

Lei de Benford e Auditoria de Obras Públicas: uma análise de sobrepreço na reforma do Maracanã



Flávia Ceccato Rodrigues da Cunha

é servidora do Tribunal de Contas da União,

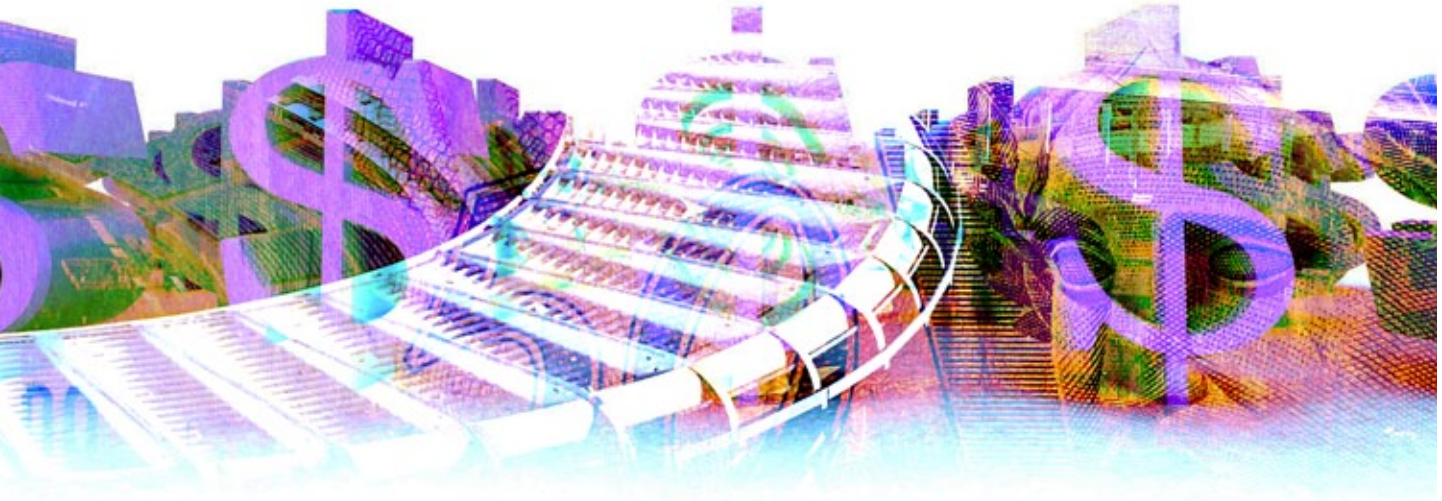


Maurício Soares Bugarin

é professor titular do Departamento de Economia da Universidade de Brasília (UnB) e líder do grupo de pesquisa do CNPq *Economics and Politics Research Group*.

RESUMO

As análises de preços nas auditorias de obras públicas ocupam algumas vezes semanas de trabalho do auditor, pois, em muitos casos, as planilhas orçamentárias são extensas e de difícil análise. A Lei Newcomb-Benford é uma ferramenta de mineração de dados para atuar em parceria com a Curva ABC e contribuir com uma seleção mais eficiente dos serviços das planilhas para análise de sobrepreço. Essa Lei propõe que as frequências dos primeiros dígitos em uma infinidade de bancos de dados são decrescentes do 1 ao 9; o dígito 1 aparece em, aproximadamente, 30% dos dados, enquanto o 9 não atinge 5% desses valores. O presente trabalho visa demonstrar a aplicação da Lei Newcomb-Benford às planilhas de preços de obras públicas. A metodologia consistiu: na introdução à Lei Newcomb-Benford e suas principais aplicações; na apresentação dos testes pertinentes da Lei em tela; no desenvolvimento e aplicação desses testes à obra de reforma do Maracanã; no confronto dos resultados obtidos com a análise de preços realizada pelo Tribunal de Contas da União (TCU); e na apresentação dos resultados, conclusões e próximas pesquisas. Constatou-se, de uma maneira geral, que essa obra apresentou uma conformidade marginal com a Lei de Benford. No entanto, os testes permitiram identificar 17 serviços da planilha que não respeitaram a Lei e que representaram 71,54% do sobrepreço total apurado pelo TCU (R\$ 149.972.318,01).



Palavras-chave: Auditoria. Obras públicas. Lei Newcomb-Benford. Sobrepreço.

1. INTRODUÇÃO

Se jogarmos um dado ao acaso, a chance de tirarmos o número 5, por exemplo, é de $1/6$, uma vez que o dado possui seis lados. Ao se jogar uma moeda não viesada, há 50% de chance de se tirar cara ou coroa. Dessa maneira, tendemos a pensar que, em um banco de dados numéricos, a probabilidade de se escolher aleatoriamente um valor e ele possuir como primeiro dígito o número 1 é de $1/9$, o mesmo se aplicando a qualquer outro valor de 2 a 9.

