Research, Society and Development, v. 9, n. 8, e314985276, 2020
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i8.5276

Aplicação da lei de Newcomb-Benford no auxílio à detecção de fraudes

Application of Newcomb-Benford law to aid fraud detection

Aplicación de la ley Newcomb-Benford para ayudar a la detección de fraudes

Recebido: 04/06/2020 | Revisado: 23/06/2020 | Aceito: 24/06/2020 | Publicado: 06/07/2020

Roberto Campos Leoni
ORCID: https://orcid.org/0000-0001-6600-2963
Academia Militar das Agulhas Negras, Resende, RJ, Brasil
E-mail: rcleoni@yahoo.com.br
Graciele Silva Dos Santos
ORCID: https://orcid.org/0000-0002-2919-0675
Associação Educacional Dom Bosco, Resende, RJ, Brasil
E-mail: graciele.santos@aedb.br
Marialda Mathias Mendonça
ORCID: https://orcid.org/0000-0002-7879-3990
Centro Universitário Teresa D'Ávila, Lorena, SP, Brasil
E-mail: marialda.arquiteta@gmail.com
Nilo Antonio Souza Sampaio
ORCID: https://orcid.org/0000-0002-6168-785X
Universidade do Estado do Rio de Janeiro, Resende, RJ, Brasil
E-mail: nilo.samp@terra.com.br
José Wilson de Jesus Silva
ORCID: https://orcid.org/0000-0002-0033-2270
Centro Universitário Teresa D'Ávila, Lorena, SP, Brasil
E-mail: jwjsilva@gmail.com

Resumo

A lei de *Newcomb-Benford* estabelece que as frequências esperadas para os primeiros dígitos dos valores que compõem determinados conjuntos de dados não seguem uma distribuição uniforme. Satisfeitos os pressupostos da lei de *Newcomb-Benford*, a probabilidade de ocorrência dos dígitos de 1 a 9 é decrescente, sendo igual a 30,1% para o dígito 1, 17,6% para o dígito 2 e de apenas 4,6% para o dígito 9, por exemplo. Neste artigo, apresenta-se a lei de *Newcomb-Benford* como uma possível ferramenta a ser utilizada no campo da auditoria

(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i8.5276
contábil. O modelo contabilométrico utilizado fundamenta-se na relação entre a lei de *Newcomb-Benford* e os testes estatísticos de hipóteses: Teste-Z e o Teste- χ^2. Em uma aplicação prática, analisaram-se 589 valores de notas de empenho da Academia Militar das Agulhas Negras no período de 01 de janeiro de 2017 a 30 de junho de 2017. Concluiu-se que os dados apresentaram conformidade com o modelo proposto em relação aos dígitos líderes dos valores analisados e que o modelo contabilométrico pode ser usado por auditores na detecção de desvios contábeis.
Palavras-chave: Newcomb-Benford; Modelo contabilométrico; Teste-Z; Teste- χ^2.

Abstract

The Newcomb-Benford law states that the expected frequencies for the first digits of values that make up certain data sets do not follow a uniform distribution. Satisfied the assumptions of the Newcomb-Benford law, the probability of occurrence of digits 1 to 9 is decreasing, being equal to 30.1% for digit 1, 17.6% for digit 2 and only 4.6% for digit 9, for example. In this article, the Newcomb-Benford law is presented as a possible tool to be used in the field of accounting auditing. The accounting model used is based on relationship between the Newcomb-Benford law and the statistical hypothesis tests: Test-Z and χ^2-Test. In a practical application, 589 values of commitment grades from the Military Academy of Agulhas Negras were analyzed from January 1, 2017 to June 30, 2017. It was concluded that the data presented conformity with the proposed model in relation to leading digits of the analyzed values and that the accounting model can be used by auditors to detect accounting deviations.
Keywords: Newcomb-Benford; Accounting model; Z-test, χ^2-Test.

