

# ANÁLISE MULTIFATORIAL DA CONSISTÊNCIA DOS DADOS DE MORTALIDADE DO SARS-COV-2 NOS PRINCIPAIS PAÍSES DA ORGANIZAÇÃO MUNDIAL DA SAÚDE: UMA ABORDAGEM UTILIZANDO A LEI DE NEWCOMB-BENFORD E ÍNDICES DEMOGRÁFICOS.

Bruno Bezerra Trigueiro<sup>1</sup>, Prof. Dr. Fernando Gonzales  
Tavares<sup>2</sup>, Prof. Dr. José Augusto Theodosio Pazetti<sup>3</sup>

<sup>1,3</sup> Faculdade de Tecnologia da Baixada Santista –  
Rubens Lara

<sup>2</sup> Universidade Santa Cecília  
bruno.trigueiro@pm.me  
josepazetti@gmail.com

## 1. Introdução

A Lei de Newcomb-Benford (LNB), uma distribuição logarítmica que descreve a frequência esperada dos dígitos iniciais em conjuntos numéricos, tem sido amplamente aplicada em diversas áreas, como contabilidade, demografia e física, para detectar fraudes, validar dados e identificar anomalias (Cleary & Thibodeau, 2005; Nigrini, 1994; Skousen et al., 2004; Drake & Nigrini, 2000). No contexto de crises sanitárias, como a pandemia de COVID-19, a integridade dos dados de mortalidade se torna uma questão central para a gestão eficaz e transparente da crise. A precisão e a confiabilidade dos dados de mortalidade são fundamentais para monitorar a progressão da doença, avaliar a eficácia das políticas públicas e guiar decisões de saúde pública em tempo real.

A história mostra que, em crises sanitárias anteriores, como o surto de SARS em 2003 e a pandemia de gripe H1N1 em 2009, a coleta de dados enfrentou desafios semelhantes. Esses eventos demonstraram que inconsistências nos dados, seja por falhas de notificação ou divergências metodológicas, comprometem a compreensão do verdadeiro impacto da crise e dificultam a alocação eficiente de recursos (World Health Organization, 2010). No caso da SARS, as lacunas na comunicação inicial dos dados atrasaram as respostas globais. Da mesma forma, a pandemia de H1N1 revelou

dificuldades na coleta e na padronização dos dados de mortalidade em nível internacional. Esses exemplos sublinham a importância crítica de dados confiáveis para enfrentar adequadamente crises de saúde pública.

Ao aplicar a LNB aos dados diários de óbitos durante a pandemia de COVID-19, este estudo busca avaliar a integridade desses dados e investigar se os padrões observados seguem a distribuição esperada. Isso é particularmente relevante quando consideramos a enorme variabilidade nos dados de mortalidade reportados em diferentes países, influenciados por metodologias divergentes de coleta e fatores políticos. A comparação entre fontes de dados, como o Consórcio de Veículos de Imprensa e o Ministério da Saúde no Brasil, revela como essas diferenças podem impactar os números reportados e, consequentemente, a aderência à LNB.

A utilização de métodos estatísticos, como limites de tolerância e o cálculo do Desvio Absoluto Médio (MAD), permitirá medir a conformidade dos dados de mortalidade com a LNB, fornecendo uma visão robusta da integridade dos dados (Barney & Schulzke, 2016; Druicã et al., 2018). Ao compreender como políticas de notificação e variáveis demográficas influenciam a confiabilidade dos dados, este estudo oferece insights que podem orientar políticas públicas e melhorar a gestão de crises sanitárias futuras.

## 2. Metodologia

A seleção dos países e regiões para análise foi estratégica, considerando a diversidade geográfica, o tamanho da população e a disponibilidade de dados confiáveis. Incluíram-se tanto países com alta incidência de casos quanto aqueles com menor número de registros, buscando uma visão abrangente do fenômeno.

### 2.1. Coleta e Preparação de Dados

Para esta análise, dados diários de óbitos foram meticulosamente coletados de fontes renomadas, incluindo o conjunto de dados COVID-19 da Universidade Johns Hopkins (JHU) (Anexo B, 2020-2024; Dong et al., 2020), do Ministério da Saúde do Governo do Brasil (Anexo D) e do Centro de Controle e Prevenção de Doenças (CDC) (Anexo A). A abrangência geográfica desses conjuntos de dados garante uma amostra diversificada e representativa. Para a aplicação da Lei de Benford, o primeiro dígito significativo foi extraído de cada contagem diária de óbitos (Wang & Ma, 2024).

