Brazilian Journal of Biometrics, **42**, 78-87 (2024)
doi:10.28951/bjb.v42i1.659

BRAZILIAN JOURNAL OF
BIOMΣTRICS
ISSN:2764-5290

ARTICLE

A novel approach to evaluate data integrity: evidence from COVID-19 in China

Lucas Emanuel de Oliveira Silva*[1] and Dalson Figueiredo[2]

[1]Postgraduate Program in Medical Sciences, Federal University of Alagoas, Maceió - AL, Brazil
[2] Department of Political Science, Federal University of Pernambuco, Recife - PE, Brazil
*Corresponding author. Email: lucas.silva@academico.uncisal.edu.br

(Received: February 23, 2023; Revised: July 3, 2023; Accepted: July 12, 2023; Published: March 15, 2024)

Abstract

The COVID-19 pandemic has generated an unprecedented amount of epidemiological data. Yet, concerns regarding the validity and reliability of the information reported by health surveillance systems have emerged worldwide. In this paper, we develop a novel approach to evaluating data integrity by combining the Newcomb-Benford Law with outlier methods. We demonstrate the advantages of our framework using a case study from China. To ensure more robust findings, we employ multiple diagnostic procedures, including three conformity estimates, four goodness-of-fit tests, and two distance measures (Cook and Mahalanobis). To promote transparency, we have made all computational scripts publicly available. Our findings indicates a significant deviation in the distribution of new deaths from the theoretical expectations of Benford's Law. Importantly, these results remain accurate even when considering alternative model specifications and conducting various statistical tests. Furthermore, the procedures developed here are easily applicable in other areas of knowledge and can be scaled to assess data quality in both the public and private sectors.

Keywords: COVID-19; Newcomb-Benford Law; Data quality; Epidemiological surveillance; China.

1. Introduction

Despite having a population of over 1.4 billion people, China has not experienced the same level of impact from the COVID-19 pandemic as some other countries. According to Our World In Data, as of January 20, 2023, China has reported approximately 2 million cases and over 5,200 deaths (Mathieu *et al.*, 2020). When compared to countries with similar territorial and social dynamics, the disparity becomes evident. For instance, China's prevalence rate is 30 times lower than that of India (Mathieu *et al.*, 2020).

This situation has prompted concerns within the international medical community regarding the reliability of information provided by Chinese authorities (Power, 2023; Stanway & Lapid, 2022). In response, the World Health Organization (WHO) has made explicit requests to China for detailed information regarding outpatient clinics, hospitalizations, emergency and intensive

---

care treatment, as well as COVID-19-related hospital deaths (WHO, 2023).

In this paper, we develop a novel framework to evaluate data integrity by combining the Newcomb-Benford Law with outlier diagnosis methods. We show the advantages of our approach using a case study from COVID-19 data in China. To ensure the accuracy of our findings, we employ three methods for measuring conformity (mean absolute deviation, distortion factor, and mantissa) and conduct four different tests to validate our results (Pearson chi-square, Kolmogorov-Smirnov D statistic, Chebyshev distance m statistic, and the Euclidean distance), in addition to two distance measures (Mahalanobis and Cook).

The NBL has found applications in various fields, including Economics (Kaiser, 2019), Political Science (Deckert *et al.*, 2011; Mebane, 2011), Health (Balashov *et al.*, 2021; Figueiredo Filho *et al.*, 2022; Silva & Figueiredo Filho, 2020) and Natural Sciences (Sambridge & Jackson, 2020) (Sambridge *et al.*, 2010). It has been used in forensic accounting (Nigrini, 2012) and fraud detection (Durtschi *et al.*, 2004). Scholars have also explored its use in assessing the integrity of campaign finance data (Cho & Gaines, 2007) and electoral outcomes (Pericchi & Torres, 2011). There are also applications to academic dishonesty (Horton *et al.*, 2020) and international trade (Cerioli *et al.*, 2018; Lacasa & Fernández-Gracia, 2019). Interestingly, even the citations to articles citing Benford Law follow a NBL distribution (Mir, 2014). Hill (1995) developed a rigorous proof of the law and Nigrini (2012) provides the most comprehensive review of both theory and applications of Benford Law in empirical research.

Outlier detection has been a historically significant field of study in statistics (Figueiredo Filho *et al.*, 2023). Its objective is to identify observations that exhibit substantial deviations from the expected theoretical patterns or trends in the data. Referred to as aberrant cases or anomalies, outliers can occur due to various factors such as measurement errors, sudden environmental changes, data entry mistakes, and even deliberate malicious actions (Hodge & Austin, 2004). To the best of our knowledge, there has been no prior use of the joint application of NBL and outlier techniques to assess data integrity. This paper advances our understanding of the role of statistical tools in assessing the accuracy and reliability of data reported by health surveillance systems.

The remainder of the paper is structured as follows: the next section describes the materials and methods. Then we present the main findings. The fourth section interprets the results, and the final section concludes.

2. Materials and Methods

The Newcomb-Benford Law (NBL), also known as the First-Digit Law, is widely employed as a mechanism to identify atypical patterns in data (Figueiredo Filho *et al.*, 2022; Nigrini, 2012). It accomplishes this by comparing the distribution of numbers in the data to the expected theoretical distribution (Hill, 1995). More technically, the exact distribution for the NBL for the first digit is given by:

$$P(d) = log10\ (1 + \frac{1}{d})\ for\ d \in \{1, \dots ,9\} \quad (1)$$

In accordance with the Newcomb-Benford Law (NBL), specific digits exhibit a higher likelihood of appearing as the first digit in a set of numbers (Benford, 1938; Newcomb, 1881). For instance, the digit 1 emerges as the most frequent first digit, occurring approximately 30% of the time. Conversely, the digit 9 stands as the least common first digit, appearing in less than 5% of cases. Table 1 shows the expected frequency of the first, second and third digits.

