UMA APLICAÇÃO DA TEORIA DAS PROBABILIDADES NA CONTABILOMETRIA:

A LEI NEWCOMB-BENFORD COMO MEDIDA PARA ANÁLISE DE DADOS NO CAMPO DA AUDITORIA CONTÁBIL

Josenildo dos Santos, José Nelson Barbosa Tenório & Luiz Gustavo Cordeiro da Silva *

INTRODUÇÃO

Este artigo tem como eixos basilares: 1. o estudo da Lei de Newcomb-Benford;[1] 2. o desenvolvimento de um modelo contabilométrico similar ao criado por Nigrini[2] que se fundamenta na relação entre a Lei de Newcomb-Benford e Testes de Hipóteses (Z – Teste e χ^2 – *Teste*) no contexto da auditoria contábil; 3. aplicação do modelo contabilométrico a uma empresa x do estado de Per-

* Josenildo dos Santos é professor doutor do Departamento de Ciências Contábeis da UFPE e do Mestrado Multiinstitucional e Inter-regional em Ciências Contábeis da UnB/UFPB/UFPE/UFRN; José Nelson Barnosa Tenório é aluno do Mestrado Multiinstitucional e Inter-regional em Ciências Contábeis da UnB/UFPB/UFPE/UFRN; e Luiz Gustavo Cordeiro da Silva é professor mestre do Departamento de Ciências Contábeis da UFPE e doutorando em engenharia de Produção.

1 F. Benford, *The law of anomalous numbers. Proceeding of the American Philosophical Society*, n. 78, 1938, pp. 551-472; S. Newcomb, "Note on the frequency of use of the different digits in natural numbers", *AJM*, n. 4, 1881, pp. 39-40.

2 M. J. Nigrini, *Digital analysis using Benford's Law: Tests ? statistics for auditors*, Canada, Global Audit Publication, 2000.

Departamento de Ciências Contábeis e Atuariais
Universidade de Brasília – UnB.
www.unb.br/cca
Volume 6, número 1. Primeiro semestre 2003.

nambuco por meio de um estudo de caso.

Estes eixos foram estudados da seguinte maneira:

1. A Lei de Newcomb-Benford é uma anomalia das probabilidades acarretando que os dígitos 1, 2 e 3 são mais comuns que os dígitos 4, 5, 6 ...9 como primeiro dígito de uma distribuição de números de bom tamanho, isto é, a chance de um número ter sido tirado ao acaso e o primeiro dígito significativo ser 1, 2 e 3 é aproximadamente 60,2%. Este fenômeno foi descoberto empiricamente pelo astrônomo e matemático Simon Newcomb[3] quando ao acaso observou que as primeiras páginas das tábuas de logaritmos, nas bibliotecas, eram mais usadas pois as encontravam mais manuseadas – mais sujas e estragadas que as outras. Isto é, as pessoas iam muito mais em busca dos valores dos logaritmos que começam por 1 do que aqueles que começam com dígito 9. Este fato foi também constatado pelo físico Frank Benford,[4] passados 57 anos, em pesquisa independente da de Newcomb. O trabalho de Benford foi mais aprofundado tendo estudado um conjunto de dados de 20.229 observações advindo de diversas fontes: áreas de rios, números de casas de uma rua, número de artigos de uma revista e outros. Diferentemente de Newcomb seu trabalho foi amplamente divulgado e aplicado por outros pesquisadores.
   Em síntese, este artigo não só apresenta a Lei de Newcomb-Benford como descreve sua aplicação no contexto da Auditoria Contábil, Newcomb, Benford, Hill,[5] Nigrini,[6] Pinkham[7] e Raimi.[8]
2. Análise das diferenças entre as distribuições de probabilidades observadas (*po*) e esperadas (*pe*) segundo a Lei de Newcomb-Benford é estudada a partir de Testes de Hipóteses. Mais precisamente, o $Z-Teste$ é utilizado para medir o grau de significância entre as diferenças $po-pe$. Isto é, se existe suficiente evidencia para rejeitar a hipótese nula $H_0: po = pe$.

3 S. Newcomb, op. cit.

4 Frank Benford, op. cit.

5 T. P. Hill, "The first digit phenomenon", *American Scientist*, n. 86, 1998, pp. 358-363.

6 M. J. Nigrini, op. cit.

7 R. S. Pinkhan, "On the distribution of first significant digits", *Annals of Mathematical Statistics*, n. 32, 1961, pp 1.223-1.230.

8 R. Raimi, "The peculiar distribution of first significant digits", *Scientific American*, Vol. 221, n. 6, 1969, pp. 109-120.

Além disso, o χ^2 – Teste é utilizado para medir o nível de significância se na sua totalidade as duas distribuições estão ou não compatíveis.

