

LEI DE BENFORD E INTELIGÊNCIA ARTIFICIAL: UMA INTEGRAÇÃO NO TRABALHO DE AUDITORIA EM UMA PERSPECTIVA REGIONAL

Paulo César Roxo Ramos¹
Roberto de Goes Ellery Junior²
Antônio Nascimento Junior³

Resumo

Com a evolução da ciência e tecnologia novas formas para a detecção de fraudes contábeis foram concebidas. A Lei de Newcomb-Benford (LNB) ou simplesmente Lei de Benford, mostra-se uma ferramenta simples e eficaz e pode ser adotada para identificação de fraudes contábeis, através da comparação da frequência dos primeiros dígitos em relação a um padrão estabelecido empiricamente por Benford. Com a utilização de metodologias de Inteligência Artificial (IA) e mecanismos de aprendizado de máquinas, pode-se desenvolver ferramentas adaptativas para diferentes tipos de fraudes. Esse trabalho apresenta um modelo para análise estatística dos dados fornecidos pelo Tribunal Superior Eleitoral e emprega o modelo para analisar os dados das últimas eleições no Brasil. Os achados da pesquisa sugerem, de modo geral, que as distribuições encontradas em 2016 e 2020 seguem a Lei de Benford, enquanto em 2014 e 2018 as análises sugerem inconformidades.

Palavras-chave: Benford, Estatística, Inteligência Artificial, Detecção de Fraude, Eleições, Processos Naturais.

Abstract

The evolution of science and technology has brought new ways to detect accounting fraud. The Newcomb-Benford Law (LNB) is a simple and effective tool and can be adopted to identify accounting frauds by comparing the frequency of the first digit against a standard empirically pattern established by Benford. The use of Artificial Intelligence (AI) methodologies and machine learning method allows the development of adaptive tools for different types of fraud. This work presents a model for analyzing data provided by the Superior Electoral Court and employs the model for analyzing data from the latest updates in Brazil. The research finds that distributions found in 2016 and 2020 follow Benford's Law, while in 2014 and 2018 may not.

Keywords: Benford, Statistics, Artificial Intelligence, Fraud Detection, Elections, Natural Process.

1. Introdução

A auditoria é um método sistêmico para a obtenção de informações sobre a gerência de uma organização, seja ela pública ou privada [Mohd-Sanusi et al. 2012]. O objetivo da auditoria é aperfeiçoar e garantir a credibilidade das informações que são geradas em um dado processo. O resultado da auditoria ajuda a expressar uma opinião sobre a veracidade dos dados ou ativos da

¹ Mestrado em Economia: ECO/FACE/UnB, pauloroxo@me.com

² Doutor em Economia. Professor da ECO/FACE/UnB, ellery@unb.br

³ Doutor em Economia Professor ADM/FACE/UnB, anjunior@unb.br

organização. Os demonstrativos financeiros gerados no processo de auditoria constituem um conjunto de informações importantes para os investidores e economistas [Karkacier et al. 2017], onde através do balanço, eles julgam sobre a possibilidade de investir ou não em tal ativo. As auditorias fornecem uma visão externa sobre a saúde e capacidade da corporação, com isso, os investidores possuem subsídios adicionais para decidir sobre onde aplicar o capital [Arens et al. 2017, Mohd-Sanusi et al. 2012].

Os analistas de mercado buscam por oportunidades financeiras capazes de maximizar a rentabilidade e diminuir a exposição ao risco, um dos riscos diz respeito ao conflito de interesses que são inerentes aos agentes internos da corporação e que pode levar a distorções financeiras em seus balanços ou declarações. Tais inconsistências ou fraudes podem corroer a estabilidade da empresa, impactar negativamente sua imagem, diminuir o lucro ou, até mesmo, levar ao prejuízo [Robu e Robu 2015]. A auditoria existe para reduzir ou mitigar os ruídos oriundos do processo administrativo e gerencial. O método é aplicado por pessoas capacitadas, objetivas e independentes da organização. Os resultados e demonstrativos da auditoria corroboram e endossam as decisões dos investidores [Arens et al. 2017, Robu e Robu 2015].

Geralmente, não existe apenas um fator de decisão associado à deliberação de investir em dado fundo, o investidor leva em conta diversas variáveis para apoiar sua escolha. Além das informações associadas à empresa é necessário analisar o contexto no qual a mesma está inserida, como por exemplo, a estabilidade política do país e a lisura dos seus governantes. Tais variáveis exógenas refletem de modo direto e indireto nos resultados alcançados e divididos com os investidores. Nesse sentido, surge a possibilidade de auditar variados processos políticos ou de governança pública. Uma questão relevante é a confiabilidade no processo eleitoral de uma nação, já que isso impacta diretamente na soberania do seu povo e suas garantias fundamentais.

Considerando que os sistemas de governança desenvolvidos ao decorrer dos séculos possuem um histórico de fraudes e que existe uma divergência pública com relação aos resultados das eleições, pode-se criar mecanismos para identificar falhas no sistema eleitoral. Cardoso [Cardoso 2021] destaca que não há sistema completamente seguro e isento de falhas:

“Uma questão primordial a ser considerada é que nenhum sistema eletrônico pode ser avaliado como sendo completamente seguro, pois, de certa forma, nem a NASA, nem o FBI, ou o Pentágono Americano possuem sistemas de informática que sejam destacados como invioláveis” [Cardoso 2021].

A Justiça Eleitoral, criada em 1930, que entre os seus princípios visa moralizar as eleições, vem através dos anos criando mecanismos para modernização tecnológica do aparato eleitoral. Vale destacar que no ano 2000 o TSE passou a realizar eleições puramente eletrônicas, sendo este um marco determinante na história tecnológica do sistema eleitoral brasileiro. Outros incrementos foram feitos em sentidos similares, como a adoção de um software de código aberto (GNU/Linux) em 2008 e o cadastramento de biometria obrigatório a partir de 2020 [Cardoso 2021].

A urna eletrônica permitiu o cadastramento nacional único dos eleitores e a votação é feita de modo distribuído (por seção e zona), com tal advento a apuração passou a ser centralizada e rápida, em comparação com o método manual que emprega cédulas [Ferrão et al. 2019].

Por outro lado, na mesma medida em que o sistema foi automatizado, surgem especulações a respeito de fraudes, já que a ausência de papel torna o processo de auditoria algo abstrato e incompreensível por grande parte da população e políticos. Desse modo, os métodos de auditoria devem ser sistematizados para atingir níveis de excelência e confiabilidade, tomando como base sempre mais de uma fonte certificadora e, preferencialmente, de modo independente [Karkacier et al. 2017]. Além disso, a formalização dos resultados deve ser feita de modo empírico e simples de explicar.

O objetivo dessa pesquisa é verificar se os dados eleitorais seguem a distribuição de Benford e empregar o aprendizado de máquina para procurar por inconsistências ou divergências nos resultados que possam indicar algum ato fraudulento. O que se deseja é a interação sinérgica entre campos distintos do conhecimento, e que o produto sirva de usufruto para a sociedade e para agregar valor ao conhecimento, matéria primeira dessa dissertação. É justamente essa lacuna que esse artigo visa endereçar, considerando que a questão da urna eletrônica e o processo eleitoral no Brasil são temas recorrentes em discussão sobre transparência e fraudes, deseja-se aplicar a Lei de Benford associada a ferramentas estatísticas e métodos de Inteligência Artificial (IA) nos dados fornecidos pelo Tribunal Superior Eleitoral (TSE) do Brasil.

2. Referencial Teórico

Simon Newcomb foi um astrônomo americano-canadiano do século XIX que fez uma observação a respeito das tábuas de logaritmos. Newcomb percebeu que as primeiras páginas das mesmas se encontravam mais desgastadas do que as últimas páginas, a partir disso assumiu que os valores iniciados em 1 (um) eram mais utilizados do que os iniciados nos números posteriores e o desgaste das páginas eram inversamente proporcionais ao primeiro algarismo significativo [Cunha 2013]. À época, Newcomb documentou suas observações, porém sem apresentar outras evidências ou dados numéricos que comprovasse a sua veracidade. Apenas em 1938 evidências mais concretas foram relatadas por Frank Benford, que de forma acidental chegou às mesmas conclusões de Newcomb [Cunha 2013].

Benford publicou um artigo em que chamava o fenômeno observado como “A lei dos números anômalos” [Jošić e Žmuk 2020, Kruger e Yadavalli 2017]. O estudo desenvolvido por Benford consistiu em utilizar diferentes dados de fontes distintas, como área de rios, populações, medidas de pressão, constantes físicas, taxas de mortalidade, entre outras. Com esses dados em mãos, Benford numerou às vezes em que o primeiro dígito era o número 1, 2, 3, 4, 5, 6, 7, 8 ou 9. A partir disso, elaborou uma tabela com a porcentagem de vezes em que esses números apareciam em relação ao total. Verificou-se que o primeiro dígito era o número 1 (um) em 30,6% (trinta vírgula seis por cento) das vezes em média, o número 2 (dois) era o primeiro dígito em

aproximadamente 18,5% (dezoito vírgula cinco por cento) dos dados... seguindo um valor decrescente até o 9 (nove), que foi computado como primeiro dígito em 4,7% (quatro vírgula sete por cento) das vezes [Benford 1938]. Após essas observações, Benford formulou uma equação logarítmica (1) que determina aproximadamente a frequência dos primeiros dígitos. Sendo $F_{(a)}$ a frequência em que o dígito a aparece como o primeiro dígito utilizado.

$$F_{(a)} = \log\left(\frac{a+1}{a}\right) \quad (1)$$

A Lei de Benford não é aplicável a qualquer tipo de conjunto de dados. Os conjuntos gerados de forma aleatória não respeitam essa lei, possuindo a mesma probabilidade de ocorrência para todos primeiros dígitos (1/9). Números preestabelecidos pelo homem, como números de telefones e números de contas bancárias também não respeitam a Lei de Benford [Filho 2013].