Resumen

La ley de Newcomb-Benford establece que las frecuencias esperadas para los primeros dígitos de los valores que componen ciertos conjuntos de datos no siguen una distribución uniforme. Satisfechos los supuestos de la ley de Newcomb-Benford, la probabilidad de ocurrencia de los dígitos 1 a 9 está disminuyendo, siendo igual a 30.1% para el dígito 1, 17.6% para el dígito 2 y solo 4.6% para el dígito 9, por ejemplo. En este artículo, la ley de Newcomb-Benford se presenta como una posible herramienta para ser utilizada en el campo de la auditoría contable. El modelo contable utilizado se basa en la relación entre la ley de Newcomb-Benford y las pruebas de hipótesis estadísticas: Test-Z y Test- χ^2. En una aplicación práctica, se analizaron 589 valores de grados de compromiso de la Academia Militar de Agulhas Negras del 1 de enero de 2017 al 30 de junio de 2017. Se concluyó que los datos presentaban conformidad con el modelo propuesto en relación con dígitos iniciales de los valores analizados y que el modelo contable puede ser utilizado por los auditores para detectar desviaciones contables.

Palabras clave: Newcomb-Benford; Modelo contable; Prueba Z; Prueba - χ^2.

1. Introdução

Fraudes contábeis são comuns tanto na iniciativa privada como nos órgãos públicos, elas ocorrem em balanços de empresas, declarações de imposto de renda, licitações e etc.. Essas fraudes causam prejuízos financeiros e diminuem a confiança de investidores, credores e órgãos reguladores, entre outros.

Na tentativa de detectar possíveis manipulações financeiras são empregadas técnicas de auditora que envolvem modelos quantitativos, cujos procedimentos, em grande parte, utilizam tópicos de inferência estatística, como, por exemplo, os testes de hipóteses.

Para Figueiredo e Moura (2001), a expressão “modelos quantitativos” é utilizada nas ciências gerenciais e descreve técnicas estatísticas e matemáticas que são usadas para a solução de problemas. Corrar e Theóphilo (2004) apresentam as principais técnicas quantitativas, bem como suas aplicações a serem utilizadas nas tomadas de decisões administrativas e contábeis, sendo elas: amostragem, análise de regressão e programação multiobjetiva. Nossa (1999) também cita algumas técnicas quantitativas, entre elas, programação dinâmica, simulação, teoria da decisão, teoria das filas e análise de séries temporais.

O objetivo deste trabalho é apresentar uma ferramenta que possibilite identificar fraude em demonstrações financeiras. Com esse intuito, empregou-se a lei de *Newcomb-Benford* para construir um modelo que sinalize para a possibilidade de fraudes em contabilidade.

Este trabalho está organizado como se segue: além da introdução, a seção 2 apresenta uma revisão da literatura com a origem da lei de *Newcomb-Benford*, um breve histórico de sua evolução e os pressupostos para sua aplicação; a seção 3 discute a metodologia utilizada nesta pesquisa; na seção 4, são apresentados os resultados obtidos por meio do Teste-Z e do Teste-χ^2 e, por fim, a seção 5 contém as considerações finais do trabalho.

A revisão da literatura apresenta o significado e as implicações da lei de *Newcomb-Benford*, explora suas origens, aplicabilidade e restrições.

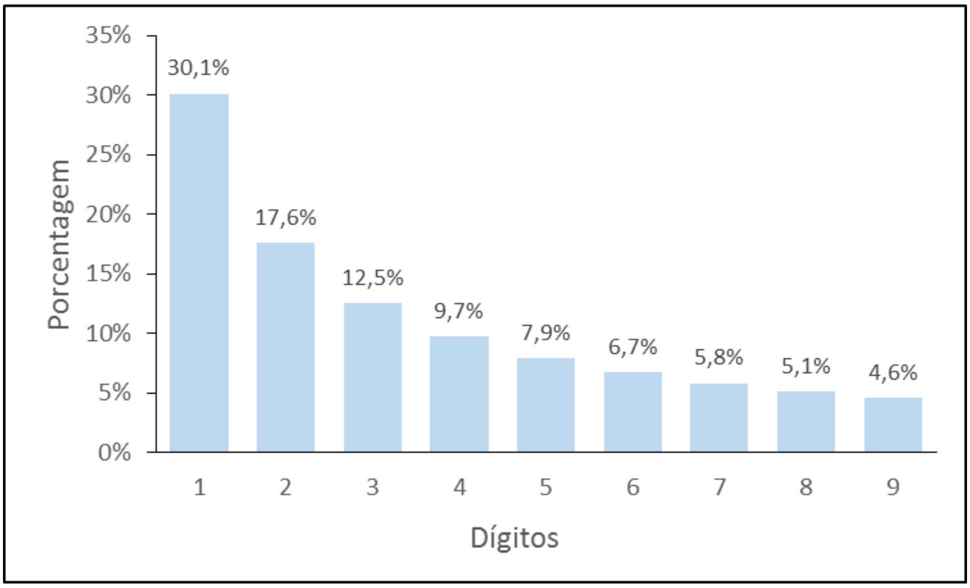
1.1 Origem e Pressupostos da Lei de Newcomb-Benford (LEI-NB)

