

questões eleitorais

A MATEMÁTICA DA URNA

Método estatístico do século XIX desbanca alegação de Bolsonaro sobre fraude nos votos eletrônicos

DALSON FIGUEIREDO, LUCAS SILVA E JULIANO DOMINGUES

20maio2021_13h59

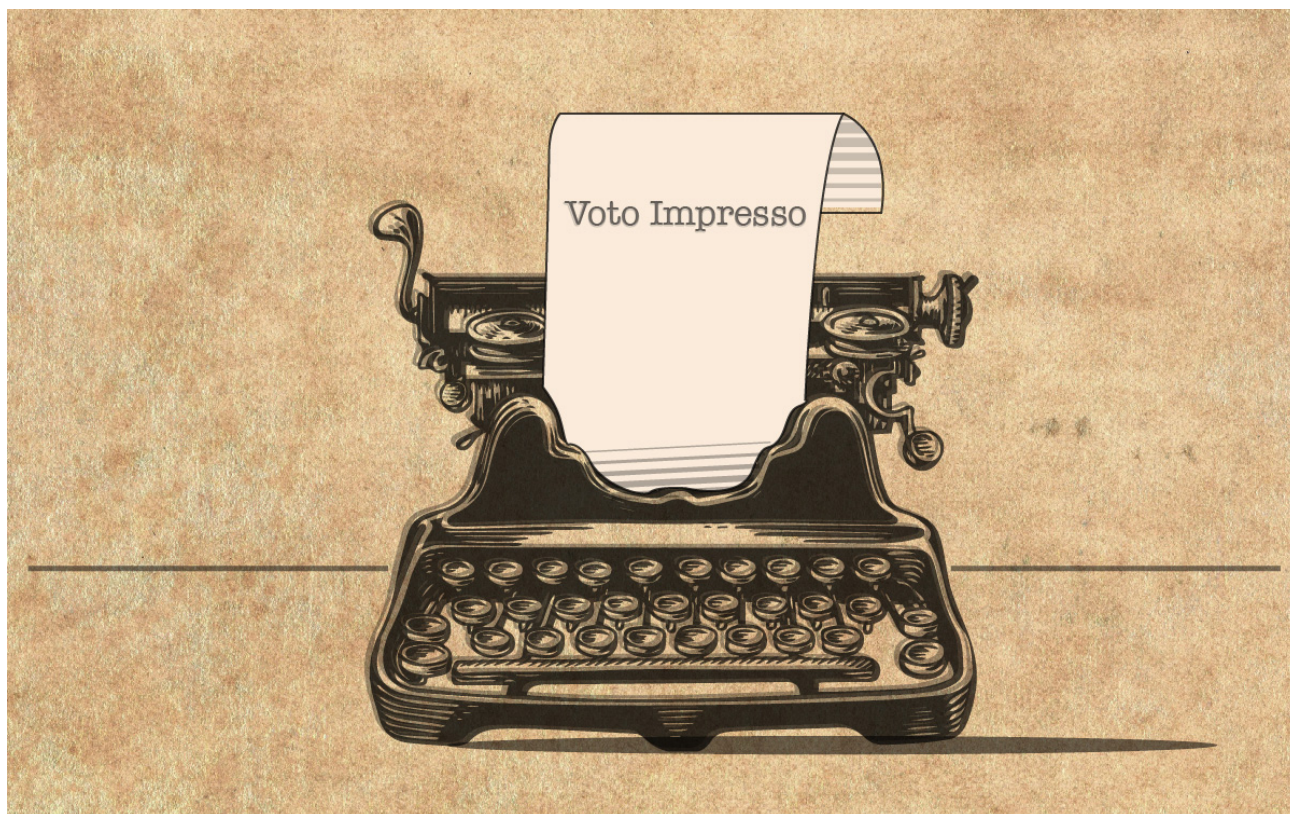


Ilustração de Carvall

O deputado Eduardo Bolsonaro (PSL-SP) afirmou em sessão da Comissão do Voto Impresso nesta segunda-feira (17) que não há como auditar as urnas eletrônicas. “Da mesma maneira que nós não temos como comprovar que houve fraude, o outro lado também não tem como comprovar que não houve fraude”, apontou. O deputado erra feio: não somente é possível auditar o resultado eleitoral como há, sim, modos de comprovar que não houve fraude. É o que este artigo vai demonstrar.