### 2.2. Distribuição pela Lei de Benford

A LNB dita que a probabilidade de um dígito  $d$  (onde  $d$  varia de 1 a 9) ser o primeiro dígito em um número é dada por (Newcomb, 1881):

$$P(d) = \log\left(1 + \frac{1}{d}\right) \quad (1)$$

Essa distribuição logarítmica gera as frequências observadas na Tabela 1.:

**Tabela 01** - Distribuição de dígitos esperados pela Lei de Newcomb-Benford

<sup>1</sup> Aluno de Iniciação Tecnológica (IT) com bolsa CNPq - PIBITI

d	P(d)	Probabilidade de ser o primeiro dígito
1	0.301029995664	30.1%
2	0.176091259056	17.6%
3	0.124938736608	12.5%
4	0.0969100130081	9.7%
5	0.0791812460476	7.9%
6	0.0669467896306	6.7%
7	0.0579919469777	5.8%
8	0.0511525224474	5.1%
9	0.0457574905607	4.6%

Fonte: autor

### 3. Testes Estatísticos e Limites de Tolerância

Para avaliar a conformidade das frequências observadas do primeiro dígito com a Lei de Benford, foi empregada uma bateria de testes estatísticos. O desvio padrão foi utilizado no cálculo de limites de tolerância, visando avaliar a qualidade do ajuste da distribuição observada à distribuição esperada. Adicionalmente, testes Z (denominados neste trabalho como MAD, sigla em inglês para Desvio Absoluto Mediano), utilizado por padrão na análise de dados relacionados a LNB (Barney & Schulzke, 2016; Druică et al., 2018; Kössler et al., 2021, 2024; Nigrini, 2012; Tošić & Vičič, 2021), foram realizados para examinar a significância estatística das discrepâncias entre os valores observados e os esperados.

A fim de considerar a variabilidade inerente aos dados reais, foram estabelecidos limites de tolerância baseados em um intervalo de confiança de 95% (Equações 2 e 3). Conforme o estudo de Henselmann et al. (2012), esses limites delimitam a faixa aceitável de variação em relação às frequências esperadas pela Lei de Benford para cada dígito. O cálculo dos limites de tolerância (L) é dado por:

Limite inferior (LI):

$$L(d) = \log\left(1 + \frac{1}{d}\right) - [1.96 * s(d1) - \frac{1}{2n}] \quad (2)$$

Limite superior (LS):

$$L(d) = \log\left(1 + \frac{1}{d}\right) + [1.96 * s(d1) + \frac{1}{2n}] \quad (3)$$

Onde:

- 1.96 é o valor crítico para um intervalo de confiança de 95%.
- sd1 é o desvio padrão para um determinado primeiro dígito, calculado por:

$$s(d1) = \left( \frac{\log\left(1 + \frac{1}{d}\right) * (1 - \log\left(1 + \frac{1}{d}\right))}{n} \right)^{\frac{1}{2}} \quad (4)$$

- Sendo n o tamanho da amostra. Ou seja, o número total de dias analisados.

Quanto ao MAD, a função considera a aplicação da equação base da LNB, conforme (1), e é calculado da seguinte maneira:

$$MAD = \frac{1}{N} \sum_{i=1}^9 |p(d) - \log(1 + \frac{1}{d})| \quad (5)$$

Onde:

- p(d) é a proporção observada da frequência do dígito do país.
- N é o número de dígitos possíveis (N = 9).

### 4. Fontes Específicas de Dados para Análise

O presente estudo utiliza dados de diversas fontes para analisar a aderência dos dados diários de mortes por COVID-19 à LNB. Cada fonte apresenta características e limitações específicas.

#### 4.1. Consórcio de Veículos de Imprensa

O Consórcio de Veículos de Imprensa (CVI) foi uma parceria estabelecida entre 8 de junho de 2020 e 28 de janeiro de 2023, composto pelos veículos jornalísticos: O Estado de S. Paulo, G1, O Globo, Extra, Folha de S. Paulo e UOL. O objetivo inicial da parceria era relatar dados sobre a pandemia de COVID-19 no Brasil recebidos das secretarias estaduais de saúde, devido à restrição que o Ministério da Saúde havia promovido na época na divulgação de dados sobre o número de casos e mortes (Barbosa et al., 2022; Dos Santos, 2022; Roncalli & Lacerda, 2020).

O acesso aos dados divulgados pelo CVI, tanto na mídia quanto em seu website oficial, mostrou-se limitado. Não houve resposta às solicitações de contato, e o website não oferece opção de download direto dos dados.

Diante disso, os dados foram extraídos através de interceptação da requisição http entre o cliente do pesquisador e a base de dados do site oficial do CVI (Anexo C), possibilitando a coleta dos dados sobre a evolução da pandemia no Brasil.

Os dados obtidos incluem média semanal, médias móveis, total de mortes e variação absoluta diária. Para a análise, utilizou-se a variação absoluta diária, representando o número de mortes diárias. O conjunto de dados abrange o período de 17/03/2020 a 28/01/2023.

#### 4.2. Ministério da Saúde do Brasil

Apesar do conjunto de dados da John Hopkins University (JHU) utilizar o Ministério da Saúde como fonte, sua limitação para 1146 dias, em comparação aos 1568 dias disponíveis no website oficial do Ministério, resulta em uma menor taxa de amostragem. Os 422 dias adicionais, compreendendo o período de 2023/1 a 2024/1, embora fora do pico da pandemia, permanecem relevantes para a análise. Portanto, optou-se pela base de dados do Ministério da Saúde, visando maior precisão nos resultados.