A partir desses conceitos estabelecidos por Benford, diversos estudos foram feitos para verificar o uso da Lei na identificação de manipulações numéricas em diversos campos do conhecimento e uma grande gama de trabalhos sobre a aplicação da Lei de Benford em serviços de auditorias foi publicada [Cunha 2013]. Um estudo feito por [Costa 2012] inventariou as publicações relativas à Lei de Newcomb-Benford em periódicos internacionais e nacionais, o trabalho mostra a relevância da aplicabilidade do tema em diversos campos (matemática, economia, administração, política e outros)

Carslaw [Carslaw 2016] introduziu o uso de testes estatísticos utilizando a Lei de Newcomb para as frequências esperadas e observadas. Na detecção de fraudes através da Lei de Benford faz-se uso de um teste de alto nível, o teste do primeiro dígito, que possibilita que anomalias no conjunto de dados sejam identificadas. O teste do primeiro dígito compara a distribuição de frequência dos dígitos (entre 1 e 9) de uma amostra de teste com relação aos valores observados pela Lei Benford [Silva et al. 2017, Cunha 2013]. A partir dessa comparação, através do cálculo de desvios e análise gráfica, é possível observar se ocorreram grandes discrepâncias entre as medidas esperadas e as obtidas e, assim, efetuar uma análise minuciosa nas amostras que tais discrepâncias ocorreram, para verificar se de fato são encontradas fraudes ou erros [Diekmann e Andreas, 2007].

É um teste macro, capaz de identificar anomalias óbvias, contudo, investigações mais detalhadas devem utilizar essa técnica em conjunto com outras ferramentas estatísticas. Matematicamente pode-se expressar essa relação como em (2).

$$\text{Prob}(D_1 = d_1) = \log\left(1 + \frac{1}{d_1}\right) \quad (2)$$

onde D_1 representa o primeiro dígito e $d_1 \in \{1, 2, 3, \dots, 9\}$.

O teste do segundo dígito avalia a frequência de repetição dos números considerando o segundo dígito dos valores em uma amostra [Shikano et al. 2011]. É considerado um teste macro

e é útil para testar certas distorções [Cunha 2013]. A equação (3) apresenta a Lei de Benford com relação ao segundo dígito.

$$\text{Prob}(D_2 = \{d_2\}) = \sum_{d_1=1}^9 \frac{1}{d_1} \log\left(1 + \frac{1}{d_1 \cdot d_2}\right) \quad (3)$$

onde D_2 representa o segundo dígito e $d_2 \in \{1, 2, 3, \dots, 9\}$.

Embora a contagem numérica dos dígitos nos forneça uma boa aproximação, considerando uma tabela e um plot (gráfico), é necessário estabelecer limites aceitáveis de simetria. Nesse sentido, existem métodos estatísticos que podem ser combinados com a LNB para dar mais aderência e relevância matemática [Pernet 2016]. Ou seja, existe o interesse em dizer até que pontos os dados são significativos e quais hipóteses podemos aceitar ou rejeitar.

Pode-se adotar uma amostra aleatória da população no teste de hipótese e existem diferentes teste estatísticos para ajudar na decisão. Essa dissertação irá empregar o teste Qui-quadrado, o teste Z e uma rede neural artificial como métodos de teste estatístico. Alternativamente à hipótese nula temos a hipótese alternativa (H_1), que deve ser contraditória à primeira. No caso de Benford, a hipótese nula representa justamente a aderência à LNB, ou seja, nos diz que há uma semelhança entre a amostra e a distribuição de Benford.

Uma auditoria pode ser definida como estudo, avaliações sistemáticas e coleções de transações, procedimentos, rotinas de demonstrações financeiras de uma entidade com objetivo de fornecer às partes uma opinião imparcial sobre a conformidade dos dados com base em normas e princípios [Costa 2012, Junior 2012]. Dados genuínos deveriam seguir o padrão da Lei de Benford para o primeiro dígito, enquanto dados fraudulentos apresentariam uma distorção desta lei.

A lei de Benford não é uma lei universal. Divergências na LNB apenas indicam anomalias no padrão esperado dos dados, dando indícios que os valores que estão sendo analisados podem ter sido manipulados. Variações desta lei funcionam como um sinal e não um dado concreto a respeito da fraude, assim, uma atenção especial deve ser dada aos conjuntos de dados que apresentam distorção do comportamento esperado. Vê-se então que a aplicação da Lei de Benford pode ser uma poderosa ferramenta no auxílio ao serviço de auditoria, aumentando as chances de possíveis fraudes serem encontradas. Os avanços tecnológicos vêm causando diversas mudanças em organizações, governos e na sociedade como um todo e os setores econômicos são beneficiados com esses avanços ao terem mais facilidade de controle sobre seus dados financeiros [Junior 2010].

A contabilidade foi uma das grandes beneficiadas no ramo dos negócios onde a tecnologia de informação foi utilizada. Atualmente, grandes empresas de auditoria possuem sistemas computacionais que auxiliam na tomada de decisões, através de ferramentas que fazem uso de inteligência artificial [Omoteso 2012].

A utilização de inteligência artificial no processo de auditorias tornou-se um mecanismo extremamente importante na detecção de fraudes em dados financeiros, sendo de grande utilidade para o combate a esses tipos de crimes, visto que essas atividades ilícitas muitas vezes são elaboradas com grande criatividade, onde os padrões do processo de fraude são constantemente modificados. Sendo assim, a integração do uso da Lei de Benford com ferramentas de aprendizado de máquinas se mostra interessante para detecção e análise de dados fraudulentos [Badal-Valero, Alvarez-Jareño e Pavía 2018]. A Inteligência Artificial sistematiza e automatiza tarefas humanas e, portanto, é potencialmente relevante para qualquer atividade intelectual [Silva 2005]. O termo IA foi criado por John McCarthy em 1956 em uma conferência nos EUA [Andresen 2002], a IA direciona esforços para a criação de processos ou métodos que simulem o modelo neural, a fim de criar algum tipo de inteligência [Wang 2019].

O estudo [Soni et al. 2020] analisou uma ampla gama de setores que aplicam IA e investigou os impactos positivos e negativos sobre governos, comunidades, empresas e pessoas. O trabalho versa sobre as conquistas e inovações que foram influentes no campo do aprendizado de máquina; a sua relevância nas atividades empresariais e no mercado global. Redes Neurais Artificiais - RNAs

Embora muito se fale sobre o processo eleitoral brasileiro e seu risco associado à fraude, o assunto não é novo e muito menos característico do Brasil. Conforme observado por [Lehoucq 2003], a história apresenta diversos eventos onde o tema fraude eleitoral foi o centro das discussões. Países da América Latina, Estados Unidos e México estiveram sob olhares investigativos de suas eleições, o que sugere que esse tipo de fraude é comum e recorrente, além disso, nações desenvolvidas também estão suscetíveis ao problema. No trabalho de Rivero [Rivero 2012] é feita uma dissecação das eleições brasileiras no contexto histórico e cultural do país, Rivero argumenta que o processo eleitoral brasileiro se origina de modo mesquinho, onde o autor provoca e diz que há um agenciamento de poder na genealogia do voto:

“A invenção de um agenciamento de poder, como o voto, sempre tem por procedência um começo hermético, uma vilania que a origem vem celebrar. Existe sempre uma vilania nos inícios de algo que é de característica completamente distinta daquilo que se chama de origem...O início desse agenciamento é baixo e adiante se fará uma interpretação possível deste começo mesquinho.” - [Rivero 2012]

Uma fraude histórica, que é bem documentada, é conhecida como voto de cabresto, onde homens que possuíam o título de coronel mantinham o controle dos votos de grupos de pessoas que estavam sob controle ideológico do mandatário, o qual exercia um tipo de apadrinhamento sobre esses grupos [Rivero 2012]. Desse modo, pessoas que estivessem sob juízo do coronel deveriam votar nos candidatos escolhidos por ele. Esse tipo de fraude teve o início de seu declínio nos anos 30 com a prisão dos coronéis baianos na Revolução de 1930 [Rivero 2012].

Em 1932 o Presidente Wenceslau Brás visando moralizar o sistema eleitoral fez uma reforma que centraliza as eleições em um órgão chamado de Justiça Eleitoral, retirando o poder do

legislativo. Um dos avanços da época foi o voto secreto e feminino, porém, em 1937 Getúlio Vargas atendeu a setores conservadores e instituiu o voto indireto e aboliu os partidos [Rivero 2012], efeito que só foi reverso em 1946. Nos anos 50 houve endurecimento da lei e do processo de votação, o que dificultou a fraude de cédulas e trouxe alguns avanços.

Entre 1964 e 1985, entre o governo de João Goulart e Tancredo Neves, houve uma instabilidade eleitoral muito grande, devido ao governo militar que adequou as eleições de acordo com seus interesses [Rivero 2012]. A incredulidade e falta confiança no processo eleitoral brasileiro remonta às questões históricas e culturais do país, que são marcadas por eventos contraditórios e, por isso, a questão da urna eletrônica ainda levanta tantas dúvidas e ceticismo quanto sua transparência.

As urnas eletrônicas começaram a ser utilizadas em 1996 e no ano 2000 o sistema foi totalmente implementado [Cardoso 2021], para muitos autores o advento representa um grande avanço, que foi agregado ao processo eleitoral e facilitou a apuração de votos. Segundo Cardoso [Cardoso 2021], as últimas eleições podem ser consideradas referências mundiais graças à informatização do sistema. Por outro lado, alguns autores e profissionais de tecnologia colocam em xeque a segurança das urnas.

É exatamente neste contexto amplo que se insere essa dissertação, tendo em vista os aspectos econômicos, históricos, sociais, estatísticos e tecnológicos. O autor pretende relacionar tópicos interdisciplinares para construção de um meio sistemático e automatizado de análise. As eleições serão escrutinadas de acordo com a LNB de um modo macro, em linhas gerais, macro pode ser entendido como uma avaliação global dos resultados eleitorais, onde os métodos construídos irão fazer uma análise por estado e, em seguida, geral. O capítulo seguinte apresenta as bases necessárias para construção da abordagem proposta por esse trabalho.

