



Brasília, 9 de outubro de 2018

Sobre o uso da Lei de Benford para avaliar fraudes eleitorais

Maurício Soares Bugarin

Professor Titular, Universidade de Brasília, www.bugarinmauricio.com

Líder do grupo de pesquisa

Documento preparado para o repórter Wellington Ramalho, UOL e Projeto Comprova

Não citar sem consentimento do autor

Sobre a Lei de Benford. Em poucas palavras, trata-se de uma regularidade estatística segundo a qual a frequência do primeiro dígito¹ de um conjunto de dados numéricos obtido de fontes naturais segue uma regra bem estabelecida. Segundo essa regra, o primeiro dígito 1 aparece aproximadamente 30% das vezes, o segundo aproximadamente 18% e assim reduzindo-se a frequência sucessivamente até se chegar ao número nove que aparece menos de 5% das vezes².

Comprimento de rios, populações de cidades, áreas de municípios, são todos exemplos de conjuntos de dados de origem natural que seguem de perto a Lei de Benford.

Quando os dados são de origem artificial, como os números de telefone, por exemplo, então essa regularidade não ocorre. E quando os dados são manipulados, como no caso de dados contábeis e empresas cotadas na bolsa? Então, em geral, a manipulação tende a distorcer o caráter natural dos números originais, fazendo-os deixar de se conformar à Lei de Benford. Mas então, comparando as frequências dos dígitos em um conjunto de números, pode-se medir quão longe estão da previsão com base na Lei de Benford, de forma a se ter uma medida indireta da possibilidade desses dados terem sido manipulados.

Aplicações da Lei de Benford. Foi com base nesse princípio que Nigrini (veja Nigrini 2012) desenvolveu testes estatísticos que possibilitaram a detecção de manipulação contábil em importantes empresas cotadas na bolsa de Nova York, tornando a metodologia mundialmente famosa. No Brasil, este autor e sua co-autora Flávia Cunha desenvolveram um algoritmo baseado na Lei de Benford para detectar fraude em licitações de grandes obras públicas (Bugarin & Cunha, 2015) que, uma vez comparada com as auditorias feitas pelo TCU, permitiu reencontrar entre 70% e 80% dos sobrepreços identificados pela corte.

A aplicação da Lei de Benford usando-se dados contábeis, bem como dados relacionados a planilhas de custos de obras públicas já está relativamente sedimentada e aceita tanto na literatura acadêmica como na prática.

Lei de Benford e fraude eleitoral. Uma outra aplicação que tem sido explorada recentemente diz respeito ao uso da Lei de Benford para se detectar indícios de fraudes em eleições, e tem como principal defensor acadêmico o professor Walter Mebane (vide, por exemplo, Mebane, 2010). Essa literatura, no entanto, ainda não está bem consolidada e é fruto de constantes críticas. A principal

¹ Também se aplicada para o segundo dígito, o “n-ésimo” dígito, $n > 2$, os dois primeiros dígitos, o último dígito, etc.

² Existe uma expressão logarítmica que pode ser consultada em Bugarin & Cunha (2015), por exemplo, mas que foi primeiramente proposta já em 1881 por Newcomb (Newcomb, 1881).



delas é que a Lei de Benford não funciona bem quando os números na base de dados são limitados por alguma razão. Por exemplo, número de crianças em salas de aulas são limitados pelas capacidades das salas. Em um universo em que as salas têm tamanho padronizado, então espera-se que o limite de capacidade distorça os dados, de forma que a Lei de Benford (no que se segue, simplesmente denotada por LB) não seja satisfeita.

O mesmo pode ocorrer com as urnas, uma vez que há um limite superior ao número de eleitores em cada seção eleitoral (em 2014, era de 400 eleitores). Essa limitação prejudica a capacidade da LB detectar fraudes. Alternativamente, pode favorecer a não-conformidade.

Por essa razão, Mebene sugere que seja usado o teste do segundo dígito, no lugar do teste do primeiro dígito, mas não está claro que essa estratégia resolva o problema. Por exemplo, Joseph Deckert, Mikhail Myagkov e Peter C. Ordeshook fizeram cuidadosos testes inclusive envolvendo simulações (Deckert, Myagkov e Ordeshook, 2011), chegando a conclusão de que o poder preditivo do testes baseados na LB é muito baixo. Eu mesmo orientei trabalhos de graduação e mestrado na área sem grande sucesso.

Por essa razão, desconfia-se da capacidade da Lei de Benford detectar com precisão a existência de fraude eleitoral. Para reforçar essa desconfiança, nota-se que grande parte dos estudos sobre o uso da LB em eleições encontram-se disponíveis na internet em formato de working papers, sem terem sido publicados em periódicos acadêmicos que envolvem o escrutínio de papers (peer review). Esse é o caso de Hoeschl, Bueno e Tonelli (2015), bem como as referências ali contidas.

Dito isto, a Lei de Benford, muito além do que um simples teste estatístico, representa uma metodologia de tratamento de dados mais geral como buscas a evidências de irregularidades. Ela pode e é aplicada em Hoeschl, Bueno e Tonelli (2015) de forma criativa e perspicaz, buscando reforço por meio de vários enfoques alternativos (votos nulos, votos em branco, número de eleitores por seção, etc.), levantando questões que devem ser consideradas pelos brasileiros preocupados com a lisura de seu processo eleitoral. Dentre elas, a mais importante é a necessidade de se desenvolver um mecanismo, ainda que amostral, que permita a auditoria física do resultado de uma urna eletrônica, como o voto em papel impresso pela urna eletrônica, mas depositado em urna física pelo eleitor, após verificar seu voto.

Conclusão: Um sinalizador, mas não uma prova. Em guisa de conclusão, devo enfatizar que, por ser um instrumental estatístico, a Lei de Benford não deve ser considerada como uma prova absoluta de existência de fraude, nem mesmo nos casos contábeis e de licitações públicas, mas sim um sinal de alerta que nos indica onde buscar as fraudes para então, por meio de auditoria clássica e outros métodos, se comprovar essas fraudes.

Do ponto de vista acadêmico, são necessários mais estudos controlados, de forma a se identificar o nível de precisão do uso da LB em estudos de eleições. No caso das planilhas de custos de licitações públicas, por exemplo, há a auditoria convencional feita pelo TCU que permite a comparação entre os sobrepreços sugeridos pela LB e aqueles encontrados pelo TCU. Infelizmente, no caso das eleições, não dispomos ainda de uma tecnologia alternativas e inquestionável de detecção de fraude para compararmos com a que é sugerido pela análise baseada na LB.