<sup>1</sup> Aluno de Iniciação Tecnológica (IT) com bolsa CNPq - PIBITI

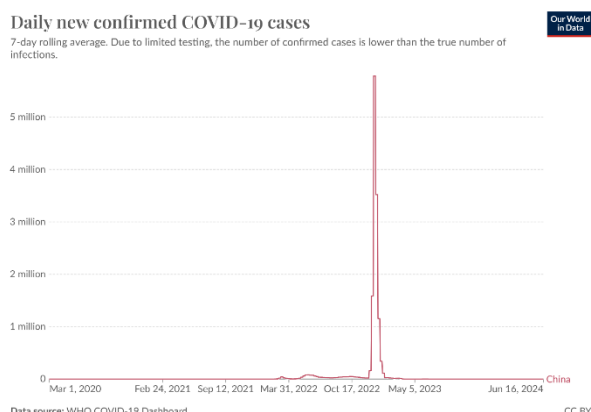
### 4.3. Centro de Controle e Prevenção de Doenças (CDC)

A base de dados do CDC foi escolhida para suprir a falta de informações sobre a disseminação da COVID-19 em todo o território estadunidense presente na base de dados da JHU. Apesar da frequência semanal dos dados do CDC (235 semanas) ser compatível com a aplicação da LNB, ela aumenta a suscetibilidade a outliers devido ao tamanho amostral reduzido, um fenômeno similar à maior propensão a falsos positivos observada na aplicação da Lei de Benford a conjuntos de dados pequenos e moderados (Barney & Schulzke, 2016; Henselmann et al., 2012; Günther et al., 2019).

## 5. Remoção Da China No Conjunto De Dados

Em nome do rigor científico, determinados conjuntos de dados foram excluídos da análise devido a inconsistências intrínsecas. Os dados da China, por exemplo, foram considerados inadequados após um aumento abrupto e inexplicado nos casos confirmados, saltando de aproximadamente 31 mil para 3,52 milhões em uma única semana de dezembro de 2022, como ilustrado na Figura 1.

**Figura 01** - Número de óbitos confirmados pela COVID-19 na China.



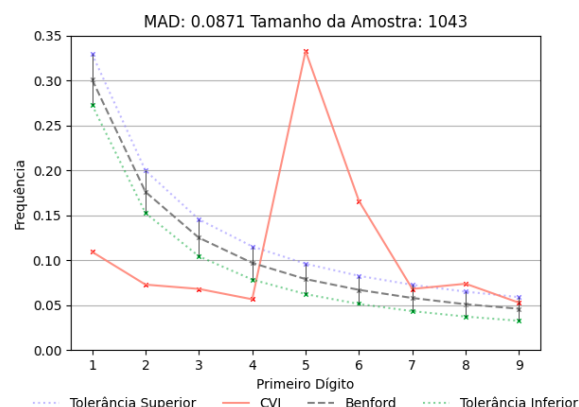
Fonte: Our World in Data.

Tal anomalia distorce os resultados e obscurece padrões genuínos de aderência.

## 6. Resultados e Discussões

A aplicação do algoritmo aos dados do Consórcio de Veículos de Imprensa (CVI) revelou um cenário preocupante quanto à integridade dos dados divulgados, conforme evidenciado pela Figura 2.

**Figura 02** - Distribuição da Lei de Benford para o Consórcio de Veículos de Imprensa.



Fonte: autor.

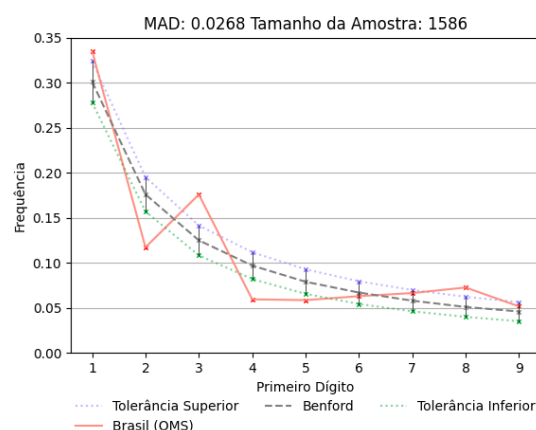
Quase todos os dígitos no conjunto de dados do CVI estão fora dos limites de tolerância superiores e inferiores estabelecidos, com exceção dos dígitos 7 e 9. Essas discrepâncias entre as frequências observadas e as frequências esperadas pela LNB são significativas.

De particular interesse é o outlier observado na frequência do dígito 5. A LNB prevê uma ocorrência de 7,9% para o primeiro dígito ser 5; no entanto, no conjunto de dados do CVI, ele aparece em 33% dos casos registrados. Essa diferença de 25 pontos percentuais levanta questões sobre a validade dos números relatados no site oficial do Consórcio de Veículos de Imprensa e na grande mídia de televisão.

### 6.1. Dados do Brasil fornecidos pelo Ministério da Saúde

A análise dos dados diários de mortalidade atribuídos à COVID-19 no Brasil, obtidos da Organização Mundial da Saúde (OMS), revela uma adesão intrigante à LNB. Com um tamanho amostral de 1586 registros diários, o desvio absoluto médio calculado é de 0,0268.