3. Metodologia

Os dados foram obtidos do site do TSE junto ao link <<https://www.tse.jus.br/eleicoes/estatisticas/repositorio-de-dados-eleitorais-1>>, lá é possível baixar os resultados relativos às eleições, desde 1993 até 2020. A metodologia optou por analisar as eleições de 2014 em diante e, dentre os arquivos disponíveis, escolheu trabalhar com dados agrupados de modo geral. Os dados representam votos de modo global, sem nenhuma filtragem. Isso foi feito para simplificar a computação e análise.

Os dados do TSE representam as eleições de modo global, com resultados por candidato e de acordo com o cargo (vereador, prefeito etc.), assim, a metodologia optou por avaliar as eleições de maneira geral, ou seja, sem filtrar as planilhas. Isso foi feito visando enxugar o experimento, diminuir a complexidade e fornecer ao leitor uma visão ampla sobre o assunto, além disso, esse trabalho não visa esgotar o tema ou fazer uma auditoria completa e detalhada.

O próximo passo do trabalho foi descobrir a proporção dos dígitos das amostras obtidas. Como o foco deste trabalho não é a programação, durante a revisão bibliográfica a metodologia encontrou uma biblioteca livre [CodeDrome 2019] que já faz essa análise, retorna a relação de Benford e a relação relativa à amostra, além de outros dados discricionários, como a porcentagem da diferença e a diferença absoluta. Conhecidas a porcentagem de frequência de cada dígito e com a proporção de Benford é possível fazer o plot (gráfico) das duas distribuições. Através desses resultados iniciais é plausível ter uma ideia qualitativa sobre a compatibilidade ou não dos dados. A Figura 1 ilustra graficamente o método, onde o verde representa a distribuição real e o vermelho a amostra.

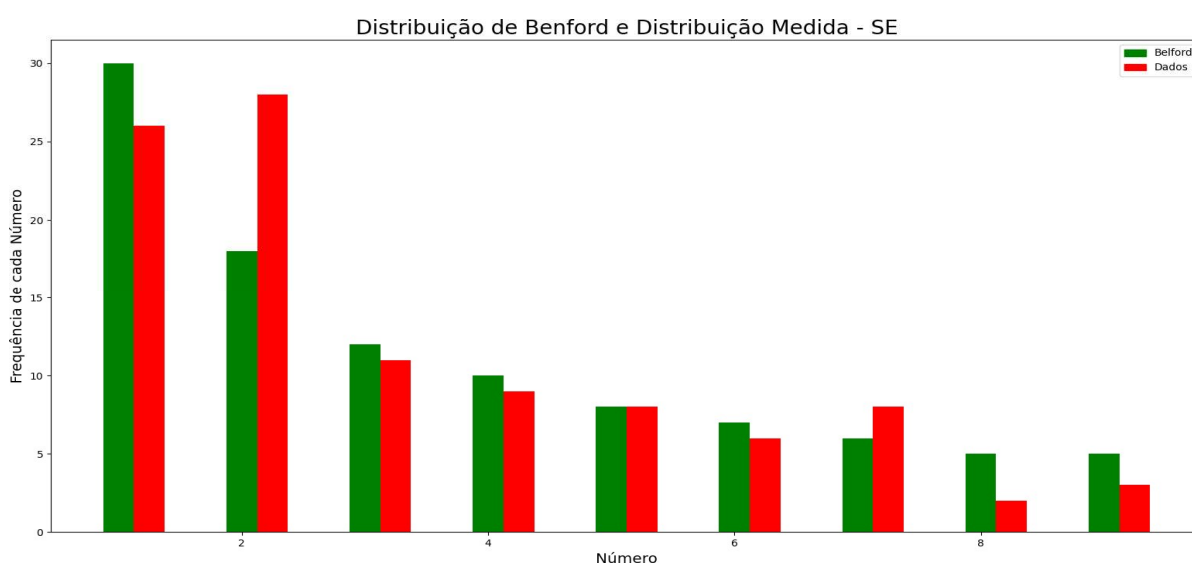


Figura 1 – Comparação visual entre a amostra de Sergipe e Benford.

Como o espaço amostral total (maior que 50 mil, em alguns casos) é muito mais elevado que a população de teste (200), optou-se por aderir a outra estratégia estatística, a finalidade foi reduzir a variabilidade que apenas um teste potencialmente poderia apresentar. Para isso a metodologia adotou os seguintes passos:

1. Realizou o experimento em 20 amostras aleatórias (ou randômicas);
2. Calculou o valor de Qui-quadrado e Z para cada amostras;
3. Com os valores calculados pode-se obter a média dos 20 testes.

Essa decisão diminui a chance de um teste aleatório não ser representativo. A mesma estatística foi adotada com o teste Z. Com o número de amostras (200) foi possível selecionar de forma aleatória conjuntos de linhas, calcular a frequência de cada dígito, comparar com a frequência esperada e calcular o valor de χ^2 . Para obter o resultado do teste Qui-quadrado optou-se por utilizar outra biblioteca estatística, facilitando assim a programação.

O resultado fica armazenado na variável (qui) χ^2 e é usado para calcular a média dos 20 testes sob a população do estado considerado. Os parâmetros e hipóteses do teste foram organizados como segue: H_0 : Não há risco de fraude nas eleições brasileiras; H_1 : Há risco de fraude nas eleições brasileiras; nível de significância fixado em 5% ($P=0.05$); oito graus de liberdade χ^2 crítico tabelado igual 15.507.

Os mesmos dados que alimentaram o teste anterior são essenciais para calcular os valores de Z para cada um dos dígitos. Nesse caso, foi necessário desenvolver o método como sugerido por Nigrini [Nigrini 2012] e outros autores [Cella e Zanolla 2018]. Nigrini sugere que, considerando a auditoria através da LNB, o cálculo do valor de Z deve ser feito individualmente para cada dígito. Com isso, pode-se avaliar anomalias em números específicos.

Durante os testes iniciais verificou-se que os resultados relativos ao segundo dígito quase nunca apresentavam desvios, a partir disso, o autor descobriu que o referido teste é mais adequado para casos em que há problemas de arredondamento, como em demonstrativos financeiros (onde há uma tendência de arredondar o 9). Por isso a metodologia exclui esse teste da avaliação, já que no caso das eleições é impossível acontecer um caso de arredondamento e, provavelmente, é por isso que nas avaliações iniciais não houve muita variação entre os resultados.

Para construir os métodos inteligentes esse trabalho adotou como ferramenta de apoio a biblioteca [Pedregosa et al. 2011] (scikit-learn), pois além da adequação a linguagem escolhida ela fornece a maioria dos algoritmos, dispensando assim a programação a priori. Para treinar uma rede com aptidão de classificação foi necessário, primeiramente, definir quais seriam as entradas ou características que alimentariam o método e a sua saída correspondente. Lembrando que essa pesquisa se enquadra em um caso típico de treinamento supervisionado, onde através de pares (entradas, resultado) a rede é treinada com exemplos.

Segundo o site [Pedregosa et al. 2011] que fornece as RNAs e demais ferramentas estatísticas, uma das etapas essenciais que antecede o processo de treino é a normalização dos dados. Basicamente o que essa etapa faz é colocar os valores das características dentro de uma mesma escala (entre 0 ou 1, por exemplo), desse modo, os dados irão possuir o mesmo nível absoluto e a contribuição de cada entrada se torna equivalente. Essa ação permite que a rede não faça suposições sobre uma característica ou outra, assim, um valor máximo de uma entrada igual a 1 é equivalente a outro valor máximo igual a 1000, pois ambas serão ajustadas para faixas equivalentes.

Pode-se citar dois métodos: a padronização dos dados e a normalização. Como esse trabalho desconhecia a distribuição das entradas, optou-se por testar os dois métodos e ver qual deles apresentava um resultado mais satisfatório durante o treino. Novamente, as funções utilizadas para normalizar foram importadas da biblioteca do software scikit-learn [Pedregosa et al. 2011]. Considerando a matriz com as entradas, discutidas na seção anterior, o programa automatiza todo processo de normalização sem a preocupação com o equacionamento.

Com as entradas normalizadas é necessário dividir a matriz (dataframe) de entrada em duas porções distintas. A primeira se refere aos dados que serão utilizados durante o processo de treino, ou seja, serão fornecidos à rede de modo iterativo para ajustar seus parâmetros. A segunda porção é empregada para testar se os resultados são fiéis, desse modo, é possível descobrir se o algoritmo é eficiente ou não, já que o teste valida o treino. Utilizou-se as proporções sugeridas no site [Pedregosa et al. 2011] da seguinte maneira: 70% das entradas são usadas no treino e 30% das entradas servem para testar.

Utilizando o programa [CodeDrome 2019] a pesquisa conseguiu gerar 90 mil amostras aleatórias que seguem relativamente a LNB (distribuição entre 0.8 e 1.2, lembrando que 1 significa aderência máxima) e 10 mil amostras que não seguem padrão algum (proporção aleatória de 1/9). A escolha desses valores se deu pelos seguintes fatos:

1. Geralmente, as tentativas de fraudes representam uma fração bem menor das transações, ou seja, são anomalias (1 fraude em 10 transações, por exemplo).
2. Com poucos valores (menores que 10 mil) a rede não obteve uma boa precisão (80%), o que sugeriu que seria necessário aumentar o número de amostras para permitir mais dados ao algoritmo.

Com os dados organizados é razoável escolher uma entre diversas redes do scikit-learn, assim, pode-se testar qual obtém a melhor resposta. Esse trabalho analisou apenas os métodos discutidos no capítulo anterior, não foi feito ajuste das configurações da rede e foram utilizados os padrões fornecidos para não estender o assunto. Finalmente, com os dados organizados, o modelo montado e as redes treinadas, a metodologia organizou os resultados em forma de gráficos (plot) e tabelas. Na seção seguinte serão apresentados os resultados e uma discussão sucinta sobre os principais achados. A metodologia optou por apresentar os resultados em conjunto, ou seja, considerando os métodos estatísticos e a IA.