Em 1881, o astrônomo e matemático *Simon Newcomb* (1835-1909), ao consultar tabelas de logaritmos em bibliotecas, observou que as primeiras páginas eram mais sujas e desgastadas do que as demais, inferindo, então, que elas eram as mais utilizadas. Uma vez que as primeiras páginas continham os números que começavam com dígitos menores, *Newcomb* constatou que números iniciados com o dígito "1" eram mais consultados do que os começados pelo dígito "2", seguindo uma escala decrescente de uso até o dígito "9". Ou seja, números que iniciavam com dígitos menores apareciam com mais frequência do que os com dígitos maiores.

Em 1938, o físico *Frank Benford* comprovou o mesmo fato que *Newcomb* havia observado há 57 anos. *Benford*, porém, se aprofundou mais no assunto, estudou um conjunto de dados de 20.229 observações sobre áreas de rios, números de casas de uma rua, tabelas de constantes físicas, cálculos científicos, dentre outros (CORRAR et al., 2007).

Analisando os primeiros dígitos dos números dos conjuntos de dados, *Benford* verificou uma distribuição não uniforme de suas frequências, ou seja, a ocorrência dos dígitos de 1 a 9 não seguia a probabilidade de 11,1%. Segundo *Benford*, a probabilidade de ocorrência do primeiro dígito, *d*, é dada por $\log_{10}(1 + 1/d)$. A distribuição das frequências dos primeiros dígitos é ilustrada na Figura 1.

Figura 1. Lei de *Newcomb-Benford* para os primeiros dígitos.



Fonte: autoria própria.

(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i8.5276

Devido à grande contribuição de *Benford* no desenvolvimento desse assunto, atualmente a literatura se refere à lei de *Benford*, mas fazendo justiça à *Newcomb* nos referimos à Lei *Newcomb-Benford*, ou simplesmente Lei-NB.

O uso da Lei-NB possui limitações, para que possa ser aplicada com efeito as seguintes condições devem ser verificadas:

a. Os dados devem ser numéricos;
b. O conjunto de dados não pode ser pequeno;
c. Os números não podem ser realmente fortuitos (não se aplica, por exemplo, em resultados de loterias);
d. Os dados não podem ter limite máximo ou mínimo (não se aplica, por exemplo, a altura de pessoas);
e. Os números devem ter origem em fenômenos naturais (não se aplica, por exemplo, a uma lista telefônica regional).

1.2 Breve Histórico da Evolução da Lei de *Newcomb-Benford*

Carslaw (1988) analisou demonstrações financeiras de 220 companhias da Nova Zelândia e verificou que alguns dígitos não estavam conforme o estabelecido pela Lei-NB. Fato que seria explicado por uma tendência de se arredondar para mais com finalidade de se atingir metas financeiras.

Christian e Gupta (1993) analisaram dados de contribuintes para achar indícios de evasão tributária. Ettredge e Srivastaval (1999) mostraram que se um conjunto de dados contábeis não estiver em conformidade com a distribuição de Benford não significa que há fraude. Nigrini (1999) relatou a descoberta de que em vários conjuntos de dados pode ser aplicar a Lei NB, como, por exemplo, imposto de renda, dados científicos e demográficos, mercado de ações, negócios corporativos e dados de vendas.

Santos et al. (2003) apud Corrar et al., (2011) desenvolveram o primeiro trabalho no contexto de auditoria digital no setor público brasileiro. Investigando 104.104 notas de empenho de 20 municípios paraibanos, para mapear as despesas públicas.

Durtschi, Hillison e Pacini (2004) realizaram análises em duas contas de um grande centro médico nos Estados Unidos. Os autores discutiram algumas particularidades para o emprego adequado da lei de *Benford*, tais como: a quantidade de dados empregados na análise, o poder dos testes estatísticos e os tipos de fraudes.