**Figura 03** - Distribuição da Lei de Benford para os dados do Ministério da Saúde do Brasil.



Fonte: autor.

Conforme observado na Figura 3, os dados não acompanham integralmente a curva teórica. O dígito inicial 1 excede o limite superior de tolerância por pouco, enquanto o dígito inicial 2 apresenta um desvio notável, ficando abaixo do limite inferior esperado em quase 4%.

<sup>1</sup> Aluno de Iniciação Tecnológica (IT) com bolsa CNPq - PIBITI

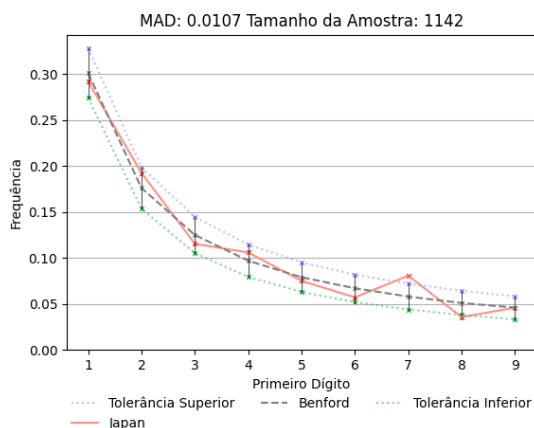
Os dígitos iniciais 6, 7 e 9, por sua vez, estão alinhados dentro dos limites de tolerância prescritos.

As discrepâncias observadas, particularmente para o dígito inicial 2, ressaltam a necessidade de uma investigação adicional para verificar os fatores subjacentes que contribuem para essas divergências.

## 6.2. Adequação do Japão a Lei de Benford

Entre as nações analisadas, o Japão apresentou o maior grau de conformidade, com um MAD de 0,107 calculado a partir de uma amostra de 1.142 números diários de mortalidade.

**Figura 04 - Distribuição da Lei de Benford para o Japão.**



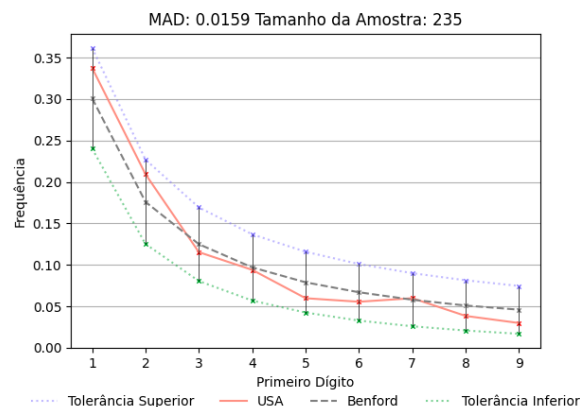
Fonte: autor.

Conforme ilustrado na Figura 4, observam-se pequenas divergências, especialmente para o dígito 7, que excedeu o limite superior de tolerância da curva teórica. Embora frequências individuais de dígitos que excedem os intervalos de confiança possam indicar imprecisões nos dados, o baixo valor do MAD sugere fortemente uma aderência estatisticamente significativa à LNB nos dados de mortalidade do Japão. Esta aderência corrobora estudos prévios que demonstram sua aplicabilidade a diversos conjuntos de dados naturais, incluindo fenômenos demográficos e epidemiológicos (Henselmann et al., 2012; Kolias, 2022; Menezes, 2021; Nigrini, 1994; Pires et al., 2022; Skousen et al., 2004).

## 6.3. Aderência dos casos semanais dos Estados Unidos

Na análise dos dados diários de mortalidade dos Estados Unidos, obteve-se uma taxa amostral de 235 semanas de óbitos registrados, como evidenciado na Figura 5. Calculou-se um desvio médio absoluto de, indicando uma adesão relativa à distribuição esperada dos dígitos.

**Figura 05 - Distribuição da Lei de Benford para os Estados Unidos da América.**



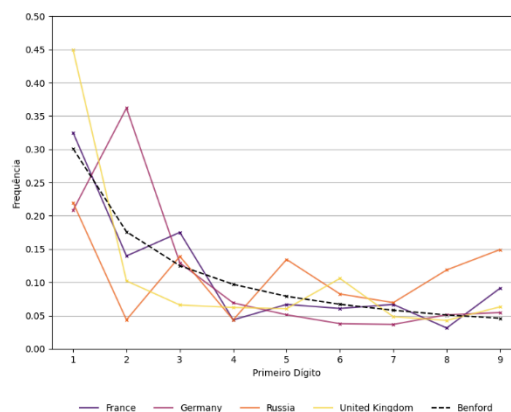
Fonte: autor

Ressalta-se que essa adesão está presente apesar do tamanho da amostra ser limitado, o que naturalmente implica em maior susceptibilidade a *outliers* e variabilidade, podendo impactar a precisão da análise. Esse fato reflete-se nos intervalos de tolerância superior (LS) e inferior (LI) mais amplos, que representam os limites estatísticos dentro dos quais a distribuição observada é considerada compatível com a LNB, destacando a limitação estatística inerente a essa análise.