Simon Newcomb (1881), um astrônomo e matemático do século XIX, observou que as primeiras páginas das tábuas de logaritmos se apresentavam mais desgastadas do que as últimas, indicando que o valor usualmente mais acessado era o 1, e que a frequência diminuía até o 9, contrariando o entendimento comum de uma distribuição uniforme da frequência desses dígitos. Como Newcomb não reuniu dados numéricos ou forneceu qualquer outra evidência de sua descoberta, o fato só começou a ganhar importância mais de meio século depois, quando o físico Frank Benford (1938) incidentalmente chegou à mesma conclusão. Benford publicou artigo seminal em 1938, denominado *The Law of Anomalous Numbers*, em que utilizou dados coletados de diferentes tipos de fontes. Esses dados

eram aleatórios e não possuíam nenhuma relação entre si, e variavam desde números obtidos nas páginas principais dos jornais e todos os números de um tópico importante do *Reader's Digest* até tabelas matemáticas e constantes científicas. Seu trabalho analisou os primeiros dígitos dos dados coletados e mostrou que 30,6% dos números possuíam 1 como primeiro dígito; o primeiro dígito 2 ocorria em 18,5% dos casos; e que, em contraste, somente 4,7% dos números possuíam como primeiro dígito o número 9. Essas frequências dos primeiros dígitos se aplicam a uma variedade de fontes de dados, incluindo contas de energia, endereços, preços de ações, valores populacionais, taxas de mortalidade, entre outras. Essa distribuição é hoje conhecida como **distribuição de Benford**. Já a propriedade descoberta por Newcomb e por Benford é conhecida como **Lei Newcomb-Benford** ou, mais simplesmente, **Lei de Benford**.

Para melhor entender o porquê das diferenças de frequência, suponha que você invista 10000 reais em um fundo de pensão que lhe ofereça um retorno prefixado de 7% ao ano. Então seu investimento dobrará mais ou menos a cada dez anos. Portanto, após dez anos tendo o 1 como o primeiro dígito, o montante de seu investimento finalmente chegará a 20.000. Passados mais 10 anos, o recurso dobrará para 40.000 (em uma maior parte desses dez anos os números começarão com 2, e na outra menor parte começarão com 3). Após outra década, o montante chegará a 80.000

(obter-se-ão os valores 4, 5, 6 e 7 como os primeiros dígitos em apenas dez anos). Em um dado momento, chegar-se-á ao valor de 100.000, com o primeiro dígito 1 incidindo por mais dez anos. Dessa maneira, ao se escolher uma data aleatória, é mais provável que o montante de seu investimento desse dia possua o 1 como primeiro dígito, do que qualquer outro dígito. Essa mesma lógica se aplica a diversos dados presentes na natureza, tal como o tamanho de populações, dentre outros.

Um banco de dados tem mais chance de representar uma distribuição de Benford se os dados forem coletados de diferentes distribuições (Hill, 1995). Por outro lado, números atribuídos pela intervenção do homem, tais como números da Seguridade Social, códigos postais, contas bancárias, números telefônicos ou números fabricados por estudantes em experimentos geralmente não se conformam com a Lei de Benford (Nigrini, 2000). Essa observação sugere que a Lei dos Números Anômalos pode ser usada para se detectar indícios de manipulação humana de dados.

Diversos estudos foram realizados adotando a hipótese de que dados fabricados são identificados mediante o desvio dos dígitos em relação à distribuição de Benford. Nigrini (1992, 2000, 2012), assumindo que dados contábeis verdadeiros seguem a distribuição de Benford bem de perto (como sua pesquisa indicou que seguiam), argumentou que desvios substanciais em relação a essa Lei sugeririam possíveis fraudes ou dados fabricados. O autor desenvolveu vários testes para mensurar a conformidade com a Lei de Benford, e o *Wall Street Journal* (Berton, 1995) noticiou que o escritório da Procuradoria do Brooklyn, em Nova York, detectou fraudes em sete companhias de Nova York usando esses testes. Como evidência, descobriu-se, nesse caso, que dados fraudulentos e aleatórios possuíam poucos valores começando com 1 e muitos números começando com 6. Com base nesses sucessos anteriores, Nigrini foi chamado a dar consultoria a órgãos de arrecadação tributária de diversos países e a instalar os testes da Lei de Benford na maioria dos programas computacionais de detecção de fraude.

Rauch Götttsche, Brähler e Engel (2011) publicaram artigo na *German Economic Review*, no qual demonstraram que a Lei de Benford poderia ser utilizada para testar dados macroeconômicos, revelando quais deles necessitavam de uma inspeção mais rigorosa. Eles analisaram a conformidade com a Lei de Benford dos primeiros dígitos de dados macroeconômicos reportados ao Gabinete de Estatísticas da União Europeia

- Eurostat pelos países membros da UE. Construíram um *ranking* dos 27 países membros de acordo com a extensão do desvio encontrado. O país que teve o maior desvio foi a Grécia, cuja manipulação dos dados havia sido oficialmente confirmada pela Comissão Europeia (2010).