3. Finalmente, na prática da Auditoria Contábil o modelo contabilométrico desenvolvido aqui é aplicado a uma empresa x do estado de Pernambuco no período de 1998 a 2001. Existindo, assim, um forte indício de divergências entre os dados esperados e os dados encontrados na empresa x no ano de 2000. Essa etapa de natureza prática será realizada de forma a comprovar as premissas que serão fixadas no decorrer do desenvolvimento deste trabalho. O raciocínio dedutivo será a linha mestre que governará todo o desenvolvimento desta pesquisa que será fundamentado em conhecimento das ciências matemáticas para comprovar a veracidade dos resultados que serão obtidos.

Metodologia

Este trabalho propõe-se a disseminar uma metodologia pouco utilizada no campo da Auditoria Contábil e da Auditoria Digital no Brasil. De fato, uma anomalia das probabilidades, a Lei de Newcomb-Benford, será usada como uma distribuição de medidas para detectar desvios de padrões – erros ou fraudes.

Para se ter idéia desta lei, imagine um conjunto numérico de muitos elementos, por exemplo contadores e auditores estão freqüentemente se deparando com lista de dados – informações financeiras denominadas dados monetários. Tais listas incluem por exemplo, segundo Nigrini, as notas fiscais de venda ou notas fiscais de compra. Escolhido ao acaso um número em uma dessas condições acima, pergunta-se: qual é a probabilidade de que o primeiro dígito deste número seja igual a 1? Se você responder 1/9 (11,1%), argumenta-se que deseja a ocorrência de um entre nove números possíveis que é a natural ocorrência, prepare-se para uma surpresa: segundo a Lei de Newcomb-Benford a resposta correta é 30,1%. Portanto, é uma anomalia das probabilidades.

Fundamentado na premissa de que "números" que começam com os dígitos 1, 2 ou 3 são mais comuns do que aqueles que começam com os dígitos 4, 5, 6, ..., 9, a Lei de Newcomb-Benford tem sido amplamente aplicada por Nigrini no campo da Auditoria Contábil, mais precisamente na descoberta de fraudes contábeis nos Estados Unidos. Entretanto, no Brasil, é incipiente o uso desta Lei por contadores e por auditores, bem como não

se encontra literatura disponível escrita em língua portuguesa sobre o tema.

Nesta ótica, o auditor pode propor que as notas de entrada e de saída sejam analisadas por testes de hipóteses em conformidade – consonância – com a freqüência teórica do primeiro dígito, segundo a Lei de Newcomb-Benford. Isto é, não se deve descartar que as probabilidades observadas – reais – e as declaradas pela lei podem representar viés, pois muitas vezes são utilizadas amostras de tamanho razoáveis e que dissolvem as discrepâncias, quando as reais são comparadas com os resultados da lei. Isto porque a Lei de Newcomb-Benford apresenta um elevado grau de precisão em sua aplicação, em consonância com o tamanho da amostra em estudo. Este é o princípio básico na qual a referida lei se apóia.

Destarte, a aplicação de Testes de Hipóteses pode melhorar substancialmente a eficiência da Lei de Newcomb-Benford, quando a mesma é usada para detectar desvios em determinadas situações contábeis. Por exemplo, o uso destes testes permitirá aos Auditores avaliarem até que ponto as diferenças detectadas entre as probabilidades observadas (*po*) e as declaradas pela Lei (*pe*) são provenientes da aleatoriedade, dissolvendo assim, as impropriedades da amostragem, quando são originadas de erros sistêmicos. Então, serão demonstradas as circunstâncias nas quais os testes: $Z - Teste$ e $\chi^2 - Quadrado$ podem ser aplicados, reportando-se, também, às limitações de uso de cada um destes testes.

De modo geral, a Lei de Newcomb-Benford para os primeiros dígitos garante que:

Tabela 1

Lei de Newcomb-Benford para o primeiro dígito

Dígito	Probabilidade
1	0,30103
2	0,176091
3	0,124939
4	0,09691
5	0,0791812
6	0,0669468
7	0,0579919
8	0,0515525
9	0,0407575

Portanto, segundo a Lei Newcomb-Benford, a probabilidade de cada primeiro dígito, *d*, não é 1/9 e sim o logaritmo na base 10 da quantidade $(d+1)/d$. Assim, a probabilidade de 1 ser o primeiro dígito significativo é 30,1%, enquanto a probabilidade de encontrar o 9 nesse lugar é, aproximadamente, 4,6%. Este obscuro resultado da teoria da probabilidade está a disposição dos auditores para ajudar a combater a fraude fiscal, erros contábeis e erros de digitação. Por exemplo, um excesso de 8 e uma tremenda escassez de 6, como primeiro dígito, devem levar o auditor a suspeitar de um superfaturamento nas contas com 8 como primeiro dígito e, conseqüentemente, é proposta uma extensão dos testes a todas as notas fiscais.