## 6.4. Demais Países

Esta seção apresenta uma análise exploratória da distribuição de dígitos em dados de mortalidade por COVID-19 para países selecionados com base em sua relevância geopolítica como membros permanentes do Conselho de Segurança da ONU, cujos resultados estão disponíveis na Figura 6.

**Figura 06 - Distribuição da Lei de Benford para França, Alemanha, Rússia e Reino Unido.**



Fonte: autor.

A análise dos dados de MAD para França (0,0285), Alemanha (0,0442), Rússia (0,0594) e Reino Unido (0,0455) revela nuances interessantes na distribuição dos dígitos iniciais dos registros de mortalidade por COVID-19 em comparação com a LNB.

Na França, o dígito 1 apresenta a maior concentração (32,5%), consideravelmente acima do esperado pela LNB (30,1%). O dígito 3 também mostra uma frequência ligeiramente elevada (17,5% contra 12,5% esperados),

<sup>1</sup> Aluno de Iniciação Tecnológica (IT) com bolsa CNPq - PIBITI

enquanto os dígitos 4, 5, 7 e 8 ocorrem com menor frequência.

A Alemanha exibe a maior disparidade na distribuição dos dígitos, com o dígito 1 ocorrendo em 44,9% dos casos, um excesso notável em relação aos 30,1% esperados. Em contraste, o dígito 2 aparece em apenas 10,2% dos registros, significativamente abaixo dos 17,6% previstos.

A Rússia apresenta um padrão distinto, com uma distribuição mais equilibrada dos dígitos iniciais. O dígito 1 lidera com 22%, seguido de perto pelo dígito 3 com 13,9%. Nenhum dígito apresenta desvios tão acentuados quanto os observados na França e na Alemanha, sugerindo uma maior conformidade com a LNB.

No Reino Unido, a concentração no dígito 1 é a menor entre os países analisados (20,8%), mas ainda acima do esperado. O dígito 2 também ocorre com menor frequência (4,4%), enquanto o dígito 9 mostra uma frequência ligeiramente elevada (14,9%).

Em termos de disparidade, a Alemanha se sobressai com a maior variação entre as frequências de seus dígitos, seguida pela França. A Rússia apresenta a menor disparidade, indicando uma distribuição mais homogênea dos dígitos iniciais.

A comparação entre os países revela que França e Alemanha compartilham uma tendência de super-representação do dígito 1, embora em graus diferentes. Rússia e Reino Unido, por sua vez, apresentam padrões mais equilibrados, com menor concentração no dígito 1 e frequências mais próximas do esperado para os demais dígitos.

Em relação à LNB, todos os países analisados apresentaram desvios nas frequências dos dígitos iniciais. A Alemanha se destaca com o maior desvio, apresentando um grande excesso de 1 e um déficit acentuado de 2. A Rússia, por outro lado, exibe os menores desvios, sugerindo uma maior proximidade com a distribuição teórica.

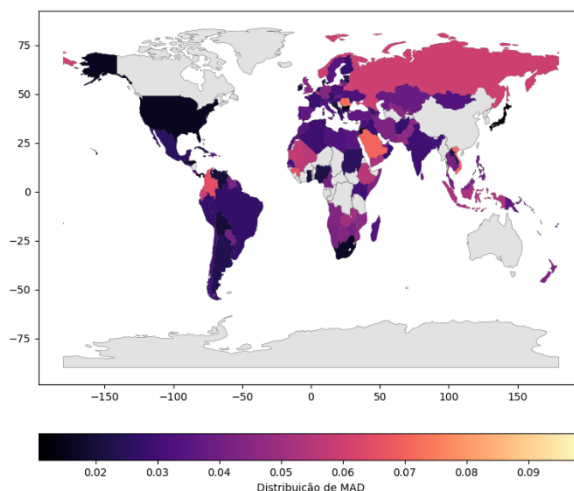
## 7. Análise Em Escala Global

Dada a robustez e efetividade do algoritmo desenvolvido, o presente trabalho apresenta uma análise global para avaliar a conformidade dos dados diários de mortalidade de diversos países com a LNB. Para cada país, calculou-se o MAD como medida quantificável dessa conformidade. Visando uma comparação relativa, os valores de MAD nas figuras 8 e 9 foram normalizados em uma escala de 0 a 1, onde 0 indica a maior e 1 a menor conformidade.

Os valores normalizados resultantes foram visualizados em um mapa global, empregando um mapa de cores para representar os diferentes graus de conformidade.

Essa visualização enfatiza a comparação proporcional entre países, e não a aderência absoluta à LNB. Portanto, o menor valor de MAD corresponde ao valor mais baixo na escala normalizada.

**Figura 07** - Análise em escala global da integridade dos dados de mortalidade dos países através da Lei de Benford visualizada pelo índice MAD.