Diekmann (2007) aplicou a lei de *Benford* para detectar dados científicos falsificados ou fabricados, publicados em dois volumes de um periódico. O autor demonstrou que os dígitos dos coeficientes de regressão ou outras estimativas estatísticas são geralmente distribuídos de acordo com a lei de *Benford*. Nye e Moul (2007) empregaram a lei de *Benford* para avaliar a qualidade de dados macroeconômicos.

Nigrini e Miller (2007) analisaram as frequências de dígitos de grandes conjuntos de dados relacionados à hidrologia de superfície e a descoberta foi uma estreita conformidade com a lei de *Benford*. Sambridge, Tkalčić e Jackson (2010) analisaram 15 conjuntos de observações tiradas dos campos da física, astronomia, geofísica, química, engenharia e matemática, e mostram que a lei de *Benford* é válida para todos eles.

2. Metodologia

Para atingir os objetivos do estudo, foi elaborado um modelo com base no teste do primeiro dígito utilizando 589 valores obtidos do Portal da Transparência do Governo Federal, referentes a pagamentos de despesas de material de consumo da Academia Militar das Agulhas Negras, no período de 01 de janeiro de 2017 a 30 de junho de 2017. O modelo contabilométrico utilizado se baseia na relação entre a lei de *Newcomb-Benford* e os testes de hipóteses, Teste-*Z* e Teste-χ^2 (teste qui-quadrado), são os mesmos adotados por Santos et al. (2003) e Nigrini (2000).

2.1 O Teste-*Z*

O Teste-*Z* é usado para mensurar o grau de significância entre as diferenças da distribuição de probabilidade observada e a distribuição de probabilidade esperada, $p_o - p_e$, de cada um dos dígitos analisados, de maneira a se definir as hipóteses a serem testadas (Stevenson, 2001). Definimos a hipótese nula H_0 e hipótese alternativa H_1 do seguinte modo:

$H_0 \Rightarrow$ Não existe diferença estatisticamente significativa entre distribuições de probabilidades observadas (p_o) e esperadas (p_e), ou seja, $p_o = p_e$;
$H_1 \Rightarrow$ Existe diferença estatisticamente significativa entre distribuições de probabilidades observadas (p_o) e esperadas (p_e), ou seja, $p_o \neq p_e$.

Assumindo que a distribuição amostral siga o modelo normal de probabilidade, e considerando um nível de significância de $\alpha = 0{,}05$, tem-se um valor crítico de $z_{0,05} = 1{,}96$.

Esse é o valor a ser comparado com o módulo de z na análise pontual de cada dígito. Se $|z| > z_{0,05}$ rejeita-se H_0, sendo, assim, viável que se proceda uma análise criteriosa para cada um dos casos em que tal situação foi observada.

Para verificar o nível de significância entre as diferenças $p_o = p_e$ utiliza-se o Teste-Z:

$$z = \frac{|p_o - p_e| - \frac{1}{2n}}{\sqrt{\frac{p_e(1-p_e)}{n}}} \qquad (1)$$

sendo n o número de observações; $1/2n$ é o termo de correção de continuidade, utilizado quando $\frac{1}{2n} < |p_o - p_e|$.

2.2 O Teste-χ^2

Para verificar se a distribuição de probabilidade observada (p_0) é equivalente a distribuição esperada (pe), segundo a lei de *Newcomb-Benford*, utiliza-se o Teste- χ^2.

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

sendo: *PO* (proporções observadas) = (*po*) × (nº da população), e *PE* (proporções esperadas) = (*pe*) × (nº da população).

3. Resultados e Discussão

Aplicou-se o Teste-*Z* de uma amostra para a proporção da população composta pelos valores pagos em material de consumo da Academia Militar das Agulhas Negras no período de 01 de janeiro de 2017 a 30 de junho de 2017. O Teste-*Z* possibilita realizar a inferência para cada dígito. O modelo contabilométrico testa a afirmação de que a ocorrência dos primeiros dígitos obedece aos pressupostos da probabilidade apresentados pela Lei-NB, de acordo com a Tabela 1.

Tabela 1. Probabilidade de ocorrência dos primeiros dígitos segundo a Lei-NB.

Dígitos	1	2	3	4	5	6	7	8	9
Probabilidade	0,301	0,176	0,125	0,097	0,079	0,067	0,058	0,052	0,046

Fonte: Autoria própria.

As hipóteses anunciadas para a ocorrência de cada dígito são apresentadas na Tabela 2.

Tabela 2. Hipóteses anunciadas para o Teste-*Z*.

