0,003	1,089	3,781	0,003
75	7	0,004	0,0035	0,001	0,075	0,264	0,001
76	4	0,002	0,0037	-0,001	0,287	0,646	0,001
77	7	0,004	0,0037	0,001	-0,020	0,137	0,001
78	9	0,006	0,0038	0,002	0,531	1,295	0,002
79	6	0,004	0,0039	0,000	0,109	0,023	0,000
80	9	0,006	0,0039	0,002	0,479	1,125	0,002
81	7	0,004	0,0040	0,000	-0,129	0,042	0,000
82	16	0,010	0,0040	0,006	2,360	13,583	0,006
83	10	0,006	0,0043	0,002	0,562	1,344	0,002
84	4	0,002	0,0043	-0,002	0,794	1,233	0,002
85	8	0,005	0,0044	0,001	-0,032	0,104	0,001
86	5	0,003	0,0043	-0,001	0,293	0,607	0,001
87	7	0,004	0,0044	0,000	-0,224	0,003	0,000
88	4	0,002	0,0046	-0,002	0,918	1,649	0,002
89	5	0,003	0,0046	-0,001	0,381	0,797	0,001
90	9	0,006	0,0046	0,001	0,137	0,304	0,001
91	9	0,006	0,0047	0,001	0,079	0,216	0,001
92	2	0,001	0,0048	-0,004	1,240	4,345	0,004
93	7	0,004	0,0048	0,000	0,207	0,084	0,000
94	10	0,006	0,0049	0,001	0,286	0,563	0,001
95	11	0,007	0,0049	0,002	0,747	1,090	0,002
96	7	0,004	0,0050	-0,001	0,026	0,146	0,001
97	7	0,004	0,0052	-0,001	0,101	0,231	0,001
98	7	0,004	0,0051	-0,001	0,091	0,219	0,001
99	11	0,007	0,0052	0,002	0,390	0,781	0,002
<b>N</b>	<b>1627</b>				<b>CS Real</b>	<b>212,750</b>	<b>0,258</b>
					<b>CS Limite</b>	<b>112,02</b>	<b>0,002868</b>

Figura 22 – Distribuição dos dois Primeiros Dígitos - Elaboração Própria