Walter Mebane, um estatístico americano da Universidade de Michigan, analisou os dados das eleições iranianas em 2009 e encontrou anomalias que indicavam fortemente a ocorrência de fraude na vitória de Ahmadinejad (Mebane, 2009). Mebane, anteriormente, havia estudado dados eleitorais de vários países, incluindo os Estados Unidos, Rússia e México. Em 2006, ele descobriu que a contagem dos votos tendia a seguir a Lei de Benford no segundo dígito (Mebane, 2006). Ao testar os dados iranianos referentes a 2009, Mebane verificou que, nas cidades com poucos votos inválidos, os números de Ahmadinejad passavam longe da distribuição de Benford e que o candidato, nessas situações, possuía uma grande vantagem nos votos.

O trabalho em apreço visa apresentar um estudo de caso da aplicação da Lei NB às obras públicas, utilizando a planilha orçamentária da obra de reforma do Maracanã. Essa obra foi selecionada por apresentar um volume de dados relevante e pela possibilidade de comparação dos resultados dos testes com a análise de preços empreendida pelo TCU. Serão apresentados primeiramente os Testes da Lei de Benford pertinentes. A seguir serão aplicados tais testes àquela planilha, realizando-se o confronto dos resultados com a análise de preços do TCU, e expostas as conclusões.

2. TESTES DA LEI DE BENFORD BASEADOS NAS PROBABILIDADES DOS DÍGITOS

Foram aplicados no presente trabalho os testes descritos a seguir, caracterizados segundo Nigrini (2012).

2.1 TESTE DOS DOIS PRIMEIROS DÍGITOS

A frequência esperada da ocorrência de um número $D_2=d_2$ como segundo dígito em um conjunto de valores, dado que o primeiro dígito é $D_1=d_1$, segundo a Lei NB é dada por:

$$\text{Prob}(D_1D_2 = I_1d_2) = \log\left(1 + \frac{1}{d_1d_2}\right) \quad (I)$$

Sendo: D_1D_2 = Dois primeiros dígitos e $d_1d_2 \in \{10, 11, \dots, 99\}$

2.2 TESTE DA SOMA

O Teste da Soma é um Teste Avançado desenvolvido por Nigrini (2012), 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, ou seja, 1/90.

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.

Quando se confronta esse Teste com o Teste dos Dois Primeiros Dígitos, confronta-se o volume de dados contidos em cada grupo com a soma de seus valores. Assim, no caso de orçamento de obras públicas, verifica-se a materialidade e a relevância de cada grupo, de forma a selecionar os dígitos que merecem um olhar crítico mais minucioso por parte do auditor.

3. ANÁLISE DA OBRA DE REFORMA DO MARACANÃ

A análise deste trabalho incidiu sobre o orçamento da reforma do Maracanã inicialmente entregue ao TCU, no valor de R\$ 931.885.382,19, uma vez que foram apresentados outros orçamentos a posteriori, que praticamente eliminaram o sobrepreço da maioria dos itens da planilha. Selecionou-se o orçamento inicial para que a análise dos dados ocorresse da forma mais efetiva possível, abrangendo todo o sobrepreço apontado pela unidade técnica do TCU inicialmente. O trabalho em tela contemplou apenas os custos unitários, mas também poderiam ser testados os quantitativos dos serviços e os preços totais. Examinaram-se 828 itens, porque foram excluídos os valores inferiores a R\$ 10,00, os quais não possuíam segundo dígito.

3.1 TESTE DOS DOIS PRIMEIROS DÍGITOS

Paramais detalhamento da análise, de modo a reduzir o tamanho da amostra de dígitos críticos, procedeu-se ao exame dos dois primeiros dígitos. Os resultados encontram-se reportados na Tabela 1, em que: “Díg.” se refere aos dois primeiros dígitos dos valores; “C.” 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 NB; “Dif.” é a diferença entre

“Real” e “LB”; “Teste Z” se refere aos resultados do Teste Z para mensurar a conformidade com a Lei NB; “QQ.” mostra os resultados do Teste Qui-Quadrado; e “MDA” traz o resultado da Média dos Desvios Absolutos.

Segundo a Tabela 1, houve picos mais intensos nos dígitos 11, 16, 25, 28 e 42 em relação às proporções da curva descendente da Lei NB. Os resultados do Teste Z foram: 11 (2,954), 16 (2,105), 25 (2,524), 28 (2,303) e 42 (2,060).

Destarte, como apenas cinco dentre os 90 dígitos superaram o limite de 1,96, pode-se afirmar que as proporções dos primeiros dígitos dos custos unitários do Maracanã, de um modo geral, não destoaram da Lei NB para esse teste se for seguida a sugestão de Nigrini (2012), que considera aceitável a ocorrência de até cinco picos nesse teste.

O resultado do Teste Qui-Quadrado foi 106,648. O valor crítico para 89 graus de liberdade e 0,05 de significância é 112,02. Assim, como o valor calculado no teste não excedeu o valor crítico, não se pode rejeitar a hipótese nula, sugerindo conformidade com a Lei NB.