4. Resultados

Os resultados serão apresentados em ordem cronológica, de acordo com o ano das eleições e por estado, além disso, a ordem em que os dados aparecem nas tabelas estão de acordo com a sequência gerada pelo programa e, a princípio, não segue nenhum padrão alfabético. Como os dados foram apresentados em uma ordem aleatória (por estado), a pesquisa manteve a mesma organização para facilitar a coleta dos resultados. Importante observar que alguns resultados (relativos a estados e anos específicos) vieram com arquivos em branco do site do TSE, provavelmente devido a alguma instabilidade ou erro e, obviamente, não foram tabelados nos resultados.

Será apresentada a avaliação com relação ao primeiro dígito, considerando a análise gráfica, o teste Qui-quadrado, o teste Z e a resposta fornecida pela rede neural. Para não tornar a

apresentação repetitiva será mostrado graficamente o resultado de alguns estados apenas e do Brasil todo. Os resultados completos, serão então, colocados em tabelas.

4.1 Eleições 2014

A Figura 2 traz os resultados para o estado do Rio Grande do Sul. Considerando o teste do primeiro dígito, o valor de χ^2 ($17.28 > 15.7$) rompeu o limite crítico tabelado, com isso rejeita-se a hipótese nula e aceita-se a alternativa, nesse caso, afirmando que não há uma correlação entre as distribuições, ou seja, não há adequação à LNB. O teste Z informou uma anomalia no primeiro dígito ($2.44 > 1.96$), essa medida pode ser facilmente identificada no gráfico, onde a proporção de 1 se aproxima de 37%. A rede neural sugere que a amostra esteja começando a tender para uma distribuição aleatória (> 15). O resultado da IA foi igual a 16, o que indica que não há uma adequação à LNB dentro dos limites estabelecidos pela metodologia. Na Figura 2 o gráfico superior representa o teste do primeiro dígito e o inferior descreve os resultados da análise do segundo dígito.

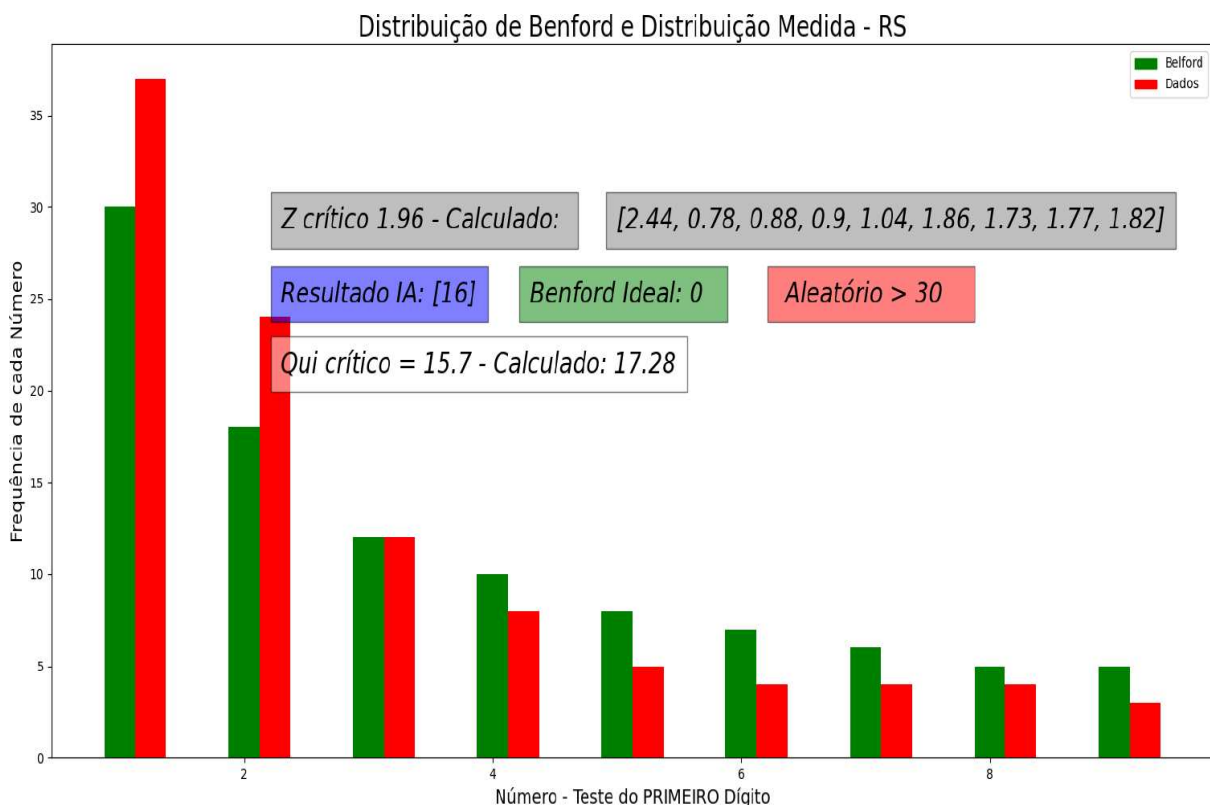


Figura 2 – Resultados do estado Rio Grande do Sul em 2014

Os resultados do Distrito Federal são exibidos pela Figura 3. Neste caso, a metodologia de pesquisa não encontrou nenhum indício que possa relacionar a distribuição de frequência numérica

a algum tipo de fraude. O valor de χ^2 ficou dentro do intervalo aceitável (6.43) A análise do valor de Z aplicado aos dígitos individuais não foi maior que 1.96 em nenhuma ocasião. O motor de IA também seguiu as previsões dos métodos estatísticos e, em ambos os testes, as amostras foram consideradas compatíveis com a LNB.

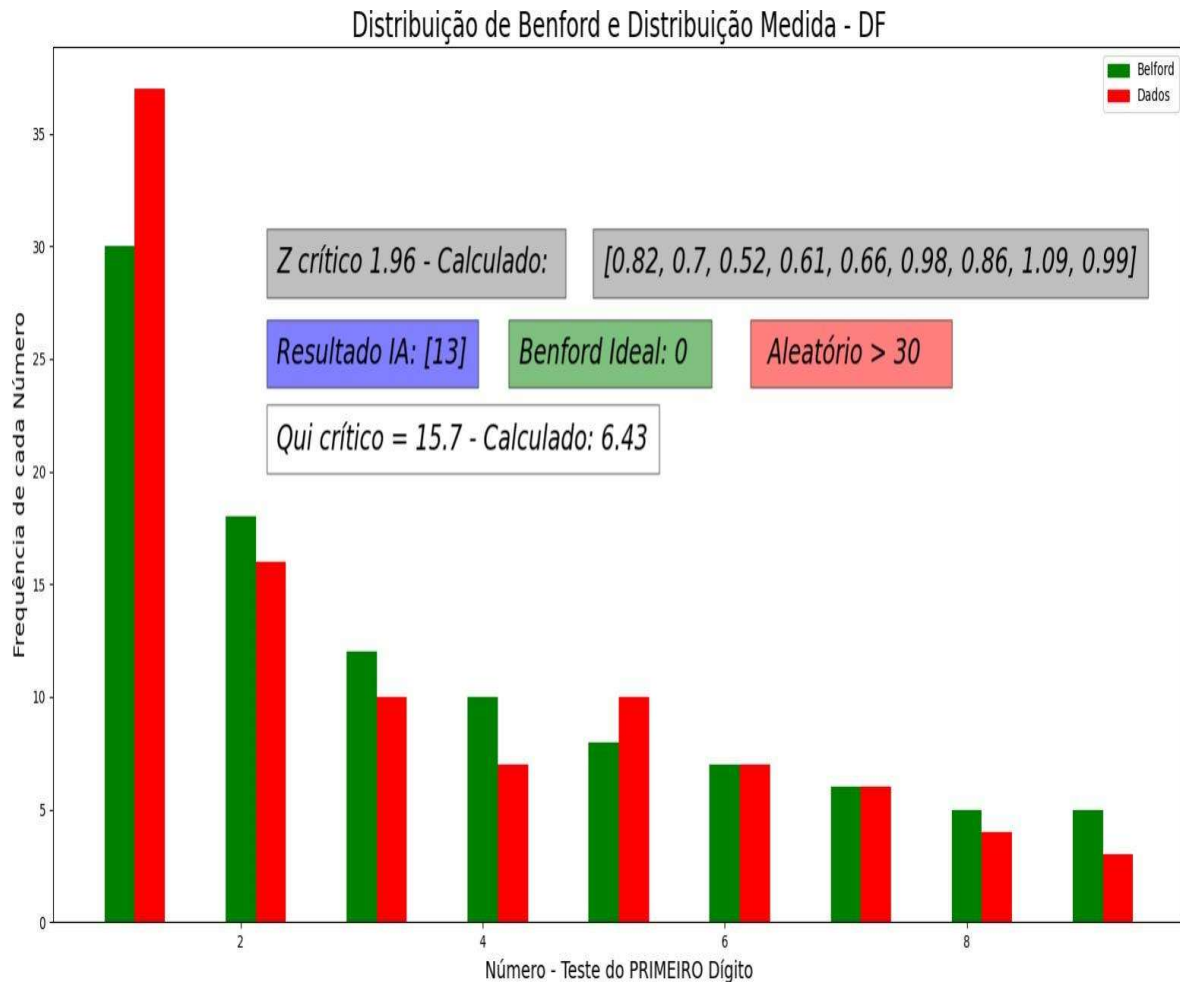


Figura 3 – Resultados do Distrito Federal em 2014

O estado de São Paulo apresentou distorções em três dígitos do teste Z (1, 8 e 9), considerando a análise do primeiro dígito. O teste do χ^2 seguiu a mesma tendência e, nesse caso, foi de 25.76 o que indica que a proporção não está correta. A IA não foi capaz de identificar anomalias e previu um valor de 13. Os resultados são apresentados na Figura 4.