Ademais, a pesquisa da revisão bibliográfica mostra que a maior parte dos crimes de colarinho branco é cometida em completa incompatibilidade com a Lei Newcomb-Benford, pois é usado o senso "comum" dos cálculos da probabilidade, isto é, que todos os primeiros dígitos significativos em uma distribuição aleatória de números têm probabilidade de aproximadamente 11%.[9]

Aplicação da Contabilometria por meio da Lei de Newcomb-Benford

Interpretação intuitiva

Considere as seguintes situações-problemas (*SP*) fundamentadas no trabalho de Nigrini:[10]

SP1 Em uma cidade com uma população de 10 mil habitantes. Suponha-se que a população está crescendo a uma taxa de 10% ao ano. A 10 mil o número de habitantes tem o primeiro dígito 1, e este primeiro dígito irá permanecer até a população atingir 20 mil habitantes. A taxa de crescimento da população de 10 mil a 20 mil é de 100%. Porém, para taxa de crescimento proposta acima, 10% ao ano, leva-se 10 anos até alcançar uma população de 20 mil habitantes;

SP2 Para uma população de 20 mil, o primeiro dígito é 2, e 2 irá continu-

9 Cf. T. P. Hill, op. cit.

10 M. J. Nigrini, op. cit.

ar como sendo o primeiro dígito até a população atingir 30 mil habitantes. A taxa de crescimento da população de 20 mil a 30 mil é de 50%. Porém, para taxa de crescimento de 10% ao ano leva-se 5 anos até alcançar uma população de 30 mil habitantes;

SP3 Para uma população de 30 mil, o primeiro dígito do número da população é 3, e 3 irá continuar como sendo o primeiro dígito até a população atingir 40 mil. A taxa de crescimento da população de 30 mil a 40 mil é 33%. Porém, para a taxa de crescimento de 10% ao ano leva-se 3 anos e 3 meses até alcançar uma população de 40 mil habitantes;

SP4 Para uma população de 40 mil, o primeiro dígito do número da população é 4, e 4 irá continuar como sendo o primeiro dígito até a população atingir 50 mil. A taxa de crescimento da população de 40 mil a 50 mil é 25%. Porém, para taxa de crescimento de 10% ao ano leva-se 2 anos e 6 meses até alcançar uma população de 50 mil habitantes;

SP5 Para uma população de 50 mil, o primeiro dígito do número de habitantes é 5, 5 irá continuar como sendo o primeiro dígito até a população atingir 60 mil. A taxa de crescimento da população de 50 mil a 60 mil é de 20%. Porém, para a taxa de crescimento de 10% ao ano leva-se 2 anos para alcançar uma população de 60 mil habitantes;

SP6 Finalizando, para a população de 90 mil, o primeiro dígito do número de habitantes é 9, 9 irá continuar como sendo o primeiro dígito até a população atingir 100 mil. A taxa de crescimento da população de 90 mil a 100 mil é de 11%. Porém, para a taxa de crescimento de 10% ao ano leva-se 1 ano e 1 mês para alcançar uma população de 100 mil habitantes.

Tabela 2

Soluções das situações-problemas

Primeiro dígito	Taxa total de crescimento	Tempo de crescimento à taxa de 10%
1	100%	10 anos
2	50%	5 anos
3	33%	3 anos e 3 meses
4	25%	2 anos e 6 meses
5	20%	2 anos
6	16,6%	1ano e 8 meses
7	15%	1 ano e 6 meses
8	11,25%	1 ano e 1 mês e 8 dias
9	11%	1 ano e 1 mês

Em resumo, tem-se, na Tabela 2, as soluções das situações-problemas acima:

A Tabela 2 mostra que, uma vez tendo a população atingindo 10 mil habitantes, é requerido o dobro do tempo para que o primeiro dígito 1 mude para o primeiro dígito 2. Este tempo vai decrescendo de modo que é requerido aproximadamente 10 anos para que o primeiro dígito 1 mude para o primeiro dígito 9. Isto leva a concordar, intuitivamente, com a Lei Newcomb-Benford, e discordar, fortemente, da maioria das pessoas que concorda que a probabilidade do primeiro dígito significativo d, onde d pertence { 1, 2, 3, ..., 9} é igual a 1/9. De fato, a Lei Newcomb-Benford nega que qualquer algarismo de 1 a 9 como primeiro dígito de um número retirado de amostra aleatória, e de bom tamanho, seja considerado igualmente prováveis.

Número de Casas de uma Rua

Considere os números das casas de uma rua, em média, pode-se supor que existem 50 casas em uma rua. Sendo assim, é fácil contar quantos números de casas começam com o dígito 1. São as casas de número 1, 10, 11, 12, 13, até a casa 19. A probabilidade de uma casa escolhida ao acaso começar pelo dígito 1 é de 11/50 ou seja 22%. Em contrapartida, haverá apenas uma casa que o número começa com o dígito 9, é a casa de número 9. A probabilidade de um número de uma casa começar por este dígito é de 1/50, apenas 2%.

De fato, para qualquer um dos nove dígitos ter a mesma probabilidade de ocorrência no início do número de uma casa é necessário que a rua tenha exatamente 9 casas, ou 99, ou 999 etc. Casos, realmente raros. Em todas as outras circunstâncias, os dígitos maiores são desfavorecidos e os menores aparecem com maior freqüência.