O último teste aplicado foi o MDA. O valor encontrado para o Maracanã foi 0,0031, o qual estava acima de 0,0022 (valor limítrofe entre conformidade e não conformidade adotado por Nigrini 2012), fato que o inseriu na faixa de não conformidade dos valores de referência.

Considerando todos os testes aplicados, pode-se afirmar que a planilha orçamentária do Maracanã não passou no Teste da Média dos Desvios Absolutos para os dois primeiros dígitos dos custos unitários, mas apresentou resultado satisfatório no Teste Qui-Quadrado e no Teste Z (Tabela 1).



Tabela 1:

Teste dos Dois Primeiros
Dígitos para os custos
unitários do Maracanã

Díg.	C	Real	LB	Dif.	Teste-Z	QQ	MDA	Díg.	C	Real	LB	Dif.	Teste-Z	QQ	MDA	
10	43	0,052	0,041	0,011	1,435	2,222	0,011	55	7	0,008	0,008	0,001	0,008	0,042	0,001	
11	48	0,058	0,038	0,020	2,954	8,925	0,020	56	11	0,013	0,008	0,006	1,645	3,376	0,006	
12	28	0,034	0,035	-0,001	0,054	0,021	0,001	57	10	0,012	0,008	0,005	1,303	2,244	0,005	
13	26	0,031	0,032	-0,001	0,029	0,016	0,001	58	8	0,010	0,007	0,002	0,548	0,559	0,002	
14	28	0,034	0,030	0,004	0,548	0,410	0,004	59	2	0,002	0,007	-0,005	1,447	2,706	0,005	
15	31	0,037	0,028	0,009	1,535	2,616	0,009	60	5	0,006	0,007	-0,001	0,183	0,150	0,001	
16	32	0,039	0,026	0,012	2,105	4,772	0,012	61	6	0,007	0,007	0,000	0,063	0,004	0,000	
17	26	0,031	0,025	0,007	1,105	1,443	0,007	62	1	0,001	0,007	-0,006	1,780	3,927	0,006	
18	18	0,022	0,023	-0,002	0,216	0,107	0,002	63	4	0,005	0,007	-0,002	0,490	0,488	0,002	
19	16	0,019	0,022	-0,003	0,458	0,324	0,003	64	6	0,007	0,007	0,001	0,181	0,032	0,001	
20	18	0,022	0,021	0,001	0,110	0,012	0,001	65	2	0,002	0,007	-0,004	1,280	2,219	0,004	
21	14	0,017	0,020	-0,003	0,550	0,445	0,003	66	5	0,006	0,007	0,000	0,176	0,031	0,000	
22	12	0,014	0,019	-0,005	0,880	0,993	0,005	67	3	0,004	0,006	-0,003	0,794	1,017	0,003	
23	10	0,012	0,018	-0,006	1,240	1,838	0,006	68	1	0,001	0,006	-0,005	1,642	3,440	0,005	
24	12	0,014	0,018	-0,003	0,574	0,489	0,003	69	7	0,008	0,006	0,002	0,585	0,644	0,002	
25	24	0,029	0,017	0,012	2,524	6,944	0,012	70	6	0,007	0,006	0,001	0,177	0,159	0,001	
26	7	0,008	0,016	-0,008	1,662	3,182	0,008	71	4	0,005	0,006	-0,001	0,237	0,211	0,001	
27	10	0,012	0,016	-0,004	0,718	0,724	0,004	72	5	0,006	0,006	0,000	0,018	0,000	0,000	
28	4	0,005	0,015	-0,010	2,303	5,887	0,010	73	7	0,008	0,006	0,003	0,729	0,908	0,003	