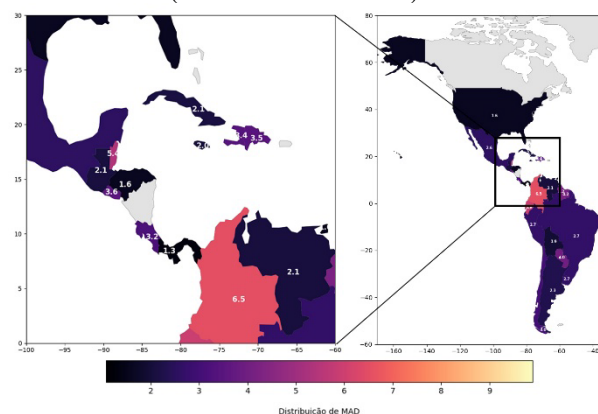


Fonte: autor.

Os índices MAD revelam variabilidade tanto entre continentes quanto dentro deles. A África e a Ásia apresentaram as maiores disparidades nos valores de MAD (Figura 7), com destaque para o contraste entre países como Japão (0.011) e Vietnã (0.071) na Ásia, e Gana (0.021) e Congo Brazzaville (0.069) na África. A Europa exibiu a menor variabilidade, com valores de MAD entre 0.011 (Montenegro) e 0.044 (Alemanha). O Brasil apresentou um MAD de 0.0268, indicando uma conformidade maior com a LNB quando comparado com outros países da América do Sul, como a Argentina.

Japão e Montenegro, ambos com MAD de 0.011, apresentaram os menores índices, sugerindo consistência nos dados (Nigrini, 2012). Em contrapartida, Vietnã (0.071) e Congo Brazzaville (0.069) exibiram os maiores índices, o que pode indicar potenciais problemas na qualidade dos dados. É importante ressaltar que mesmo o maior índice observado neste estudo (Vietnã, 0.071) não ultrapassa o índice do Consórcio de Veículos de Imprensa do Brasil (0.0871, Figura 2).

**Figura 08** - Análise das Américas dos dados de mortalidade da COVID-19 via Lei de Benford, utilizando o índice MAD (normalizado em  $10^2$ )



Fonte: autor.

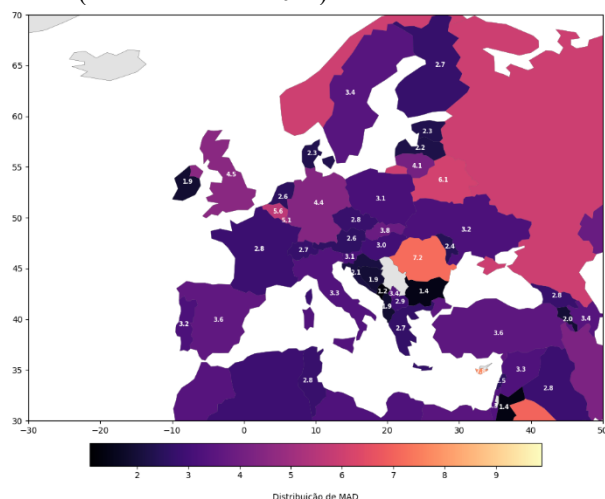
<sup>1</sup> Aluno de Iniciação Tecnológica (IT) com bolsa CNPq - PIBITI



Na América do Norte, a tendência de baixa variabilidade se mantém, conforme ilustrado na Figura 8. Os Estados Unidos apresentam um índice MAD de 0.0159, o COVID-19. Na América Central, predominam valores intermediários de MAD, como em Honduras (0.0159) e Guatemala (0.0208), com exceções notáveis como Barbados (0.0608). A América do Sul, em geral, também exibe baixa variabilidade, com o Brasil apresentando um índice MAD abaixo do esperado, mas com discrepâncias observadas na Argentina (0.0230) e Bolívia (0.0194).

O Caribe e a América Central demonstram a maior disparidade nos índices MAD, com Barbados (0.0608) e Bahamas (0.0576) apresentando valores que podem indicar não-conformidade com a LNB (Barney & Schulzke, 2016; Nigrini, 2012). Em contraste, os Estados Unidos mantêm um dos menores índices, o que sugere maior precisão nos registros.

**Figura 09** - Análise da Europa e os dados de mortalidade da COVID-19 via Lei de Benford, utilizando o índice MAD (normalizado em  $10^2$ ).



Fonte: autor.

A análise do Índice MAD nos países europeus (Figura 9) revela uma heterogeneidade considerável. Montenegro (0.012) e Bulgária (0.014) apresentam índices significativamente baixos, o que sugere parcial consistência em seus dados de mortalidade (Barney & Schulzke, 2016; Nigrini, 2012). Em contrapartida, Chipre (0.078) e Romênia (0.072) exibem os maiores valores de MAD, indicando potenciais anomalias ou inconsistências em seus registros. A maioria dos países europeus, como Áustria (0.026), Dinamarca (0.023) e Letônia (0.022), concentra-se em uma faixa que merece investigações detalhadas.

Em uma análise global, a distribuição dos valores de MAD indica uma baixa conformidade com a LNB (Barney & Schulzke, 2016; Menezes, 2021; Nigrini, 2012; Pires et al., 2022). Contudo, as disparidades regionais observadas anteriormente reforçam a

importância de se considerar o contexto geográfico na interpretação dos resultados.