Taxas de Crescimento

Suponha que um índice de produção quando foi elaborado, começava em 100, sem perda de generalidade. Este tipo de índice tem tendência a crescer. Suponha, também, que a taxa de crescimento é constante e que o valor do índice duplica todos os anos. Durante, um ano, o índice mantém-se nos cento e tal. No início do terceiro, 400, e assim por diante. O dígito 1 permanece como primeiro dígito do índice durante doze meses. Já os dígitos 2 e 3 aparecem como primeiro dígito do índice durante cer-

ca de seis meses cada. Quando se chega ao dígito 9, o seu tempo é de cerca de um mês. Agora é fácil perceber que, escolhendo um elemento ao acaso, o mais provável é encontrar um dígito baixo a liderar o índice. Portanto, se considerar os dígitos menores que 4 a chance é de 60,2%.

Analisando o fenômeno do primeiro dígito dedutivamente

O objeto desse trabalho impõe a clarificação der certas questões, como, por exemplo: qualquer um dos nove dígitos {1, 2, 3, ..., 9} tem a mesma probabilidade de ocorrência como primeiro dígito de um número? o senso comum diz que sim, mas os fenômenos reais mostram que não. Portanto, por que motivo deverá um número escolhido ao acaso ter uma maior probabilidade para começar com os primeiros dígitos: 1, 2 e 3? Neste contexto, só as Ciências Matemáticas podem contribuir para dirimir esta dúvida. Será assim dada uma justificativa matemática para a Lei Newcomb-Benford sem os rigores da teoria das probabilidades, pois isto não cabe no contexto da contabilometria.

Pesquisas no campo da teoria das probabilidades, tais como Hill,[11] Pinkham[12] e Raimi,[13] mostram que a Lei Newcomb-Benford aplica-se ao conjunto de dados que tem as seguintes propriedades:

a. é invariante escalar;

b. advém de uma escolha a partir de uma variedade de diferentes fontes.

Este resultado é obtido de uma análise mais rigorosa da Teoria do Limite Central na forma de teoremas para a mantissa de variáveis randômicas sobre o efeito da multiplicação. Neste sentido, quando o número de variáveis cresce, a função densidade tende a distribuição logarítmica. Hill demonstrou rigorosamente que a "distribuição da distribuição" obtida a partir de amostras aleatórias advindas a partir de uma variedade de diferentes distribuições é a distribuição de Newcomb-Benford.[14]

11 T. P. Hill, op. cit.; idem, "Base-Invariance implies Benford's Law", *Proceedings of the American Mathematical Society*, n. 13, 1995, pp. 887-895.

12 R. S. Pinkhan, op. cit.

13 R. Raimi, op. cit.

14 T. P. Hill, "A statistical derivation of the significant-digit law", *Statistical Science*, n. 4, 1996, pp. 354-363.

Demonstração da Lei Newcomb-Benford

Suponha-se que existe uma distribuição de probabilidade universal $p(x)$ no conjunto dos elementos caracterizados anteriormente, então p é invariante-escalar, isto é:

(1) $p(kx) = \lambda(k)p(x)$ para uma função real de variável real λ.

Como se pode supor que

$\int p(x)dx = 1$, segue-se fazendo uma mudança de variável $u = kx$ na primeira integral que:

$$\int p(kx)dx = \int \lambda(k)p(x)dx$$

$$1/k\int p(u)du = \lambda(k)\int p(x)dx$$

Isto acarreta que:

$$\lambda(k) = \frac{1}{k}$$

Derivando a equação (1) na variável k, tem-se:

$$p'(kx).x = \lambda'(k)p(x)$$

e para $\lambda = 1$ resulta a equação diferencial:

(2) $xp'(x) = -p(x)$

Resolvendo (2) vem:

$$\int_1^x \frac{p'(t)}{p(t)}dt = \int_1^x \frac{1}{t}dt$$

$$Lnp(x) = -Lnx$$

Portanto, $p(x) = \frac{1}{x}$ é solução de (2).

A distribuição $p(x) = 1/x$ não representa uma distribuição de probabilidade própria, mas como os fenômenos reais impõem um ponto de corte na distribuição pode-se, assim, ser considerada esta distribuição para efeito dos cálculos das probabilidades para a ocorrência do primeiro dígito. Então:

P (primeiro dígito significativo = d) =

$$\frac{\int_d^{d+1} p(x)dx}{\int_1^{10} p(x)dx} = \frac{Ln(d+1) - Ln(d)}{Ln10 - Ln1}$$

Limites de intenção:

$$\frac{Ln\left(\frac{d+1}{d}\right)}{Ln10} = \log_{10}\left(1 + \frac{1}{d}\right)$$

Portanto, em particular, fica demonstrado que as probabilidades de 1, 2 e 3 serem os primeiros dígitos de um número são: $\log_{10} 2 = 0{,}30103$, $\log_{10} 1{,}5 = 0{,}176091$ e $\log_{10} 4/3 = 0{,}124939$, respectiva-

mente. Isto responde às perguntas acima. Além disso, está deduzida a Lei Newcomb-Benford de maneira simplificada.