29	7	0,008	0,015	-0,006	1,353	2,210	0,006	74	4	0,005	0,006	-0,001	0,149	0,142	0,001	
30	13	0,016	0,014	0,001	0,208	0,124	0,001	75	4	0,005	0,006	-0,001	0,121	0,122	0,001	
31	9	0,011	0,014	-0,003	0,571	0,512	0,003	76	6	0,007	0,006	0,002	0,370	0,359	0,002	
32	13	0,016	0,013	0,002	0,434	0,338	0,002	77	2	0,002	0,006	-0,003	0,996	1,502	0,003	
33	15	0,018	0,013	0,005	1,157	1,694	0,005	78	3	0,004	0,006	-0,002	0,506	0,546	0,002	
34	15	0,018	0,013	0,006	1,271	2,009	0,006	79	5	0,006	0,005	0,001	0,225	0,050	0,001	
35	8	0,010	0,012	-0,003	0,515	0,448	0,003	80	3	0,004	0,005	-0,002	0,459	0,482	0,002	
36	13	0,016	0,012	0,004	0,848	1,005	0,004	81	7	0,008	0,005	0,003	0,997	1,518	0,003	
37	10	0,012	0,012	0,000	0,133	0,018	0,000	82	5	0,006	0,005	0,001	0,068	0,094	0,001	
38	9	0,011	0,011	0,000	0,112	0,012	0,000	83	2	0,002	0,005	-0,003	0,873	1,235	0,003	
39	10	0,012	0,011	0,001	0,132	0,088	0,001	84	4	0,005	0,005	0,000	0,124	0,015	0,000	
40	12	0,014	0,011	0,004	0,884	1,097	0,004	85	5	0,006	0,005	0,001	0,144	0,150	0,001	
41	10	0,012	0,010	0,002	0,285	0,206	0,002	86	5	0,006	0,005	0,001	0,169	0,171	0,001	
42	2	0,002	0,010	-0,008	2,060	4,934	0,008	87	3	0,004	0,005	-0,001	0,302	0,300	0,001	
43	7	0,008	0,010	-0,002	0,268	0,194	0,002	88	4	0,005	0,005	0,000	0,031	0,001	0,000	
44	8	0,010	0,010	0,000	0,029	0,001	0,000	89	2	0,002	0,005	-0,002	0,759	1,013	0,002	
45	7	0,008	0,010	-0,001	0,144	0,103	0,001	90	7	0,008	0,005	0,004	1,271	2,305	0,004	
46	4	0,005	0,009	-0,005	1,168	1,802	0,005	91	2	0,002	0,005	-0,002	0,723	0,948	0,002	
47	4	0,005	0,009	-0,004	1,121	1,684	0,004	92	1	0,001	0,005	-0,003	1,214	2,145	0,003	
48	6	0,007	0,009	-0,002	0,337	0,270	0,002	93	4	0,005	0,005	0,000	0,079	0,006	0,000	
49	6	0,007	0,009	-0,002	0,285	0,220	0,002	94	2	0,002	0,005	-0,002	0,671	0,856	0,002	
50	9	0,011	0,009	0,002	0,519	0,496	0,002	95	4	0,005	0,005	0,000	0,121	0,015	0,000	
51	8	0,010	0,008	0,001	0,197	0,148	0,001	96	3	0,004	0,005	-0,001	0,118	0,142	0,001	
52	12	0,014	0,008	0,006	1,784	3,873	0,006	97	1	0,001	0,004	-0,003	1,142	1,959	0,003	
53	6	0,007	0,008	-0,001	0,086	0,077	0,001	98	1	0,001	0,004	-0,003	1,128	1,925	0,003	
54	4	0,005	0,008	-0,003	0,820	1,023	0,003	99	5	0,006	0,004	0,002	0,467	0,531	0,002	
N	828												χ^2 Real	106,65	MDA	
													χ^2 Limite	112,02		0,0032