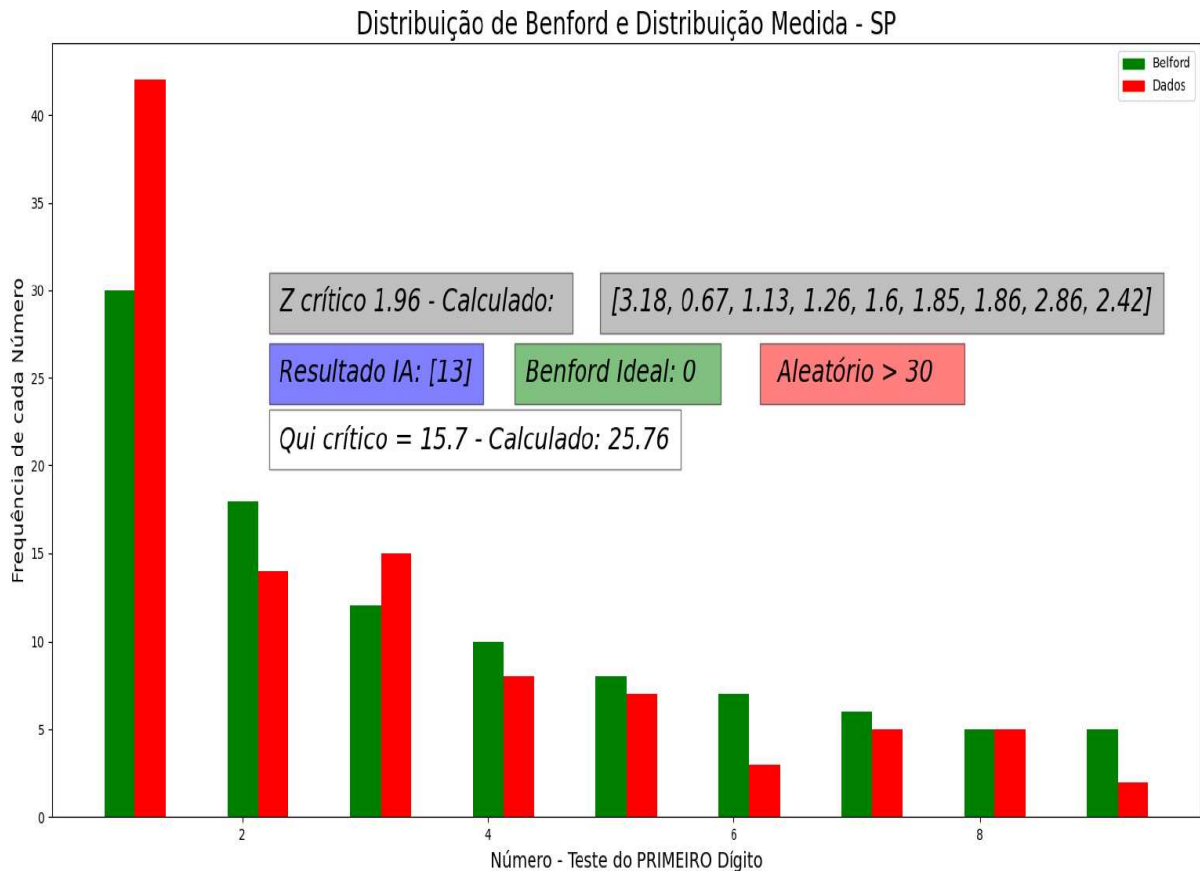


Figura 4. Resultados de São Paulo em 2014

Por último, tem-se o resumo gráfico dos resultados considerando dados de toda a nação. As informações estão resumidas na Figura 5 e indicam, em todos os testes e medidas mensurados, que não houve violação da Lei de Benford. Os valores dos χ^2 ficaram abaixo daqueles tabelados e a estatística Z não encontrou nenhum dígito cuja frequência possa ser considerada anormal dentro do contexto avaliado. Ou seja, considerando os resultados de maneira global (todo país), não houve quaisquer distorções identificadas pelos testes e pela rede neural.

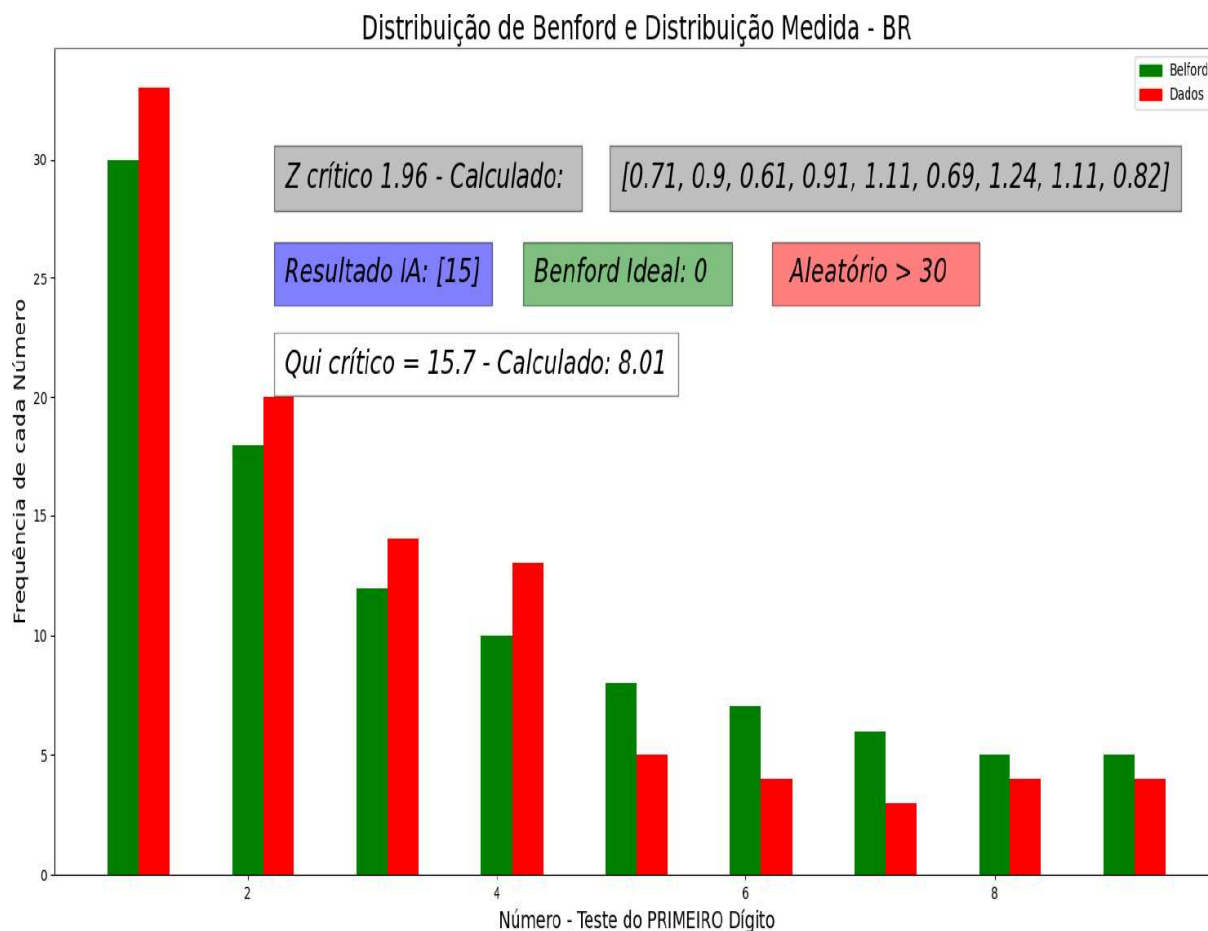


Figura 5 – Resultados Gerais do Brasil

Todos os resultados relativos ao teste do primeiro dígito do ano de 2014 podem ser encontrados na Tabela 4.1 de modo quantitativo. Os estados que apresentaram desvios, levando em conta os valores críticos empregados, foram o Rio Grande do Sul, Bahia, Pará, Ceará, Tocantins, Maranhão, São Paulo, Paraná, Santa Catarina, Rondônia, Goiás, Minas Gerais, Paraíba e Rio de Janeiro. Importante salientar que o principal dígito com distorção em relação a LNB foi 1, seguido pelo 9 e 8. Curiosamente a rede neural não indicou muitas anomalias (com bastante previsões marginais), todavia, os valores previstos não indicam aderência total à lei de Benford.