## 6.5. Conclusões

A Lei de Benford (LNB), frequentemente aplicada para detectar anomalias em conjuntos de dados, revelou uma imagem complexa ao ser aplicada aos dados diários de mortalidade por COVID-19. A análise abrangendo diversos países e fontes de dados demonstra que a aderência à LNB não foi universal. Notavelmente, os dados do Consórcio de Veículos de Imprensa (CVI) no Brasil desviaram significativamente da LNB, levantando questões sobre sua precisão. Por outro lado, grande parte dos países, notavelmente o Japão, se alinharam de perto com a LNB, reforçando uma boa resposta em identificação e contagem de casos.

A análise global dos dados de mortalidade de 128 países mostrou uma conformidade geral moderada com a LNB, mas com disparidades regionais. África e Ásia apresentaram as maiores variações, enquanto a Europa demonstrou a menor. O Japão e Montenegro tiveram os menores índices de desvio, sugerindo alta consistência nos dados, enquanto Vietnã e Congo Brazzaville apresentaram os maiores, indicando possíveis problemas na qualidade dos dados. O Brasil, por sua vez, através dos dados do Ministério da Saúde, exibiu uma boa conformidade em comparação com outros países da América do Sul.

Essas descobertas têm implicações para a avaliação da qualidade dos dados e para a pesquisa epidemiológica. Em particular, as inconsistências encontradas nos dados do CVI destacam a importância de práticas robustas de coleta de dados durante crises de saúde pública. Além disso, o presente estudo contribui para a discussão sobre a utilidade da LNB em cenários do mundo real, enfatizando a necessidade de considerar fatores contextuais ao interpretar os resultados.

É importante reconhecer que este estudo teve limitações, incluindo variações nos tamanhos das amostras entre países e o potencial impacto de atrasos na notificação dos dados. Investigações futuras devem se concentrar em analisar e abordar essas limitações, aprofundando-se na compreensão das causas subjacentes às disparidades regionais na conformidade com a LNB. Em última análise, o presente estudo serve como um apelo para investigações contínuas sobre a qualidade dos dados epidemiológicos.

## Referências

- Barbosa, M., Goulart De Andrade, A. P., & André Oliveira Cony, M. (2022). Cenas dos próximos capítulos: A criação do consórcio de imprensa. *Razón y Palabra*, 26(114). <https://doi.org/10.26807/rp.v26i114.1947>
- Barney, B. J., & Schulzke, K. S. (2016). Moderating “Cry Wolf” Events with Excess MAD in Benford’s Law Research and Practice. *Journal of Forensic Accounting Research*, 1(1), A66–A90. <https://doi.org/10.2308/jfar-51622>
- Benford, F. (1938). The Law of Anomalous Numbers. *Proceedings of the American Philosophical Society*, 78(4), Artigo 4.

<sup>1</sup> Aluno de Iniciação Tecnológica (IT) com bolsa CNPq - PIBITI

Cleary, R., & Thibodeau, J. C. (2005). Applying Digital Analysis Using Benford's Law to Detect Fraud: The Dangers of Type I Errors. *AUDITING: A Journal of Practice & Theory*, 24(1), Artigo 1. <https://doi.org/10.2308/aud.2005.24.1.77>

Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*, 20(5), Artigo 5. [https://doi.org/10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1)

Dos Santos, K. D. N. (2022). A luta entre o Poder Político e o Quarto Poder em tempos de pandemia da COVID-19 no Brasil / The struggle between Political Power and the Fourth Estate in times of the COVID-19 pandemic in Brazil. *Brazilian Journal of Development*, 8(2), 11341–11360. <https://doi.org/10.34117/bjdv8n2-191>

Drake, P. D., & Nigrini, M. J. (2000). Computer assisted analytical procedures using Benford's Law. *Journal of Accounting Education*, 18(2), Artigo 2. [https://doi.org/10.1016/S0748-5751\(00\)00008-7](https://doi.org/10.1016/S0748-5751(00)00008-7)

Druică, E., Oancea, B., & Vâlsan, C. (2018). Benford's law and the limits of digit analysis. *International Journal of Accounting Information Systems*, 31, 75–82. <https://doi.org/10.1016/j.accinf.2018.09.004>

Günther, L. C., Colangelo, E., Wiendahl, H.-H., & Bauer, C. (2019). Data quality assessment for improved decision-making: A methodology for small and medium-sized enterprises. *Procedia Manufacturing*, 29, 583–591. <https://doi.org/10.1016/j.promfg.2019.02.114>

Henselmann, K., Scherr, E., & Ditter, D. (2012). Applying Benford's Law to individual financial reports: An empirical investigation on the basis of SEC XBRL filings.