3.2 TESTE DA SOMA

Para uma boa seleção dos dígitos críticos, realizou-se, de forma complementar, o Teste da Soma, o qual verificou a materialidade de cada dupla de dígitos na planilha orçamentária. Os testes anteriores verifica-

ram o número de repetições em relação ao padrão da Lei NB sem, no entanto, examinarem a magnitude dos serviços que correspondiam a esses dígitos. Os resultados estão ilustrados na Tabela 2 a seguir.

Na Tabela anterior, a 1ª e 6ª colunas se referem aos dois primeiros dígitos dos valores; a 2ª e 7ª colunas corres-

Tabela 2: Teste da Soma para os custos unitários do Maracanã

Dígito	Soma	Real	Benford	Diferença	Dígito	Soma	Real	Benford	Diferença
10	1.117.783,14	0,002	0,011	-0,009	55	1.393,43	0,000	0,011	-0,011
11	35.228.545,85	0,060	0,011	0,049	56	11.333.846,10	0,019	0,011	0,008
12	9.996,90	0,000	0,011	-0,011	57	9.859,91	0,000	0,011	-0,011
13	1.427.480,57	0,002	0,011	-0,009	58	3.104,83	0,000	0,011	-0,011
14	149.926,09	0,000	0,011	-0,011	59	1.184,83	0,000	0,011	-0,011
15	7.327,10	0,000	0,011	-0,011	60	7.413,72	0,000	0,011	-0,011
16	12.400,03	0,000	0,011	-0,011	61	7.613,46	0,000	0,011	-0,011
17	19.334.196,90	0,033	0,011	0,022	62	626,58	0,000	0,011	-0,011
18	18.810.868,98	0,032	0,011	0,021	63	636.977,00	0,001	0,011	-0,010
19	22.983.744,00	0,039	0,011	0,028	64	2.706,05	0,000	0,011	-0,011
20	20.216.982,58	0,035	0,011	0,024	65	723,09	0,000	0,011	-0,011
21	23.621.379,30	0,040	0,011	0,029	66	8.734,92	0,000	0,011	-0,011
22	22.910.130,11	0,039	0,011	0,028	67	2.028,35	0,000	0,011	-0,011
23	3.365,19	0,000	0,011	-0,011	68	68,92	0,000	0,011	-0,011
24	2.441.496,65	0,004	0,011	-0,007	69	6.930.295,76	0,012	0,011	0,001
25	282.240.352,41	0,483	0,011	0,472	70	2.955,46	0,000	0,011	-0,011
26	3.516,32	0,000	0,011	-0,011	71	1.568,94	0,000	0,011	-0,011
27	3.699,11	0,000	0,011	-0,011	72	2.322,08	0,000	0,011	-0,011
28	629,84	0,000	0,011	-0,011	73	4.482,65	0,000	0,011	-0,011
29	9.100,20	0,000	0,011	-0,011	74	1.632,72	0,000	0,011	-0,011
30	3.041.197,41	0,005	0,011	-0,006	75	3.028,68	0,000	0,011	-0,011
31	1.124,59	0,000	0,011	-0,011	76	3.899,28	0,000	0,011	-0,011
32	32.871.298,35	0,056	0,011	0,045	77	7.708.500,52	0,013	0,011	0,002
33	40.266,66	0,000	0,011	-0,011	78	945,72	0,000	0,011	-0,011
34	3.453.111,19	0,006	0,011	-0,005	79	3.254,91	0,000	0,011	-0,011
35	8.277,56	0,000	0,011	-0,011	80	241,03	0,000	0,011	-0,011
36	14.275,68	0,000	0,011	-0,011	81	3.495,51	0,000	0,011	-0,011
37	5.748,38	0,000	0,011	-0,011	82	1.892,95	0,000	0,011	-0,011
38	40.635,97	0,000	0,011	-0,011	83	919,55	0,000	0,011	-0,011
39	2.169,59	0,000	0,011	-0,011	84	3.373,90	0,000	0,011	-0,011
40	6.314,11	0,000	0,011	-0,011	85	4.275,98	0,000	0,011	-0,011
41	85.478,77	0,000	0,011	-0,011	86	4.330,92	0,000	0,011	-0,011
42	8.514,19	0,000	0,011	-0,011	87	1.054,65	0,000	0,011	-0,011
43	6.138,93	0,000	0,011	-0,011	88	353,98	0,000	0,011	-0,011
44	4.472.961,24	0,008	0,011	-0,003	89	1.789,15	0,000	0,011	-0,011
45	4.557.697,21	0,008	0,011	-0,003	90	4.705,38	0,000	0,011	-0,011
46	605,18	0,000	0,011	-0,011	91	182,34	0,000	0,011	-0,011
47	5.356,92	0,000	0,011	-0,011	92	924,80	0,000	0,011	-0,011
48	48.849.864,06	0,084	0,011	0,073	93	2.895,08	0,000	0,011	-0,011
49	7.018,04	0,000	0,011	-0,011	94	1.034,26	0,000	0,011	-0,011
50	2.718,91	0,000	0,011	-0,011	95	11.507,03	0,000	0,011	-0,011
51	1.805,57	0,000	0,011	-0,011	96	1.154,44	0,000	0,011	-0,011
52	3.010,93	0,000	0,011	-0,011	97	976,28	0,000	0,011	-0,011
53	2.242,23	0,000	0,011	-0,011	98	98,55	0,000	0,011	-0,011
54	584.727.527,67	0,000	0,011	-0,011	99	9.975.775,86	0,017	0,011	0,006
SOMA TOTAL					584.727.527,67				

pondem 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 calculada na 2ª e 7ª colunas em relação ao somatório de todos os custos unitários da planilha; a 4ª e 9ª colunas apresentam as frequências padrão da Lei NB; e a 5ª e 10ª colunas trazem a diferença entre as proporções das Somas e as frequências da Lei NB.

Conforme se observa na Tabela 2, houve picos nos dois primeiros dígitos 11, 17, 18, 19, 20, 21, 22, 25, 32 e 48. Chama a atenção à proporção encontrada para o dígito 25, que representou 48,3% do total dos custos unitários. Constatou-se, para esse teste, que os dados em nada se aproximaram de uma sequência de Benford.

3.3 CONFRONTO ENTRE O TESTE DOS DOIS PRIMEIROS DÍGITOS E O TESTE DA SOMA

A seguir, selecionaram-se os dígitos detectados como críticos no Teste dos Dois Primeiros Dígitos e no Teste da Soma. Foi, então, realizado um confronto entre os testes para confirmar a criticidade dos dígitos, comparando-se a frequência relativa deles na planilha com a proporção em termos materiais. Os resultados estão ilustrados na Tabela 3.

Na Tabela 3 encontram-se exibidos os dígitos selecionados em ambos os Testes realizados (coluna 1). A coluna 2 mostra as frequências relativas desses dígitos na planilha, conforme a coluna “Real” da Tabela 1. A coluna 3 exibe as proporções dos dígitos na planilha orçamentária, segundo a coluna “Real” da Tabela 2. A coluna 4 apresenta a confirmação ou não da criticidade desses dígitos.

Conforme se observa no confronto entre os Testes, apenas os dígitos 28 e 42 foram descartados da amostra, porque somente foram selecionados no Teste dos Dois Primeiros Dígitos em virtude de sua insuficiente frequência na planilha.