A auditoria das eleições não depende do voto impresso, pelo contrário. Há oito procedimentos previstos pelo TSE para aferir a lisura do processo, desde a verificação do resumo digital, passando pela comparação entre o boletim impresso e o boletim recebido pelo sistema



de totalização. A partir dessa perspectiva procedimental, portanto, cai por terra a primeira parte da argumentação.

A segunda parte da fala do deputado pressupõe uma afirmação passível de teste científico. Utilizamos a Lei de Newcomb-Benford (LNB), também conhecida como Lei do Primeiro Dígito ou Lei dos Números Anômalos, para colocar à prova a premissa propalada pelo Zero Três de que há fraude eleitoral no Brasil.

Antes dos números, um pouco de história. Em meados do século XIX, ao folhear um livro de tabela de logaritmos, o astrônomo canadense-americano Simon Newcomb observou um desgaste maior nas primeiras páginas, aquelas em que havia sequência de números iniciados com menor valor. As últimas folhas, com valores maiores, estavam mais preservadas. Ele havia acabado de identificar uma curiosa regularidade matemática. E lançou a seguinte proposição: a frequência dos dez dígitos (0 até 9) em números que ocorrem naturalmente não é uniforme. A chance de ocorrência dos números é inversamente proporcional ao tamanho do algarismo, de modo que os valores mais baixos (1, 2 e 3) são mais frequentes do que valores mais altos (7, 8 e 9). Por isso, as páginas iniciais daquele livro estavam mais desgastadas: elas eram as mais consultadas.

Mais de meio século se passou até Frank Benford avaliar empiricamente a proposição teórica de Newcomb. O artigo *The Law of Anomalous Numbers* foi publicado em 1938 na *Proceedings of the American Philosophical Society*. A partir de um levantamento de mais de 20 mil observações, Benford calculou a frequência relativa dos primeiros dígitos de uma variada quantidade de dados que aparentemente não guardavam nenhuma relação entre si, como o tamanho de populações, área de rios, distância entre cidades, constantes físicas etc. As distribuições observadas se aproximaram fortemente da expectativa teórica de que alguns algarismos aparecem com maior frequência do que outros. Hoje sabemos que tanto a sequência de Fibonacci quanto as taxas de mortalidade também seguem essa distribuição.

Simon Newcomb e Frank Benford se tornaram estrelas da ciência mundial e a LNB passou a ser aplicada a uma infinidade de campos do conhecimento. No Brasil, foi usada pelo Tribunal de Contas da União (TCU) para auxiliar na auditoria de preços de obras públicas. O professor Maurício Bugarin, do departamento de economia da Universidade de Brasília (UnB), usou essa técnica para examinar o orçamento da obra de



reforma do aeroporto Tancredo Neves em Minas Gerais. Recentemente, o Instituto Brasileiro de Auditoria de Obras Públicas publicou uma cartilha que ensina o passo a passo da aplicação da LNB na seleção de amostras elegíveis para processos de investigação. Fora do Brasil, o método já foi empregado para analisar doações eleitorais, dados epidemiológicos, comércio internacional, tipos de células, inflação, evasão de divisas etc.

A LNB também é chamada de Lei do Primeiro Dígito porque mostra que, em conjunto de dados numéricos que ocorrem naturalmente, o número 1 tende a ser o primeiro dígito significativo em 30% das vezes, enquanto o 9 surge em menos de 5% dos casos. Contrariamente, se todos os dígitos tivessem a mesma probabilidade de ocorrência, deveríamos observar uma frequência de 11,1% para todos os algarismos.

Por esse raciocínio, o primeiro passo para usar a LNB para detectar fraudes eleitorais é identificar o primeiro dígito da quantidade total de votos de cada candidato em cada um dos municípios brasileiros. Ele é chamado pelos estatísticos de dígito significativo.

Em Serra da Saudade (MG), município de menor população no Brasil, Jair Bolsonaro (então no PSL) recebeu 482 votos. O dígito significativo é, portanto, 4. Fernando Haddad (PT) conquistou 112 votos – 1 é o dígito significativo. Ciro Gomes (PDT) foi a primeira opção de 75 eleitores, o que coloca o número 7 como dígito significativo.

O segundo passo consiste em repetir esse procedimento para os 5.570 municípios brasileiros. Ou seja, para cada cidade, devemos identificar o primeiro dígito da quantidade total de votos de cada candidato. Computadores e pacotes estatísticos fazem isso em questão de segundos.