Dígitos	1	2	⋯	8	9
Hipóteses	$H_0: pe_1 = 0{,}301$ $H_1: pe_1 \neq 0{,}301$	$H_0: pe_2 = 0{,}176$ $H_1: pe_2 \neq 0{,}176$	⋯	$H_0: pe_8 = 0{,}051$ $H_1: pe_8 \neq 0{,}051$	$H_0: pe_9 = 0{,}046$ $H_1: pe_9 \neq 0{,}046$

Nota: pe_i =proporção esperada para o dígito *i* (*i*=1, 2, ... ,9).
Fonte: Autoria própria.

A partir dos dados das notas de empenho, aplicamos a equação 1. Os resultados são apresentados na Tabela 3.

Tabela 3. Resultado dos testes realizados com nível de significância α=5%.

Dígitos	Proporção Esperada Lei-NB(p_e)	Proporção observada (p_o)	Desvio $\lvert p_o - p_e \rvert$	Valor do Teste *Z*	Valor *p*
1	0,301	0,309	0,008	0,377	0,7062
2	0,176	0,194	0,018	1,058	0,2900
3	0,125	0,104	0,021	1,506	0,1320
4	0,097	0,107	0,010	0,755	0,4502
5	0,079	0,083	0,004	0,284	0,7764
6	0,067	0,065	0,002	0,154	0,8776
7	0,058	0,056	0,002	0,116	0,9076
8	0,051	0,032	0,019	**1,988**	**0,0468**
9	0,046	0,051	0,005	0,503	0,6150

Fonte: autoria própria.

A Tabela 3 apresenta o Valor *p* para cada primeiro dígito. Verifica-se que somente o Valor *p* = 0,0468 para o teste de hipóteses do dígito 8 é menor que o nível de significância de α = 0,05. Podemos constatar o mesmo fato observando que o valor do Teste- *Z* =1,988 para o dígito 8 é o único maior que $z_{0,05} = 1{,}96$. Logo, segundo o Teste-*Z*, há evidência do ponto de vista do modelo contabilométrico, que o dígito 8 não atende à Lei-NB. Esse resultado não significa que houve fraude, apenas auxilia auditores no processo de avaliação de possíveis fraudes.

O Teste-χ^2, por sua vez, fornece informações para uma análise geral de todos os dígitos. Se ocorrem diferenças nas proporções entre os grupos de dígitos, o Teste-χ^2 pode ser utilizado com o objetivo de validar a magnitude dentro de um determinado nível de significância (□). Para avaliar se o conjunto de dígitos atende ao disposto na Lei-NB, testaram-se as hipóteses:

$H_0: pe_1 = po_1, \ldots, pe_9 = po_9$

H_1: Pelo menos uma das proporções não é igual ao valor alegado.

sendo: pe_i=proporção esperada para o dígito *i* e po_i=proporção observada para o dígito *i* (*i*=1, 2, ... ,9)

Para calcular a estatística do Teste-χ^2, definiu-se primeiro a tabela de contingência (ver Tabela 4).

Tabela 4. Tabela de contingência.

Dígitos	1	2	3	4	5	6	7	8	9
Proporção esperada	177	104	74	57	47	39	34	30	27
Proporção observada	182	114	61	63	49	38	33	19	30

Fonte: Autoria própria.

Usando-se a equação 2, obtém-se a estatística de teste χ^2. A Tabela 5 apresenta os cálculos dos componentes da estatística teste χ^2 para os dígitos líderes de 1 a 9. Sendo a estatística de teste χ^2=8,5250, e o grau de liberdade igual a 8, o Valor *p* é igual a 0,3839, ou seja, maior que o nível de significância 0,05; desse modo não há evidência suficiente para rejeitar a hipótese nula (H_0).

Pode-se verificar que a estatística de teste χ^2 =8,5250 não pertence a região de rejeição limitada pelo valor crítico 15,507. Portanto, há evidência estatística para concluir que a distribuição das proporções, por dígitos, atende à Lei-NB. Infere-se que as Notas de Empenho da AMAN, no período analisado, atendem no todo à lei natural de distribuição.

Tabela 5. Resultado do Teste-χ^2.