Os resultados do Teste dos Dois Primeiros Dígitos apontaram para os dígitos 11, 16 e 25 (descartados os dígitos 28 e 42). Já o Teste de Soma identificou valores excessivos para as proporções do 11, 17, 18, 19, 20, 21, 22, 25, 32 e 48. Percebeu-se que os dígitos 11 e 25 foram identificados em ambas as análises como excessivos e, por isso, possuíam maior risco de estarem superestimados. Para uma auditoria, seria recomendável uma análise a priori dos valores contendo os dois primeiros dígitos 11 e 25, para depois se examinar o restante.

Tabela 3: Confronto entre os dígitos selecionados no Teste dos Dois Primeiros Dígitos e no Teste da Soma

Dígitos	T. Dois Prim. Díg.	Teste da Soma	Dígitos críticos
11	0,058	0,060	Sim
16	0,039	0,000	Sim
17	0,031	0,033	Sim
18	0,022	0,032	Sim
19	0,019	0,039	Sim
20	0,022	0,035	Sim
21	0,017	0,040	Sim
22	0,014	0,039	Sim
25	0,029	0,483	Sim
28	0,005	0	Não
32	0,016	0,056	Sim
42	0,002	0	Não
48	0,007	0,084	Sim

3.4 CONFRONTO ENTRE OS RESULTADOS DOS TESTES DA LEI NB E A ANÁLISE DO TCU

Ao se confrontarem os dígitos 11 e 25 com o sobrepreço apurado pelo TCU, obteve-se o resultado apresentado na tabela 4.

Observou-se que sete itens da Curva ABC possuíam os dois primeiros dígitos dos custos unitários 11 ou 25, e que o somatório do sobrepreço apurado pelo TCU para esses serviços foi R\$ 41.601.247,32. Ressaltou-se que o item Sistema de cobertura tensionada teve o maior sobrepreço da Curva ABC e também representava o serviço mais caro da obra (R\$ 256.714.917,00). Seus primeiros dígitos 2 e 5, além de terem sido identificados pelo Teste de Soma como materialmente relevantes no orçamento, foram apontados pelo Teste dos Dois Primeiros Dígitos como excessivamente frequentes na planilha.

O TCU apurou um sobrepreço total de R\$ 149.972.318,01. O valor a maior de R\$ 41.601.247,32 representava 27,74% do total do sobrepreço.

Após o exame dos dois primeiros dígitos mais críticos, procedeu-se à análise dos demais.

Com base na Tabela 5, identificaram-se mais 10 itens da Curva ABC, cujo sobrepreço fora apontado pelo TCU. Somente para os dois primeiros dígitos 18 e 20 não se encontrou sobrepreço. Contudo, os serviços 18.052.222-6 - Sistema de painéis informativos, displays (...) e 18.052.259-6 - Transformadores, grupos geradores, no-breaks e (...), cujos custos unitários eram respectivamente R\$ 20.206.546,09 e R\$ 18.600.382,98, não foram analisados pelo TCU. Sendo assim, não se pode afirmar se seus preços estavam adequados ou não.

O sobrepreço apurado nos itens da Tabela 5 foi R\$ 65.692.812,51, que representava 43,8% do que havia sido apontado pela Corte de Contas (R\$ 149.972.318,01).

Somando-se o sobrepreço dos serviços das Tabelas 4 e 5, encontrou-se o valor de R\$ 107.294.059,83, que representava 71,54% do sobrepreço total apurado pelo TCU.

4. CONCLUSÃO

O presente trabalho testou a aplicação da Lei de Newcomb-Benford, como ferramenta de mineração de dados, aos custos unitários da planilha orçamentária da obra de reforma do Maracanã. Realizaram-se o Teste dos Dois Primeiros Dígitos e o Teste da Soma, ambos da Lei de Benford. Para mensurar a conformidade com a Lei NB, aplicaram-se o Teste Z, Teste Qui-Quadrado e Média

Tabela 4:

Confronto entre os resultados dos Testes da Lei NB e a análise do TCU para os dígitos 11 e 25

Dígitos	Serviço	Custo Unitário	Sobrepreço TCU
25	Sistema de cobertura tensionada, incluindo estrutura metálica, cabos "fully locked" e membrana "PTFE", conforme projeto da Hightex, ref. jop/juv, p/Estádio do Maracanã Copa 2014.	256.714.917,00	26.961.972,80
	Mobiliário esportivo, e Mobiliário do auditório para o Estádio do Maracanã Copa 2014 - conforme proposta da Mackey Indústria de Móveis Ltda.	25.518.649,48	4.057.726,58
	Demolição c/equip. de ar compr., de massas de concr. armado, exceto pisos ou pav.	258,38	699.201,19
11	Esquadrias de alumínio para o Estádio Mário Filho - Maracanã conforme proposta da Itelaf nº t 2011-0131-000 e caixilhos complementares.	11.920.282,97	1.935.661,95
	Pré-moldado em concreto armado compreendendo desde a fabricação até instalação no local, para arquibancada do Maracanã, conforme planilha orçamentária 001/2011- Maracanã Copa 2014.	11.771.177,22	3.605.433,06
	Sistema de sonorização (áreas internas e arquibancadas), para o Estádio Mário Filho - Maracanã, conforme proposta pp 95981/10 rev.1 da Prosegur Forn. e Inst. (Maracanã Copa 2014)	11.512.097,16	3.748.888,14
	Piso monolítico de alta resistência poliuretânico flakes3000c/acab. verniz pu acetinado e = 4 mm nsbrasil ou sim. aplic. sobre contrapiso nivelado e.3 mm.incl.este.Maracanã 2014.f/i.	113,28	592.363,60
Total			41.601.247,32

Tabela 5:

Confronto entre os resultados dos Testes da Lei NB e a análise do TCU para os demais dígitos.