Um esquema geométrico para a Lei Newcomb-Benford

Embora haja muitas "explanações longas e eruditas" da Lei Newcomb-Benford na literatura, principalmente nos Estados Unidos, ela pode ser explicada de maneira simples pelo seguinte esquema geométrico:

1	2	3	4	5	6	7	8	9

De fato, uma distribuição de ocorrências em consonância com a Lei Newcomb-Benford tem como premissa subjacente que a população estará distribuída razoavelmente uniforme em uma escala logarítmica. Conseqüentemente,

Pr (primeiro dígito significativo = *d*) =

$$\frac{\log_{10}(d+1)-\log_{10} d}{\log_{10} 10-\log_{10} 1}$$

$$\log_{10}\left(1+\frac{1}{d}\right)$$

Assim:

Pr (primeiro dígito significativo = 1) =

$$\log_{10}(1+1)=\log_{10} 2$$

Pr (primeiro dígito significativo = 2) =

$$\log_{10}(1+1/2)=\log_{10} 3/2=\log_{10} 1{,}5$$

Então, isto comprova o trabalho de Newcomb, em 1881, quando os valores dos logaritmos que iniciam com os dígitos de 1 a 9 são plotados numa escala logarítmica da seguinte maneira:

30,1%	17,6%	12,5%	7,9%	6,7%	5,8%	5,1%	4,6%	
1	2	3	4	5	6	7	8	9

Isto significa, por exemplo, que todos os números que começam pelo dígito 1 ocupam 30,1% do total do tamanho desta escala. Em particular, os números como 1,8005, 1,534 e 1,21 estão nesta região.

Estudo de caso

Analisar uma população de aproximadamente 8 mil notas fiscais de venda emitidas no período de 1998 a 2001, que tenham sido objeto de verificação pelos postos fiscais de controle da Secretaria da Fazenda, visando identificar possíveis subfaturamento dos valores de venda que devem ser informados à Secretaria da Fazenda de Pernambuco.

O novo paradigma da tecnologia da informação permite ao ramo da auditoria digital, o que antes era inviável, que o auditor possa realizar seus trabalhos com análise da população inteira. Entretanto o uso de amostras continua sendo o caminho mais lógico visando comprovar a existência física dos dados e a sua apreciação em um tempo considerado razoável. Todavia, com a utilização de amostras – ainda que estas sejam significativas – nos trabalhos de auditoria, não se pode descartar a possibilidade de ocorrência de erros, que podem conduzir o auditor a conclusões não condizentes com a realidade. Isso porque os indivíduos de uma população – os itens que a compõe – variam.

Assim, na estimação da probabilidade de ocorrência de um determinado evento correspondente a um dado item da população sobre a experiência – partindo da amostra –, devem ser levadas em consideração essas variações e necessário se faz que o possível erro cometido seja estimado.

Estes erros podem ocorrer ou serem considerados como função do tipo de amostra escolhida pelo auditor – que pode ser probabilística, ou não – ou podem estar relacionados tanto à aleatoriedade como à não representatividade do tamanho da mesma, ou seja, o tamanho da amostra pode não ser suficiente para garantir conclusões fidedignas sobre a população – que é o todo sobre o qual se pretende inferir –, o que, mais uma vez, deve chamar a atenção do auditor.

Nesta ótica, uma forma racional do auditor se orientar é utilizar a contabilometria como metodologia científica, em particular neste trabalho será utilizado Testes de Hipóteses que permitem a análise sobre o grau de significância das divergências que poderão ocorrer entre as distribuições de probabilidades esperadas (p_2), com base na Lei Newcomb-Benford, e as observadas (p_0) do fenômeno contábil que está sob auditoria visando à verificação de desvios padrões nos valores das notas fiscais de venda da empresa *x* ao longo dos anos de 1998 a 2001.

Em outros termos, a análise da empresa *x* fundamenta-se nos seguintes pilares: comparação entre as distribuições observadas e a esperadas segundo a Lei Newcomb-Benford; e verifica-

ção da significância das diferenças entre a probabilidade observada – real –, p_0, e a probabilidade esperada, p_e, segundo a Lei Newcomb-Benford.

Para tanto, aplica-se Teste de Hipótese da seguinte forma:

A hipótese nula H_0 é definida como sendo: as duas distribuições de probabilidades observadas e esperadas são iguais. Isto é,

$$H_0 : (p^o) = (p^e)$$

Para estudar o nível de significância entre as diferenças $p^o - p^e$ utiliza-se o $Z - Teste$:

$$Z = \frac{|p^o - p^e| - 1/2n}{\sqrt{\frac{p^e(1-p^e)}{n}}}$$

onde n é o número de observações.