Tabela 1 – Tabela 2014 - Teste do Primeiro Dígito

Sigla	Teste do Primeiro Dígito - 2014									Qui (X)	AI
	Z (1.96)										
	1	2	3	4	5	6	7	8	9	(15.7)	-30
RS	2.44	0.78	0.88	0.90	1.04	1.86	1.73	1.77	1.82	17.28	16
DF	0.82	0.70	0.52	0.61	0.66	0.98	0.86	1.09	0.99	6.43	13
AP	1.61	0.70	1.30	1.44	1.0	0.83	1.1	0.93	1.03	11.13	12

MT	2.01	0.75	0.89	1.21	1.1	1.09	0.94	1.55	1.52	13.09	13
BA	2.47	0.70	0.89	0.89	1.16	1.78	1.60	2.05	1.61	17.13	15
MS	1.84	0.76	0.53	0.91	1.97	1.09	1.89	1.18	1.53	13.53	12
SE	1.63	0.76	0.90	1.05	1.72	1.35	1.62	1.21	0.65	11.76	15
PA	2.21	1.03	0.81	1.32	1.37	1.74	1.34	1.39	1.44	16.15	16
RN	1.99	0.93	0.69	0.78	0.89	1.98	1.09	1.92	2.35	14.85	12
AM	1.48	0.66	0.81	0.89	0.60	1.81	1.12	1.50	1.56	11.45	15
CE	2.69	0.79	1.01	1.25	1.46	1.37	1.20	1.48	1.89	17.81	13
TO	2.17	0.82	0.95	1.23	1.54	0.97	1.38	1.28	1.04	18.93	13
MA	2.26	0.88	0.58	1.06	1.32	1.61	1.74	1.51	2.09	16.48	14
PI	2.36	0.71	1.09	1.34	1.37	1.06	1.76	1.21	1.29	15.3	12
SP	3.18	0.67	1.13	1.26	1.60	1.85	1.86	2.86	2.42	25.76	13
ES	1.80	0.89	0.90	1.48	1.41	1.0	1.68	1.50	1.25	14.43	12
RR	1.76	1.03	0.78	1.25	0.92	1.04	1.50	1.42	1.61	12.62	16
BR	0.71	0.90	0.61	0.91	1.11	0.69	1.24	1.11	0.82	8.01	15
AC	1.42	0.93	0.73	0.71	1.39	1.91	1.82	1.29	1.10	12.57	15
PR	2.66	1.06	1.06	1.18	1.57	1.58	2.03	2.0	2.29	21.06	14
SC	2.05	0.78	0.60	1.31	0.70	1.0	1.13	2.16	1.75	13.39	12
RO	1.21	0.65	0.80	0.99	1.02	0.79	1.33	1.21	2.01	10.15	12
GO	2.20	0.95	0.73	1.09	1.27	1.27	1.57	2.01	2.15	16.29	14
MG	3.51	0.75	1.14	1.54	1.91	1.90	2.01	2.31	2.73	28.81	13
PE	1.67	0.88	0.69	1.08	1.12	1.06	1.53	1.57	1.83	12.51	15
PB	2.08	0.85	0.96	1.08	1.82	1.68	1.41	1.53	1.82	15.89	15
RJ	2.88	1.07	1.23	1.36	1.19	1.78	2.08	2.44	3.48	25.27	15
AL	2.13	0.64	1.0	1.08	1.36	1.20	1.71	1.18	1.67	14.37	12

Fonte: tabulações dos autores

4.2 Eleições 2016

Todas as métricas estatísticas e a IA apontam para a lisura do processo eleitoral brasileiro no ano de 2016, além disso, nenhum dígito individualmente fugiu à regra. Os resultados gerais dos testes estão discriminados na Tabela 2, pode-se inferir que sob nenhum aspecto a LNB foi violada, os valores críticos de cada método não foram atingidos por nenhum indicador. Embora a IA não tenha indicado nenhuma semelhança perfeita (igual a zero) em todos os testes, os valores ficaram dentro da faixa aceitável para aceitar a hipótese nula (as frequências amostrais são iguais).

Tabela 2 – Tabela 2016 - Teste do Primeiro Dígito

Teste do Primeiro Dígito - 2016											
Sigla	Z (1.96)									Qui (X)	AI
	1	2	3	4	5	6	7	8	9	(15.7)	-30
SE	0.68	0.94	1.01	0.84	0.62	0.79	0.87	1.21	1.19	9.1	13
PB	0.82	0.89	0.56	0.84	0.9	0.83	0.83	0.9	1.23	7.62	13
PR	0.61	0.80	0.79	0.91	0.85	0.76	1.06	1.1	0.87	7.72	15
AL	0.84	0.95	0.81	0.78	0.69	0.68	1.06	0.79	0.76	6.95	12
PA	0.96	0.68	0.67	0.60	0.82	0.83	1.37	0.92	1.01	8.11	11

MG	0.65	0.69	0.79	1.19	1.12	0.75	0.70	0.56	0.75	6.79	12
TO	1.23	0.71	1.55	1.48	0.67	0.99	1.42	1.24	1.0	11.95	15
AC	0.81	0.85	0.80	0.84	0.87	0.54	0.54	0.72	0.81	6.39	12
CE	0.64	0.77	0.89	0.64	0.68	0.97	0.62	1.0	0.89	7.86	12
RO	0.66	0.59	0.94	0.66	1.05	0.70	1.16	0.82	1.25	7.35	13
BR	0.77	0.92	0.73	1.11	0.58	0.9	1.06	0.83	0.54	7.43	13
RR	1.31	0.56	1.4	0.82	0.93	1.09	1.64	0.69	0.68	9.55	12
RN	0.82	0.69	0.92	0.74	0.77	0.97	1.38	1.2	1.44	10.1	13
GO	0.70	0.80	0.72	1.08	1.0	0.89	1.15	0.56	1.03	8.51	13
ES	0.70	0.68	0.82	0.57	0.77	0.76	0.80	0.73	99	7.2	14
AP	1.13	0.72	0.67	0.99	1.70	0.80	0.92	1.05	0.73	9.31	11
RJ	1.39	0.90	0.70	0.95	0.63	1.28	1.69	1.96	1.32	8.65	13
AM	0.63	0.50	0.80	0.70	1.14	0.70	0.72	0.87	1.14	6.84	13
MA	1.02	1.02	0.88	0.7	0.71	0.61	0.92	1.02	0.87	8.21	15
SP	0.71	0.99	0.81	0.72	0.96	0.72	0.71	1.13	1.1	7.91	11
SC	0.69	0.81	0.74	0.66	0.70	0.73	0.76	1.06	1.31	7.26	6
MT	0.78	0.61	0.87	0.91	1.15	0.71	0.80	1.03	0.81	7.68	9
PI	1.1	0.70	0.55	1.37	0.76	0.83	1.52	1.32	1.18	9.27	13
RS	0.69	0.94	0.76	0.59	0.78	0.81	0.70	1.09	0.83	7.21	15
MS	0.83	1.0	0.79	0.84	0.91	1.14	0.70	1.19	0.71	8.46	15
PE	0.91	0.71	0.94	0.84	0.97	0.94	0.75	1.25	1.6	9.91	13
BA	0.81	0.72	0.81	0.79	0.73	0.79	0.85	0.67	1.52	9.95	13

Fonte: tabulações dos autores

4.3 Eleições 2018

Neste item a apresentação dos resultados será sumarizada, em invés de discutir alguns estados individualmente a apresentação será geral, os casos específicos podem ser facilmente checados na Tabela 3. Com isso, espera-se tornar a apresentação mais fluida e objetiva. Considerando os dados amostrados de toda a nação, não houve sequer uma anomalia nas eleições de 2018, seja pelo método do primeiro dígito ou através da análise de IA. Os resultados estão dispostos na Figura 6 onde verifica-se a aderência à LNB.

Fazendo a avaliação individual por estado começam a surgir discrepâncias. No caso do primeiro dígito alguns estados como Paraná (19.86), São Paulo (23.78), Piauí (19.49), Bahia (18.72), Minas Gerais (26.50) e Goiás (16.45) apresentaram inconformidade com a LNB, com valores de elevados de χ^2 . Além disso, alguns dígitos individuais apresentaram a estatística Z acima do limiar, como Paraná (2.55 e 2.09), Tocantins (2.29), Rio de Janeiro (2.45), São Paulo (2.81), Espírito Santo (2.11), Amapá (2.50), Rio Grande do Sul (2.33 e 2.11), Piauí (2.36 e 2.08), Bahia (2.11 e 2.17), Minas Gerais (2.92, 19.96 e 2.16) e Goiás (2.09, 1.56, 2.12 e 2.01), os resultados integrais estão na Tabela 3. Avaliando a saída da rede neural, os estados de Roraima e São Paulo foram responsáveis por medidas anômalas.

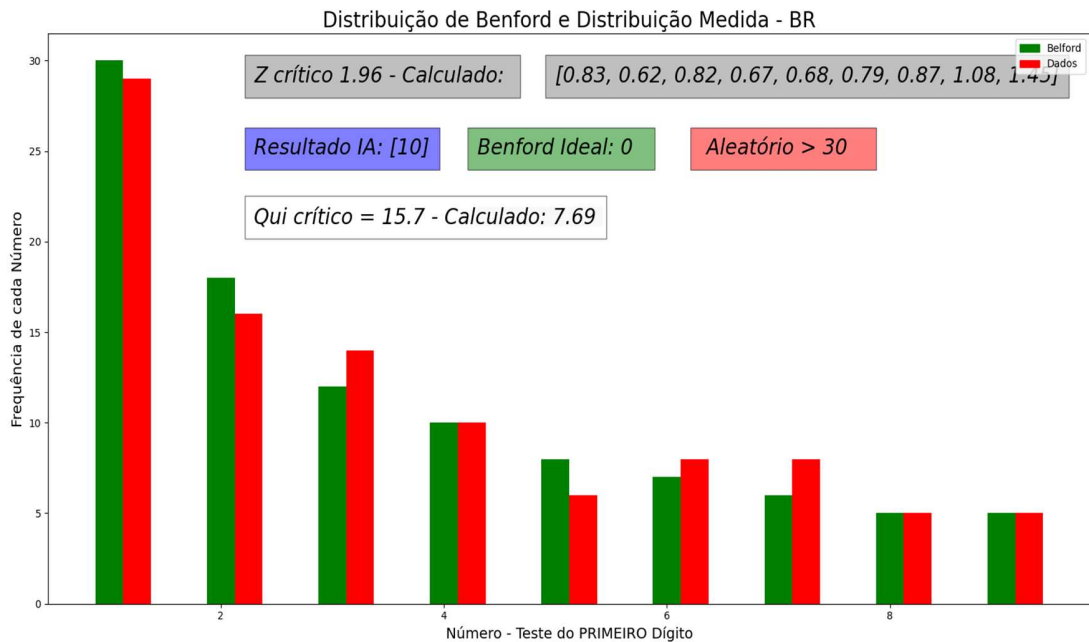


Figura 6 – Resultados do Brasil em 2018.