Dígitos	Serviço	Custo unitário	Sobrepreço TCU
16	Forro em grade metálica eletrofundida malha 30x100mm, h=20 cm, sendo barra principal 20x2mm, em aço carbono SAE 1006/1020, acab.pint.eletr.st.fornec.e instalação, p/ Maracanã copa 2014.	163,52	1.514.293,41
17	Sistema de controle de acesso restrito e controle de acesso do público, para o Estádio Mário Filho - Maracanã, conforme proposta pp 95981/10 rev.1 da Prosegur. Forn. e Inst.(Copa 2014)	17.544.505,41	12.370.588,70
	Perfuração em solo para execução de estaca raiz, diam.410 mm.p/obra do Maracanã Copa 2014.	177,96	373.048,34
19	Serviços (cablagens, conectorizações, instalação e interligação de softwares, comissionamentos, startup e operações assistidos), conforme prop. pp 95981/10, Prosegur. For. Inst. Maracanã.	19.081.957,09	4.241.578,74
	Revest.de piso em placas de granito branco polar polido, med.(60x60)cm, e=2,00cm.assentado c/nata de cimento sobre argam.de cim. areia e saibro, traço 1:2:2 e rejunte cimento branco.	196,72	435.709,44
	Drenagem, top soil, grama natural e sintética e sistema de irrigação para o gramado do Estádio Maracanã - Copa 2014, conforme proposta da firma Campanelli.	1.919.536,42	722.897,49
21	Demolição de cobertura em concreto armado sobre a arquibancada do Maracanã (marquise), conforme planilha orçamentária n002/2009-EMOP.	21.439.722,66	13.464.684,58
22	Recuperação da estrutura do Estádio Mário Filho - Maracanã - Copa 2014, conforme planilha 004/2011 - EMOP.	22.904.212,57	10.839.664,36
32	Administração local p/obras de reforma c/modificações e acréscimos para o Estádio Mário Filho, Maracanã Copa 2014, conforme planilha nº 003/2011 EMOP.	32.863.882,27	12.563.282,35
48	Sistema de ar condicionado central e ventilação mecânica completa, p/o Estádio Mário Filho - Maracanã Copa 2014, conforme proposta PR-078620 (Ambienter) fornecimento e montagem.	48.844.340,89	9.167.065,10
Total			R\$ 65.692.812,51

dos Desvios Absolutos. Os testes, de uma forma geral, apontaram uma conformidade relativamente marginal dos custos unitários com a Lei de Benford.

Na análise dos dígitos dos custos unitários de forma individualizada, os dois primeiros dígitos 11 e 25 foram detectados no Teste dos Dois Primeiros Dígitos e no Teste de Soma. Constaram em sete itens analisados pelo TCU, incluindo o serviço com o maior sobrepreço apontado, o Sistema de Cobertura Tensionada (...). O sobrepreço identificado nesses itens somava R\$ 41.601.247,32 e representava 27,74% do sobrepreço total apurado pelo órgão de controle.

Além disso, o Teste de Soma identificou sozinho mais 10 serviços apontados pelo TCU com valores acima dos referenciais de mercado. O sobrepreço total dos 17 itens detectados pela Lei NB foi R\$ 107.294.059,83, que representava 71,54% do sobrepreço total apurado pelo TCU (R\$ 149.972.318,01).

Este trabalho constitui tão somente o início de um estudo sobre a aplicação da Lei NB à auditoria de obras públicas. Buscou-se apresentar a aplicabilidade dessa ferramenta às planilhas orçamentárias para que pesquisas futuras aprofundem no desenvolvimento de métodos efetivos de seleção de amostra de auditoria utilizando a Lei NB. Essas pesquisas poderiam estudar se existe uma interdependência entre os Testes da Lei NB propostos neste trabalho. Além disso, poderiam ser adotados outros níveis de significância nos testes estatísticos, além do nível de 0,05, para mensurar a conformidade, tais como 0,01 e 0,10, por exemplo. Outra pesquisa interessante seria testar isoladamente a variável quantidade das planilhas orçamentárias, como alternativa à variável preço, para verificar a aplicabilidade e efetividade da Lei NB na identificação de possíveis tendências, e fazer um confronto com as análises empreendidas pelos órgãos de controle na detecção de sobrepreço por majoração de quantitativo.