$\frac{1}{2n}$ é o termo de correção de continuidade e só é usado quando ele é menor que $|p^0 - p^e|$.

O nível de significância é $\alpha = 0{,}05$ e Zc crítico é igual a 1,96.

Para estudar se as duas distribuições de probabilidade estão em conformidade uma com a outra, ou se a distribuição de probabilidade observada (p^o) é "igual" a distribuição esperada (p^e) segundo a Lei Newcomb-Benford utiliza-se o $\chi^2 - Teste$ da seguinte maneira:

$$\chi^2 = \sum_{d=1}^{9} \frac{(PO - PE)^2}{PE}$$

onde PO e PE são as proporções observadas e esperadas definidas por:

$PO = (p^o)x$ (população)

$PE = (p^e)x$ (população)

O nível de significância é $\alpha = 0{,}05$, o grau de liberdade é 8 e valor crítico do Qui-quadrado é 15,507.

Portanto, fazendo os cálculos para empresa x ao longo dos anos 1998 a 2001, vem:

1. Analisando a empresa X no ano de 1998, tem-se:

Tabela 3

Aplicação da Lei Newcomb-Benford na empresa X: ano 1998 - Z-teste

Dígito	Freqüência	Lei Newcomb- Benford Prob.Esp. (p^e)	Prob.Observada (p^o)	Desvio $p^e - p^o$	Z-Teste
1	208	0,301	0,326	0,025	1,335
2	99	0,176	0,155	-0,021	1,342
3	75	0,125	0,117	-0,008	0,552
4	62	0,097	0,097	0,000	-0,067
5	47	0,079	0,074	-0,005	0,395
6	41	0,067	0,064	-0,003	0,224
7	48	0,058	0,075	0,017	1,754
8	27	0,051	0,042	-0,009	0,944
9	32	0,046	0,050	0,004	0,388
Soma	639	1	1	0	

Tabela 4

Aplicação da Lei Newcomb-

Dígito	Frequência	Lei Newcomb- Benford Prob.Esp. (p^e)	Prob.Observada (p^o)	Valor Esperado (V^e)	Chi-Squ:	Diferença $p^o - p^e$
1	208	0,301	0,326	192,339	1,275	0,025
2	99	0,176	0,155	112,464	1,612	-0,021
3	75	0,125	0,117	79,875	0,298	-0,008
4	62	0,097	0,097	61,983	0,000	0,000
5	47	0,079	0,074	50,481	0,240	-0,005
6	41	0,067	0,064	42,813	0,077	-0,003
7	48	0,058	0,075	37,062	3,228	0,017
8	27	0,051	0,042	32,589	0,959	-0,009
9	32	0,046	0,050	29,394	0,231	0,004
Soma	639	1	1	642,000	7,92	

Os testes Z e $\chi - Quadrado$, de acordo com as Tabelas 3 e 4, mostram que a distribuição de notas fiscais emitidas pela empresa x no ano de 1998 está em consonância com a Lei Newcomb-Benford.

Neste caso, a referida lei não garante a existência de desvios. Isto é, não existe suficiente evidência para rejeitar a hipótese nula *H0*.

2. Analisando a empresa *X* no ano de 1999, tem-se:

Tabela 5

Aplicação da Lei Newcomb-Benford na empresa *X* ano 1999 – z teste

Dígito	Freqüência	Lei Newcomb- Benford Prob.Esp.(p^e)	Prob.Observada (p^o)	Desvio $p^e - p^o$	Z-Teste
1	685	0,301	0,294	-0,007	0,714
2	379	0,176	0,163	-0,013	1,62
3	287	0,125	0,123	-0,002	0,26
4	216	0,097	0,093	-0,004	0,617
5	218	0,079	0,094	0,015	2,645
6	184	0,067	0,079	0,012	2,274
7	133	0,058	0,057	-0,001	0,162
8	119	0,051	0,051	0	-0,047
9	107	0,046	0,046	0	-0,049
Soma	2.328	1	1		

Tabela 6

Aplicação da Lei Newcomb-Benford na empresa *X* ano 1999 – qui-quadrado

Dígito	Freqüência	Lei Newcomb- Benford Prob.Esp.(p^e)	Prob.Observada (p^o)	Valor Esperado (V^e)	Chi-Squar	Diferença $p^o - p^e$
1	685	0,301	0,294	700,728	0,353	-0,007
2	379	0,176	0,163	409,728	2,304	-0,013
3	287	0,125	0,123	291	0,055	-0,002
4	216	0,097	0,093	225,816	0,427	-0,004
5	218	0,079	0,094	183,912	6,318	0,015
6	184	0,067	0,079	155,976	5,035	0,012
7	133	0,058	0,057	135,024	0,030	-0,001
8	119	0,051	0,051	118,728	0,001	0
9	107	0,046	0,046	107,088	0,000	0
Soma	2.328	1	1	2.331	14,513	

Por um lado, o *Z-teste* mostra que as diferenças $p^o - p^e$ correspondentes aos dígitos 5 e 6 são significantes, isto é, pode-se rejeitar a hipótese nula nestes casos, pois existe suficiente evidência. Por outro lado, o χ^2 – *Teste* mostra que não existe suficiente evidência para rejeitar a hipótese nula, isto é, as duas distribuições de probabilidades, na sua totalidade, são compatíveis.