Tabela 3 – Tabela 2018 - Teste do Primeiro Dígito

Teste do Primeiro Dígito - 2018											
Sigla	Z (1.96)									Qui (X)	AI
PA	2.14	0.58	0.81	0.98	1.55	1.66	1.37	1.67	0.96	14.60	13
MS	1.86	0.79	0.80	1.02	0.69	1.59	1.47	1.40	1.65	12.48	12
RO	1.69	0.92	0.73	0.98	0.48	1.13	1.0	1.60	1.41	10.06	12
PE	1.77	0.92	0.82	1.18	0.66	1.23	1.66	1.51	1.94	13.53	13
MA	2.17	0.95	0.85	1.15	1.13	0.78	1.48	1.35	1.79	14.96	15
PR	2.35	0.99	0.85	1.29	1.42	1.95	2.55	2.09	1.90	19.86	15
RR	0.95	0.83	0.93	0.88	1.10	0.74	1.19	1.04	1.51	9.44	16
RN	1.69	0.59	1.09	1.23	0.77	0.84	0.90	1.15	2.03	11.39	12
SC	1.62	0.88	0.46	0.98	1.10	1.77	1.37	1.52	1.59	13.38	13
TO	1.86	0.77	1.03	0.79	0.96	1.42	1.26	2.29	1.43	12.91	13
MT	1.61	0.67	0.70	0.95	0.95	0.89	0.99	1.39	1.76	10.69	12
AC	1.70	0.86	1.05	1.25	1.11	1.69	1.20	1.89	1.80	14.5	14
SE	1.94	0.70	0.97	0.81	1.14	1.44	1.71	1.52	1.09	13.47	12
RJ	2.20	0.99	0.86	1.16	1.12	1.23	1.61	1.49	2.45	16.26	15
SP	3.01	0.77	0.78	1.66	1.79	1.50	1.95	1.89	2.81	23.78	16
DF	0.88	0.84	0.62	0.83	0.75	0.86	0.96	0.85	1.06	7.92	10

ES	1.63	0.68	0.66	0.85	0.82	0.87	1.47	1.90	2.11	11.61	13
RS	2.33	0.60	0.84	1.39	1.07	1.01	2.11	1.69	1.61	15.46	15
CE	1.74	1.02	0.62	0.99	0.89	1.47	1.77	1.56	1.41	13.22	13
AP	0.66	0.41	1.37	0.38	1.23	0.50	2.50	1.43	1.15	9.83	15
PI	2.39	1.04	0.82	1.77	1.28	1.27	1.13	2.36	2.08	18.49	14
BR	0.83	0.62	0.82	0.67	0.68	0.79	0.87	1.08	1.45	7.69	10
BA	2.48	0.68	0.62	1.26	1.69	1.36	2.11	2.17	1.90	18.72	14
MG	3.1	0.83	0.87	1.92	1.75	1.95	1.96	2.16	2.92	26.5	15
GO	2.09	0.80	0.80	0.81	1.12	2.09	1.56	2.12	2.01	16.45	13

Fonte: tabulações dos autores

4.4 Eleições 2020

O último teste foi o que apresentou maior conformidade com a LNB. O teste do primeiro dígito obedeceu a todas as estatísticas e nem um valor excedeu o Qui-quadrado tabelado ou o valor crítico de Z, ou seja, nenhum dígito apresentou qualquer tipo de anomalia, considerando qualquer estado ou o país inteiro. A rede neural seguiu a mesma tendência e não fez nenhuma previsão que ultrapasse 15. Os resultados integrais estão na Tabela 4.

Tabela 4 – Tabela 2020 - Teste do Primeiro Dígito

Teste do Primeiro Dígito - 2020											
Sigla	Z (1.96)									Qui (X)	AI
PI	0.82	0.99	0.65	1.24	1.03	0.81	0.90	0.84	0.92	8.49	13
MS	0.64	0.94	1.15	0.89	0.88	0.60	1.10	0.82	0.92	8.25	13
MG	0.85	0.73	0.97	0.93	0.69	0.96	0.50	0.94	1.23	7.64	13
PR	0.73	0.72	0.87	0.82	0.90	0.77	0.69	0.80	1.07	7.18	11
CE	0.74	0.69	0.63	0.64	0.66	0.65	0.91	1.1	0.95	6.84	6
ES	1.02	0.89	0.63	0.67	0.78	0.81	0.90	0.95	0.97	8.38	11
TO	1.14	0.45	1.56	1.03	0.78	0.96	0.76	1.25	88	10.06	15
AL	1.05	0.91	0.87	0.78	0.72	0.69	0.74	0.92	1.01	8.55	11
BR	0.86	0.92	0.57	0.62	0.87	1.05	0.56	1.24	1.05	8.02	12
AP	1.1	0.69	0.99	0.91	0.58	1.12	0.77	0.51	1.08	7.95	12
PE	0.66	0.74	0.64	0.80	0.65	0.74	0.96	0.83	0.69	6.62	15
RJ	1.02	0.78	0.82	0.70	1.20	1.31	1.12	0.70	1.33	9.04	13
MT	0.73	0.94	0.74	0.92	1.38	1.19	1.19	1.06	0.72	9.82	14
PA	0.87	1.10	0.91	0.83	1.10	1.11	0.79	0.48	8.91	8.91	15
AC	0.67	0.55	0.78	0.78	0.75	1.08	0.55	0.77	0.66	6.5	12
BA	0.75	0.63	0.69	0.47	0.87	0.97	0.64	1.15	1.0	7.08	13
PB	0.54	0.92	0.82	0.75	1.08	0.89	0.89	1.05	0.70	7.20	14
RO	0.53	0.96	0.83	1.11	0.92	0.69	1.11	0.78	0.90	8.05	15
AM	0.74	0.96	0.81	0.94	1.22	1.12	0.87	0.87	0.97	9.62	12

RR	0.74	0.98	0.58	0.63	0.87	0.76	0.93	0.82	0.88	7.45	13
SP	0.79	0.82	0.91	1.05	0.99	0.75	0.68	1.17	0.80	8.06	15
RN	0.86	0.92	0.76	1.03	0.95	0.72	0.92	1.08	1.1	8.63	13
RS	0.86	0.50	0.67	1.10	0.83	0.81	0.68	0.85	1.26	7.47	13
SE	0.81	0.79	0.46	0.57	1.12	0.75	0.80	1.28	0.81	7.44	15
GO	0.92	0.78	0.74	1.11	0.84	0.87	0.68	0.94	0.97	7.75	11
MA	0.97	0.94	0.73	0.72	0.77	0.87	1.07	0.77	0.77	8.29	13
SC	0.50	0.74	0.89	0.85	0.55	0.76	1.01	0.87	0.61	6.94	12

Fonte: tabulações dos autores

5. Análise dos Resultados

O ano de 2014 apresentou desvios consideráveis e podem ser considerados passíveis de uma análise ou auditoria mais criteriosa. Olhando o resultado do teste Qui para o primeiro dígito percebe-se que 12 estados extrapolaram a margem crítica aceitável e o valor de Z apresentou anomalias em diversos dígitos e estados. Isso sugere que há uma inconformidade ou fatores externos que podem estar alterando o processo natural, ou seja, pode existir algum ato fraudulento ou tentativa de manipulação. Essa observação não significa que a urna tenha sido adulterada ou sofrido algum ataque externo, mas sim que algum problema ou evento tenha ocorrido no sentido de mudar o curso da eleição, por exemplo, é possível citar outros meios de manipulação, como a compra de votos ou fake news direcionadas a alteração natural da percepção pessoal dos eleitores, o que em última instância, impactaria nos resultados.

Esse trabalho utilizou um método vanguardista para aplicação da LNB e adotou uma rede neural artificial para classificar uma classe em relação à adequação ou não a LNB, o algoritmo foi configurado para prever 0 em caso de conformidade total a lei e valores à partir de 15 estão no limiar de aceitabilidade. Nesse caso, a rede sugeriu que 3 estados não seguem a distribuição e outros 7 estão no limite máximo de aceitação. A abordagem apesar de pioneira é experimental, passível de melhorias e pode ser utilizada como um aditivo à estatística clássica, sem que a mesma seja excluída do processo avaliativo.

A LNB é uma métrica capaz de sinalizar um problema, porém, ela sozinha não pode indicar a natureza ou origem dos desvios, além disso, no curso desta pesquisa foi possível perceber que a análise macro (considerando os dados de todo Brasil) é menos suscetível a apresentar desvios, ou seja, a tendência de fraude é mais fácil de ser identificada de maneira micro, neste caso, avaliando os resultados por estado. Com isso, pode-se sugerir um trabalho futuro para ampliar a margem de confiabilidade do estudo e fazer uma comparação específica e refinada. Ou seja, o método pode ser aplicado em dados que representam exclusivamente municípios ou, até mesmo, agrupando conjuntos de seções e zonas eleitorais, o que permitiria identificar anomalias regionais e cruzar os resultados com os eventos locais que indiquem a origem de uma fraude em potencial. Além disso, como os dados do TSE apresentam todos os resultados na mesma planilha e com pouca

organização, é possível pensar em uma pesquisa futura direcionada a filtragem e tratamento dos dados.

A metodologia adotada sugere que em 2016 não houve nenhum indício que possa colocar as eleições como um evento que não adere à LNB. A estatística Qui-quadrado e Z foram praticamente unânimes em seus verídicos, o que permite aceitar a hipótese nula: não há risco de fraude na eleição de 2016. Esses achados foram em quase totalidade os mesmos para ambos os testes, além disso, a IA não fez nenhuma previsão alarmante para o primeiro dígito, que como já foi discutido, se associa mais ao caso do arredondamento.

Além disso, os resultados apresentados graficamente dão a mesma noção intuitiva de conformidade numérica, que vai de encontro com os testes utilizados. Como mencionado no item anterior, seria interessante abordar essa questão de maneira mais específica e colocar uma lupa nos dados, uma vez que esses estejam devidamente organizados e dispostos.