PO	PE	$(PO-PE)$	$(PO-PE)^2$	$(PO-PE)^2/PE$
182	177	5	25	0,1412
114	104	10	100	0,9615
61	74	-13	169	2,2838
63	57	6	36	0,6316
49	47	2	4	0,0851
38	39	-1	1	0,0256
33	34	-1	1	0,0294
19	30	-11	121	4,0333
30	27	3	9	0,3333
Soma				8,5250

Fonte: Autoria própria.

4. Considerações Finais

A análise global realizada por meio do Teste-□^{2} apontou conformidade com a Lei NB em relação aos primeiros dígitos das notas de empenho referente ao material de consumo da Academia Militar das Agulhas Negras, no período de 01 de janeiro de 2017 a 30 de junho de 2017. De uma forma geral os valores dessa população não apresentam dissonância com os esperados pela Lei NB.

Utilizando-se o Teste-*Z*, houve conformidade com a Lei NB em relação aos primeiros dígitos das notas de empenho, com exceção do dígito 8, porém, a estatística de teste apresentou valor próximo ao ponto crítico, logo, se fosse adotado um valor mais rigoroso para o valor de α (por exemplo: α =0,01) tal conclusão não seria mais a mesma para o dígito 8.

Não se pode afirmar que há fraude, baseando-se apenas na Lei NB, mas devido ao fato de ser uma ferramenta de fácil implementação e não exigir grande investimento, ela pode ser usada para indicar quais notas de empenho deveriam ser verificadas mais detalhadamente.

O modelo aqui apresentado não substitui a técnica de auditoria utilizada por auditores contábeis, entretanto é uma proposta que pode apoiar as atividades de auditoria. Como estudos futuros, propõe-se a análise da distribuição dos outros dígitos.

Referências

Carslaw, C. A. P. (1988). Anomalies in income numbers: evidence of goal-oriented behavior, The Accounting Review, 63(2), 321-327.

Christian, C. W., Gupta, S. (1993). New evidence on "secondary evasion". The Journal of the American Taxation Association, American Accounting Association, 15(1), 72.

Corrar, L., Paulo, E., Dias Filho, J. M., & Rodrigues, A. (2011). Análise multivariada para os cursos de administração, ciências contábeis e economia. São Paulo: Atlas.

Corrar, L., & Theóphilo, C. R. (2004). Pesquisa operacional para decisão em contabilidade e administração: contabilometria. São Paulo: Atlas.

Diekmann, A. (2007). Not the First Digit! Using Benford's Law to Detect Fraudulent Scientific Data, Journal of Applied Statistics, 34:3, 321-329, DOI: 10.1080/02664760601004940

Durtschi, C., Hillison, W., Pacini, C. (2004). The Effective Use of Benford's Law to Assist in Detecting Fraud in Accounting Data. Journal of Forensic Accounting, v. 99, n. 99, p. 17–34.

Ettredge, M. L., Srivastaval, R. P. (1999). Using digital analysis to enhance data integrity. Issues in Accounting Education, 14(4), 354–363.

Figueiredo, S., Moura, H. (2001). A utilização dos Métodos Quantitativos pela Contabilidade. Revista Brasileira de Contabilidade, Brasília, ano 30, 127, 51-61.

Nigrini, M. J. (1999). I've got your number. Journal of Accountancy, 187(5), 79–83.

Nigrini, M. J., Miller, S. J. (2007). Benford's Law Applied to Hydrology Data – Results and Revelance to Other Geophysical Data. Mathematical Geology, 39(5), 469–490.

Nossa, V. (1999). A necessidade de professores qualificados e atualizados para o ensino da Contabilidade. Revista de Contabilidade do CRC-SP, n. 9.

(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v9i8.5276

Nye, J., Moul, C. (2007). The Political Economy of Numbers: On the Application of Benford's Law to International Macroeconomic Statistics. The B.E. Journal of Macroeconomics, 7(1).

Sambridge, M., Tkalčić, H., Jackson, A. (2010). Benford's law in the natural sciences. Geophysical Research Letters, 379(22), 1–5.

Santos, J., Diniz, J. A., Filho, J. F. R. (2003). A lei de newcomb-benford: uma aplicação para determinar o dna-equivalente das despesas no setor público. Anais do Seminário USP de Contabilidade e Controladoria, v. 3.

Porcentagem de contribuição de cada autor no manuscrito

Roberto Campos Leoni - 20%

Graciele Silva Dos Santos – 20%

Marialda Mathias Mendonça - 20%

Nilo Antonio Souza Sampaio – 20%

José Wilson de Jesus Silva – 20%