3. Analisando a Empresa *X* no ano de 2000, tem-se:

Tabela 7

Aplicação da Lei Newcomb-Benford na empresa *X*, ano 2000 – z teste

Dígito	Freqüência	Lei Newcomb- Benford Prob.Esp.(p^e)	Prob.Observada (p^o)	Desvio $p^e - p^o$	Z-Teste
1	1.086	0,301	0,291	-0,010	1,314
2	590	0,176	0,158	-0,018	2,865
3	394	0,125	0,106	-0,019	3,484
4	335	0,097	0,090	-0,007	1,417
5	303	0,079	0,081	0,002	0,422
6	304	0,067	0,082	0,015	3,631
7	266	0,058	0,071	0,013	3,362
8	235	0,051	0,063	0,012	3,294
9	217	0,046	0,058	0,012	3,459
Soma	3.730	1	1	0	

TABELA 8

APLICAÇÃO DA LEI NEWCMB-BENFORD NA EMPRESA *X* ANO 2000 – QUI-QUADRADO

Dígito	Freqüência	Lei Newcomb- Benford Prob.Esp. (p^e)	Prob.Observada (p^o)	Valor Esperado (V^e)	Chi-Squa	Diferença $p^o - p^e$
1	1.086	0,301	0,291	1.122,73	1,202	1,314
2	590	0,176	0,158	656,48	6,732	2,865
3	394	0,125	0,106	466,25	11,196	3,484
4	335	0,097	0,090	361,81	1,987	1,417
5	303	0,079	0,081	294,67	0,235	0,422
6	304	0,067	0,082	249,91	11,707	3,631
7	266	0,058	0,071	216,34	11,399	3,362
8	235	0,051	0,063	190,23	10,536	3,294
9	217	0,046	0,058	171,58	12,023	3,459
Soma	3.730	1	1	3.733	67,017	

O ano 2000 é um ano constatadamente problemático para a empresa *x* pois, de um lado o Z – Teste mostra que só não existe suficiente evidência para rejeitar a hipótese nula, H_0, para os dígitos 1, 4 e 5 mas no restante dos dígitos as diferenças $p^o - p^e$ são significantes. De outro lado, o χ^2 – *Teste* mostra que, na totalidade, as duas distribuições de probabilidade analisadas não são compatíveis, pois existe suficiente evidencia para rejeitar a hipótese nula, $H_0 = (\chi^2 = 67{,}017)$. Portanto, existem indícios de desvios, principalmente, nas notas fiscais, pois 9 como primeiro dígito significativo das notas fiscais em tela, é aquele que mostra mais evidência de desvios.

4. ANALISANDO A EMPRESA *X* NO ANO 2001, TEM-SE:

TABELA 9

APLICAÇÃO DA LEI NEWCOMB-BENFORD NA EMPRESA X, ANO 2001 – Z TESTE

Dígito	Freqüência	Lei Newcomb- Benford Prob.Esp. (p^e)	Prob.Observada (p^o)	Desvio $p^e - p^o$	Z-Teste
1	301	0,301	0,264	-0,037	2,689
2	195	0,176	0,171	-0,005	0,404
3	130	0,125	0,14	-0,011	1,077
4	108	0,097	0,095	-0,002	0,178
5	117	0,079	0,103	0,024	2,947
6	92	0,067	0,081	0,014	1,830
7	85	0,058	0,075	0,017	2,390
8	63	0,051	0,055	0,004	0,546
9	47	0,046	0,041	-0,005	0,734
Soma	1.138	1	0,999	-0,001	

TABELA 10

APLICAÇÃO DA LEI NEWCOMB-BENFORD NA EMPRESA *X* ANO 2001 – QUI-QUADRADO

Dígito	Freqüência	Lei Newcomb- Benford Prob.Esp. (p^e)	Prob.Observada (p^o)	Valor Esperado (V^e)	Chi-Squa	Diferença $p^o - p^e$
1	301	0,301	0,264	342,538	5,037	-0,037
2	195	0,176	0,171	200,288	0,140	-0,005
3	130	0,125	0,14	142,250	1,055	-0,011
4	108	0,097	0,095	110,386	0,052	-0,002
5	117	0,079	0,103	89,902	8,168	0,024
6	92	0,067	0,081	76,246	3,255	0,014
7	85	0,058	0,075	66,004	5,467	0,017
8	63	0,051	0,055	58,038	0,424	0,004
9	47	0,046	0,041	52,348	0,546	-0,005
Soma	1.138	1	0,999	1.141	24.144	

No ano de 2001, o Z – Teste mostra que, para os dígitos 1, 5 e 7 as diferenças $p^o - p^e$ são significantes, e existe suficiente evidencia para rejeitar a hipótese nula H_0 nestes casos. Além disso, o χ^2 – *Teste* evidencia que as duas distribuições de probabilidades não estão em consonância, isto equivale a mostrar que existe suficiente evidencia para rejeitar a hipótese nula, H_0. Neste caso, o auditor pode suspeitar que existem desvios nas notas fiscais de valores que iniciam com 1, 5 e 7.