Embora mais discreto que em 2014, os resultados deste item sugerem violações da LNB em alguns estados. Como em 2014 e 2018 as eleições foram para presidente, governadores, deputados e senadores, pode-se questionar se esse tipo de eleição está mais sujeita à fraude ou se existe algum fator associado que distorce a distribuição dos votos. Na ocasião 7 estados apresentaram inconsistências relativas a métrica χ^2 , além disso, os mesmos também apresentaram valores que indicam inconsistências em alguns dígitos específicos ($Z > 1.96$), analisando-se o primeiro dígito. Já a IA classificou 2 estados como não pertencentes a uma distribuição de Benford, enquanto outros 6 estão no limite superior.

É importante ressaltar que os métodos e análises devem ser feitos em conjunto, ou seja, os resultados individuais podem ser somados para criar uma base de comparação mais sólida e, em seguida, o auditor pode se qualificar de modo mais abrangente. Eleições 2020

No ano de 2020 as eleições seguiram os resultados de 2016, o que novamente indica que os pleitos municipais tendem a apresentar uma conformidade maior à LNB. Esse ponto é curioso e pode se relacionar com a maneira como os dados são adquiridos ou com relação a quantidade de candidatos, não há uma conclusão óbvia a esse respeito. Considerando o teste do primeiro dígito, todos os métodos estatísticos e a rede neural apontaram para a lisura das eleições.

De forma geral houve no Brasil uma tendência de adequação à LNB, exceto o ano de 2014 e 2018 que apresentaram indícios, em alguns estados, de desvios quanto à lei de Newcomb-Benford.

6. Conclusão

No artigo fizemos um estudo da aderência das eleições de 2014, 2016, 2018 e 2020 através da análise estatística e aplicação da Lei de Newcomb-Benford nos dados fornecidos pelo TSE. Além disso, o trabalho introduziu um método adjunto de verificação, utilizando uma rede neural

artificial modelada de acordo com a LNB. A metodologia procurou por semelhanças e diferenças nas amostras, comparando cada frequência com a de Benford, com isso forneceu-se à literatura um inventário geral sobre os últimos pleitos. Pode-se dizer, de forma geral, que os resultados alcançados foram satisfatórios e o objetivo proposto pela dissertação foi atingido.

Com relação aos resultados, o autor observou que as eleições de 2014 e 2018 apresentaram desvios alarmantes. Alguns estados violaram a LNB em mais de uma métrica considerada, além disso, é possível verificar graficamente que existe uma boa margem de discrepância entre a amostra considerada e a frequência esperada. Em muitos casos a rede neural fez a mesma sugestão e em outros apresentou valores marginais. Com esse somatório de indícios é possível e recomendável fazer uma análise por estado, sem considerar o resultado global, já que este indicou uma boa aderência à LNB. Esse fato sugere que, caso os desvios existam, eles provavelmente acontecem em um nível micro e a manipulação se associa a regiões específicas. Além disso, Minas Gerais, São Paulo e Bahia apontaram grandes discrepâncias em mais de um teste. Pode-se concluir que a aplicação do método encontrou discrepâncias e elas devem ser dissecadas em um contexto micro, ou seja, por região e considerando seus fatores intervenientes associados ao contexto regional. Essa conclusão permite ao autor sugerir um trabalho futuro que versa sobre as cidades, colocando uma lupa mais potente no problema e analisando as três esferas separadamente (municipal, estadual e federal), além disso, uma análise por legenda ou grupos de candidatos também poderia tornar o processo mais criterioso e objetivo.

Um dos diferenciais desse trabalho foi a adoção de redes neurais artificiais em consórcio com a Lei de Benford, neste sentido os achados são satisfatórios e apresentam grande potencial de estudo. As previsões da IA na maioria dos casos seguiu os resultados dos métodos estatísticos utilizados, com poucos desvios. Como o método não foi unânime em todas as observações, recomenda-se não o utilizar sozinho e sim em conjunto com as ferramentas estatísticas. Finalmente, considerando a proposta da dissertação, pode-se concluir que o tema é relevante, tem um forte apelo social e econômico, além disso, a abordagem mista empregada colaborou com a literatura multidisciplinar, já que os métodos implementados se apoiam em ciências distintas.

7. Referências Bibliográficas

ANDRESEN, S. L. John mccarthy: father of ai. *IEEE Intelligent Systems*, v. 17, n. 5, p. 84–85, 2002.

ARENS, A. A. et al. *Auditing and assurance services: an integrated approach*. Sixteenth edition, global edition. Boston: Pearson, 2017.

BADAL-VALERO, E.; ALVAREZ-JAREÑO, J. A.; PAVÍA, J. M. Combining Benford's Law and machine learning to detect money laundering. An actual Spanish court case. *Forensic Science International*, v. 282, p. 24–34, jan. 2018.

BENFORD, F. The Law of Anomalous Numbers. Proceedings of the American Philosophical Society, v. 78, n. 4, p. 551–572, 1938.

CARDOSO, M. A incredibilidade das urnas eletrônicas no Brasil. 2021. Disponível em: <<https://monografias.brasilecola.uol.com.br/computacao/a-incredibilidade-das-urnas-eletronicas-no-brasil.htm>>.

CARSLAW, C. Anomolies in income numbers: Evidence of goal oriented behavior. In: [S.l.: s.n.], 2016.

CELLA, R.; ZANOLLA, E. Benford's Law and transparency: an analysis of municipal expenditure. Brazilian Business Review, v. 15, n. 4, p. 331–347, jul. 2018. CodeDrome 2019 CODEDROME. Benford's Law in Python. 2019. Disponível em: <<https://www.codedrome.com/benfords-law-in-python/>>.

COSTA, J. I. d. F. Desenvolvimento de metodologias contabilométricas aplicadas a auditoria contábil digital: uma proposta de análise da lei de Newcomb-Benford para os Tribunais de Contas. Tese (Dissertação) — Universidade Federal de Pernambuco, Recife, 2012.

CUNHA, F. C. R. d. Aplicações da lei Newcomb-Benford à auditoria de obras públicas. Tese (Doutorado) - Universidade de Brasília, 2013.

DIEKMANN, Andreas. Not the First Digit! Using Benford's Law to Detect Fraudulent Scientific Data. Journal of Applied Statistics. 34. 321-329, 2007.

FERRÃO, I. G. et al. Urnas eletrônicas no Brasil: linha do tempo, evolução e falhas e desafios de segurança. Revista Brasileira de Computação Aplicada, v. 11, n. 2, p. 1–12, maio de 2019.

FILHO, M. A. F. M. A Confiabilidade dos Dados Financeiros de Hospitais Filantrópicos Canadenses: Um Estudo Empírico Baseado na Lei de Benford. Sociedade, Contabilidade e Gestão, v. 8, n. 2, nov. 2013.

JOŠIĆ, H.; ŽMUK, B. The Application of the Law of Anomalous Numbers on Global Food Prices in Examining Psychological Pricing Strategies. Journal of International Food & Agribusiness Marketing, p. 1–16, jul. 2020.

JUNIOR, E. P. M. Investigação de modelo de auditoria contínua para tribunais de contas. Tese (Programa de Pós-Graduação em Ciência da Computação) - Universidade Federal de Pernambuco, Recife, 2010.

JUNIOR, J. H. P. Auditoria De Demonstrações Contábeis: Normas E Procedimentos. 5. ed. [S.l.]: ATLAS - GRUPO GEN, 2012.

KARKACIER, Atila & ERTAŞ, Fatih. Independent auditing effect on investment decisions of institutional investors. Journal of Accounting and Management Information Systems. 16. 297-319. 10.24818/jamis.2017.

KRUGER, P.; YADAVALLI, S. The power of one: Benford's law. *South African Journal of Industrial Engineering*, v. 28, n. 2, set. 2017.

LEHOUCQ, Fabrice. "E lectoral fraud : Causes, Types, and Consequences". *Annual Review of Political Science*, vol. 6, no 1, junho de 2003, p. 233–56.

MOHD-SANUSI, Z. ET AL. The importance of financial audit: Benefits, challenges and service quality towards client loyalty. In: 2012 IEEE Symposium on Business, Engineering and Industrial Applications. [S.l.: s.n.], 2012. p. 614–618.

NIGRINI, M. J. (Ed.). *Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2012.

OMOTESO, K. The application of artificial intelligence in auditing: Looking back to the future. *Expert Systems with Applications*, v. 39, n. 9, p. 8490–8495, jul. 2012.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011

PERNET, Cyril. "Null Hypothesis Significance Testing: A Short Tutorial". *F1000. Research*, vol. 4, outubro de 2016, p. 621.

RIVERO, Facundo. *Tecnologia e Política: o voto e seu suporte*. Pontifícia Universidade Católica de São Paulo, 2012, <https://tede2.pucsp.br/bitstream/handle/3406/1/Facundo%20Guerra%20Rivero.pdf>.

ROBU, M. A.; ROBU, I. B. The Influence of the Audit Report on the Relevance of Accounting Information Reported by Listed Romanian Companies. *Procedia Economics and Finance*, v. 20, p. 562–570, 2015.

SHIKANO, SUSUMU & MACK, VERENA. When Does the Second-Digit Benford's Law-Test Signal an Election Fraud? Facts or Misleading Test Results. *Journal of Economics and Statistics (Jahrbuecher fuer Nationaloekonomie und Statistik)*. 231. 719-732.

SILVA, R. A. C. *Inteligência Artificial Aplicada a Ambientes de Engenharia de Software: Uma Visão Geral*. INFOCOMP, v. 4, n. 4, p. 27–37, 2005.

SILVA, W. B. d. et al. Using the Newcomb-Benford Law as a Deviation Identification Method in Continuous Auditing Environments: A Proposal for Detecting Deviations over Time. *Revista Contabilidade & Finanças*, v. 28, n. 73, p. 11–26, abr. 2017.

SONI, N. et al. Artificial Intelligence in Business: From Research and Innovation to Market Deployment. *Procedia Computer Science*, v. 167, p. 2200–2210, 2020.

WANG, P. On Defining Artificial Intelligence. *Journal of Artificial General Intelligence*, v. 10, n. 2, p. 1–37, jan. 2019.