A última etapa é contar, no conjunto de municípios, quantas vezes cada algarismo aparece como dígito significativo para cada candidato. Se o conjunto de dados se adequa à Lei de Newcomb-Benford, o 1 será o dígito mais frequente, com cerca de 30% das menções. Mais ainda: os dígitos seguintes aparecem na mesma proporção esperada teoricamente pela Lei de Newcomb-Benford. Foi exatamente o que aconteceu. Observe os resultados:

As barras azuis ilustram a frequência do primeiro dígito para cada um dos três candidatos mais votados no primeiro turno das eleições presidenciais de 2018. A linha vermelha representa a distribuição que se conforma perfeitamente à Lei de Newcomb-Benford.

A leitura dos gráficos é simples: quanto maior a proximidade entre as barras e a linha, maior é a compatibilidade entre o resultado da eleição e o modelo teórico esperado. Por outro lado, quanto maior a distância entre as barras e a linha, maior é a diferença entre os dados observados e a distribuição esperada por Newcomb-Benford.

Para usar o jargão científico: não encontramos diferença significativa entre a distribuição de votos em Bolsonaro, Haddad e Ciro e o que prevê a LNB. O que permite concluir que não existem evidências de fraude nas urnas no primeiro turno das eleições presidenciais de 2018.

Ao aplicar o mesmo procedimento aos dados relativos ao 2º turno, chegamos à mesma conclusão: não existe diferença estatisticamente significativa entre a distribuição de votos observada para Bolsonaro, Haddad e aquela da LNB para o primeiro dígito. Devemos, então, rejeitar a hipótese de fraude nas urnas.



Os resultados descritos aqui sugerem que não há indícios de anormalidades nas eleições brasileiras. A lei pode ser utilizada apenas para o primeiro dígito, ou apenas para o segundo, que é mais indicado para lidar com dados eleitorais. Para este artigo, fizemos três análises (primeiro dígito, segundo e os dois primeiros). Todas apontam para a mesma conclusão: não houve fraude. Os dados estão todos disponíveis [aqui](#).

Essas conclusões ganham ainda mais força quando examinamos outros trabalhos sobre o tema. O professor Gauss Cordeiro (UFPE), em artigo publicado na [Revista Brasileira de Estatística](#) em conjunto com Cláudia Eirado (TRE-DF) e Gustavo Silva (USP), aplica a lei de Newcomb-Benford observando a presença do segundo dígito para analisar os resultados das eleições presidenciais no Brasil entre 2002 e 2018. As conclusões são as mesmas: não existem evidências de fraudes na contagem de votos.

Mas nem Galileu está a salvo dos negacionistas, e a matemática tampouco. Em 2018 a mesma Lei de Newcomb-Benford foi usada de modo enganoso para apontar supostas fraudes na eleição de 2014, que deu vitória à petista Dilma Rousseff – uma aplicação totalmente [distorcida](#). Entendida e aplicada de modo correto, a Lei dos Números Anômalos mostra que [não houve](#) anomalia nos resultados da eleição presidencial de 2014 nem de 2018.

O professor americano Walter Mebane, internacionalmente reconhecido como um dos maiores especialistas em fraude eleitoral no mundo, desenvolveu um modelo matemático complexo para examinar a consistência dos resultados eleitorais no Brasil em 2014. E [os achados permanecem](#): nem sinal de fraude, mesmo controlando por outros elementos – como a proporção de pessoas recebendo Bolsa Família, renda per capita, desigualdade de renda e composição racial. O professor Salomon Mizrahi, da Universidade Federal de São Carlos, também examinou, com base na Lei de Newcomb-Benford, os dados das eleições de 2014 e não encontrou sinais de irregularidades.

A afirmação de Eduardo Bolsonaro sobre o voto impresso não resiste, portanto, ao confronto com a ciência. O Zero Três mistura a falácia do falso dilema com a inversão do ônus da prova. A receita é rasteira: alguém afirma algo sobre a realidade – nesse caso, fraude eleitoral –, mas não apresenta evidências capazes de validar o que foi dito, ao mesmo tempo em que transfere o ônus da contraprova aos seus opositores no debate. Essa tem se mostrado uma prática usual na estratégia discursiva



da família Bolsonaro e de seus aliados em defesa de pautas específicas, do tratamento precoce para Covid-19 ao aquecimento global, passando pela inexistente fraude nas eleições. Ironicamente, a Lei do Primeiro Dígito demonstra ser insustentável a afirmação do deputado que atende por um número.

*

Este artigo reflete a opinião de seus autores, não necessariamente a da piauí

ASSINE NOSSA NEWSLETTER

Fique por dentro das melhores histórias da piauí

Inscreva-se em nossa newsletter e receba reportagens e outros conteúdos imperdíveis semanalmente