Relevância

A relevância dessa pesquisa vai de encontro à sistematização dos direcionadores dos trabalhos da auditoria quando a aplicação da Lei Newcomb-Benford se faz presente, e quando um dos objetivos é detecção de evasão fiscal – indícios de fraudes –, usando como ferramenta a contabilometria. Vale não olvidar que a realização de uma auditoria não se reporta apenas à detecção de fraudes contábeis ou manipulação de contas. Também funciona como mecanismo de controle de processos internos e externos da empresa, servindo de instrumento de diagnóstico e de correção.

Como o aprimoramento da eficiência da Lei Newcomb-Benford em sua aplicação na auditoria, mediante a utilização de Testes de Hipóteses, várias são as vantagens que o auditor pode desfrutar, das quais se pode citar:

- dissolução de viés em suas conclusões sobre a verossimilhança de contas auditadas;
- otimização de tempo, pois uma discrepância detectada pela comparação das distribuições observadas (p_o) e esperadas (p_e) segundo a Lei Newcomb-Benford não sugere firmemente a existência de evasão ou fraude contábil. O erro pode ser aleatório ou decorrente do tipo de amostragem utilizada;
- o auditor consegue imprimir maior cientificidade em seus trabalhos, e assim pode emitir seu parecer com mais consistência.

Considerações finais

Este artigo apresenta, introduz e mostra que a Lei Newcomb-Benford é uma ferramenta poderosa para auxiliar o auditor a detectar erros contábeis, erros de digitação ou fraudes contábeis no contexto de uma auditoria contábil tradicional ou da auditoria digital.

A Lei Newcomb-Benford é estudada de forma simples e dedutivamente neste artigo.

A contabilometria é uma metodologia científica proposta por este artigo para a auditoria contábil nos casos em que a análise qualitativa não assegura ao auditor uma tomada de decisão precisa, conforme sugerem Nigrini, Hill, Pinkham e Raimi.

A principal estratégia contabilométrica usada é a junção dos métodos quantitativos que são a Lei Newcomb-Benford e os Testes de Hipóteses (Z – Teste e χ^2 – *Teste*) à ajuda do planejamento do trabalho do auditoria contábil tradicional e da auditoria digital. Este método é usado nos Estados Unidos por Nigrini.

Finalmente, um estudo de caso mostra que realmente os propósitos teóricos acima aludidos são alcançados. Para a empresa x, em análise, o modelo expressa que do período analisado, 1998 a 2001, os anos de 2000 e 2001 apresentam divergências do esperado, sugerindo ao auditor examinar o porquê dos desvios da Lei Newcomb-Benford.

Referências bibliográficas

Aimi, R. A. "The first digit phenomenon again", *Proceedings of the American Philosophical Society*, n. 129, 1985, pp. 211-219.

Allaart, P. C. "Na invariant-sum characterization of Benford's Law", *The Journal of Applied Probability*, n. 34, 1997, pp. 288-291.

Benford, F. "The law of anomalous numbers", *Proceeding of the American Philosophical Society*, n. 78, 1938, pp. 551-472.

Brown, G. W. "History of Rand's random digits", *Nat.Bur. Stds. App. Math Series*, Vol. 12, 1951, pp. 31-32.

Hill, T. P. "Base-Invariance implies Benford's Law", *Proceedings of the American Mathematical Society*, n. 13, 1995, pp. 887-895.

———. "A statistical derivation of the significant-digit law", *Statistical Science*, n. 4, 1996, pp. 354-363.

———. "The first digit phenomenon", *American Scientist*, n. 86, 1998, pp. 358-363.

NEWCOMB, S. "Note on the frequency of use of the different digits in natural numbers", *AJM*, n. 4, 1881, pp. 39-40.

NIGRINI, M. J. *Digital analysis using Benford's Law: Tests ? statistics for auditors*, Canada, Global Audit Publication, 2000.

NIGRINI, M. J. & LINDA, J. M. "The use of Benford's Law as an aid in Analytical Procedures. Auditing", *A Journal of Practice and Theory*, n. 16, 1997, pp. 52-67.

PINKHAN, R. S. "On the distribution of first significant digits", *Annals of Mathematical Statistics*, n. 32, 1961, pp 1.223-1.230.

RAIMI, R. "The peculiar distribution of first significant digits", *Scientific American*, Vol. 221, n. 6, 1969, pp. 109-120.

TSAO, N. R. "On the distributions of digits and roundoff errors", *Communications of the Association for Computing Machinery*, n. 17, 1974, pp. 269-271.