Kolias, P. (2022). Applying Benford's law to COVID-19 data: The case of the European Union. *Journal of Public Health*, 44(2), e221–e226. <https://doi.org/10.1093/pubmed/fdac005>

Kössler, W., Lenz, H.-J., & Wang, X. D. (2021). Is the Benford Law Useful for Data Quality Assessment? Em S. Knoth & W. Schmid (Orgs.), *Frontiers in Statistical Quality Control 13* (p. 391–406). Springer International Publishing. [https://doi.org/10.1007/978-3-030-67856-2\\_22](https://doi.org/10.1007/978-3-030-67856-2_22)

Kössler, W., Lenz, H.-J., & Wang, X. D. (2024). Some new invariant sum tests and MAD tests for the assessment of Benford's law. *Computational Statistics*. <https://doi.org/10.1007/s00180-024-01463-8>

Menezes, R. O. (2021). Aplicação da Lei de Benford nos números de casos confirmados de COVID-19 em diferentes países. *REMAT: Revista Eletrônica da Matemática*, 7(1), e3005. <https://doi.org/10.35819/remat2021v7i1id4586>

Newcomb, S. (1881). Note on the Frequency of Use of the Different Digits in Natural Numbers. *American Journal of Mathematics*, 4(1), Artigo 1. <https://doi.org/10.2307/2369148>

Nigrini, M. (1994). Using digital frequencies to detect fraud. *The white paper*, 8(2), Artigo 2.

Nigrini, M. J. (2012). *Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection*. John Wiley & Sons.

Pires, R. C., Andrade, H. D. S., Duarte, A. C. R., & Borges, D. (2022, outubro 10). *Uso da Lei de Benford*

*para investigação de anomalias quantitativas nos números da COVID-19 nos estados brasileiros*. ENEGEP 2022 - Encontro Nacional de Engenharia de Produção, FOZ DO IGUAÇU/SP - BRASIL. [https://doi.org/10.14488/ENEGEP2022\\_TN\\_ST\\_384\\_1\\_900\\_43276](https://doi.org/10.14488/ENEGEP2022_TN_ST_384_1_900_43276)

Roncagli, A., & Lacerda, J. (2020). *Jornalismo e conhecimento: A divergência dos dados da covid-19 divulgados via imprensa nacional e SESAP-RN*. <https://doi.org/10.1590/SciELOPreprints.1141>

Skousen, C. J., Guan, L., & Wetzell, T. S. (2004). Anomalies and Unusual Patterns in Reported Earnings: Japanese Managers Round Earnings. *Journal of International Financial Management & Accounting*, 15(3), Artigo 3. <https://doi.org/10.1111/j.1467-646X.2004.00108.x>

Tošić, A., & Vičić, J. (2021). Use of Benford's law on academic publishing networks. *Journal of Informetrics*, 15(3), 101163. <https://doi.org/10.1016/j.joi.2021.101163>

Wang, L., & Ma, B.-Q. (2024). A concise proof of Benford's law. *Fundamental Research*, 4(4), Artigo 4. <https://doi.org/10.1016/j.fmre.2023.01.002>

### Anexos

A. CDC. (2020, março 28). *COVID Data Tracker*. Centers for Disease Control and Prevention. <https://covid.cdc.gov/covid-data-tracker>. Acesso em: 14/08/2024.

B. CSSEGISandData. (2024). *CSSEGISandData/COVID-19* [Software]. <https://github.com/CSSEGISandData/COVID-19> (Trabalho original publicado 2020). Acesso em: 14/08/2024.

C. G1. (2024). *Dados epidemiológicos da COVID-19 obtidos pelo Consórcio de Veículos de Imprensa*. <https://especiais.g1.globo.com/bemestar/coronavirus/dados/semaforo/graficos/mortesporidiamedia/brasil.csv>. Acesso em: 14/08/2024.

D. *Coronavirus Brasil*. ([s.d.]). <https://covid.saude.gov.br/>. Acesso em: 14/08/2024.

### Apêndices

A. Código-fonte do Projeto. Trigueiro, B. (2024). *Koobzaar/Lutero* [HTML]. <https://github.com/koobzaar/Lutero> (Trabalho original publicado 2024)

### Agradecimentos

À Maylis Bernardino Bezerra Trigueiro: quando você chegar a um dos muitos momentos da vida em que precisar refletir sobre si mesma, fornecer um relato do que foi, do que fez e do que significou para o mundo, peça – peça não, oro – para que não se esqueça de que você preenche os dias do seu tio com uma alegria plena, uma alegria que me foi desconhecida em todos os meus anos passados.

... à Kethelyn Bezerra Trigueiro Leite, Madlyn Bezerra Trigueiro Leite, José Trigueiro Leite e Cícera Aero Riuza Bezerra Guerra, pelo apoio e amor incondicional.

<sup>1</sup> Aluno de Iniciação Tecnológica (IT) com bolsa CNPq - PIBITI

... ao Me. Lucas Yuri Saladini, pela companhia em dias mal-dormidos, revisão metodológica e gramatical.

... ao Guto, meu orientador, por ter encontrado a dose perfeita entre cobrança e liberdade para extrair meu desempenho máximo.

... à Fatec Baixada Santista Rubens Lara, pela oportunidade de realizar minha primeira IT.

... à memória de Fernando Gonzales Tavares.