Table 1. Probability of occurrence of the first, second and third digits

Digit	First digit	Second digit	Third digit
0		12	10.2
1	30.1	11.4	10.1
2	17.6	10.9	10.1
3	12.2	10.4	10.1
4	9.7	10	10
5	7.9	9.7	10
6	6.7	9.3	9.9
7	5.8	9	9.9
8	5.1	8.8	9.9
9	4.6	8.5	9.8

Source: Nigrini (2012).

We used country level information from the Our World in Data (Mathieu et al., 2020), which is updated daily and provides accurate figures on COVID-19 confirmed cases, deaths, testing, and vaccinations. We collected time-series data from January 22, 2020 to December 31, 2022.

Following the standard approach in the scholarly literature (Cho & Gaines, 2007), we compare the observed frequency of the first digit in the daily number of cases and deaths in China and the expected frequency based on the NBL. Furthermore, we employ three different methods, namely mean absolute deviation, distortion factor, and mantissa, to assess the degree of conformity to the NBL. Additionally, we apply four statistical tests, including Pearson chi-square, Kolmogorov-Smirnov D statistic, Chebyshev distance m statistic, and Euclidean distance. The rationale behind employing multiple tests is to enhance the reliability of the results. The theoretical foundations of these tests are explained in Nigrini (2012).

We used R Statistical 4.0.5 to perform all of the computational analyses and our significance tests were two-sided. We utilized the benford.analysis package developed by Cinelli (2018) and the BenfordTests package developed by Joenssen & Muellerleile (2015) to run our NBL model. For the purpose of replication, we have made the raw data and computational scripts readily available on <https://osf.io/vtfsz/>.

3. Results

Table 2 summarizes both conformity estimates and goodness of fit tests for COVID-19 cases and deaths in China from 2020 to 2022.

Table 2. NBL goodness of fit and conformity tests for Chinese COVID-19 new cases and deaths (2020-2022)
chi-square = Pearson chi-square; ks = Kolmogorov-Smirnov D statistic; md = Chebyshev distance; mantissa = Average mantissa; mad = Mean absolute deviation; mad conformity = Conformity classification according to mad value; df = Distortion factor

Variable	Year	N	Conformity estimates				Goodness of fit tests		
			mantissa	mad	mad conformity	df	chi-square	ks	md
Cases	2020	342	0.50	0.03	Nonconformity	-10.99	32.13 ($p < 0.001$)	1.55 ($p < 0.001$)	1.55 ($p < 0.001$)
Cases	2021	365	0.51	0.03	Nonconformity	-1.60	37.28 ($p < 0.001$)	1.35 ($p < 0.01$)	1.35 ($p < 0.001$)
Cases	2022	361	0.48	0.01	Marginally acceptable conformity	-2.32	13.25 ($p = 0.1$)	0.70 ($p = 0.36$)	0.58 ($p = 0.4$)
Deaths	2020	127	0.33	0.04	Nonconformity	-17.90	22.34 ($p < 0.001$)	11.74 ($p < 0.001$)	2.31 ($p < 0.001$)
Deaths	2021	109	0.11	0.09	Nonconformity	-	103.97 ($p < 0.001$)	13.40 ($p < 0.001$)	2.37 ($p < 0.001$)
Deaths	2022	110	0.17	0.09	Nonconformity	-21.00	92.58 ($p < 0.001$)	13.35 ($p < 0.001$)	2.89 ($p < 0.001$)

Note: *In the Mantissa Arc Test, the distortion factor is a statistical parameter that varies depending on the magnitude of the data. In the year 2021, there was a lot of repetition of zero daily deaths. This resulted from the statistical function a NaN value return of the df in that year.*

To ensure a strong correspondence between a distribution and the theoretical predictions of the NBL, specific characteristics must be met. These include an average mantissa of .5, a variance of 1/12, and a skewness close to zero (Shao & Ma, 2010). According to Nigrini (2012), a mean absolute deviation (MAD) exceeding .015 suggests a deviation from the NBL and a lack of alignment in the first digit test. Another useful measure is the distortion factor, which provides insights into potential tampering with the data, whether in an upward or downward direction (Nigrini, 2012).

In the context of goodness-of-fit tests, statisticians often employ the chi-square test to compare an observed distribution with a theoretical distribution. In this case, the null hypothesis assumes that the data adheres to the NBL. Another commonly used test is the Kolmogorov-Smirnov test (K-S), which is based on the cumulative density function and is sensitive to the sample size. It is worth noting that the K-S test is particularly useful for assessing the fit between the observed and theoretical distributions. Additionally, the Chebyshev distance test, similar to the Euclidean distance test, is utilized to evaluate the similarity between two probability distributions. This test, as described by Druica *et al.* (2018), offers another approach to investigate the degree of agreement between the distributions under examination.

In terms of the frequency of new COVID-19 cases, conformity estimates based on mean absolute deviation (MAD) and distortion factor (DF) suggest nonconformity for the years 2020 and 2021, with marginally acceptable conformity observed for 2022. All goodness of fit tests indicate that we should reject the null hypothesis, which assumes that the observed distribution adheres to the NBL.

Figure 1 shows the distribution of the first digit of the number of new COVID-19 cases and deaths. The observed distribution deviates from the expected distribution based on the NBL. Specifically, the data for deaths shows a significant deviation from the NBL. The distortion factor, a measure proposed by Nigrini (2012), indicates that both the cases and deaths are underreported.

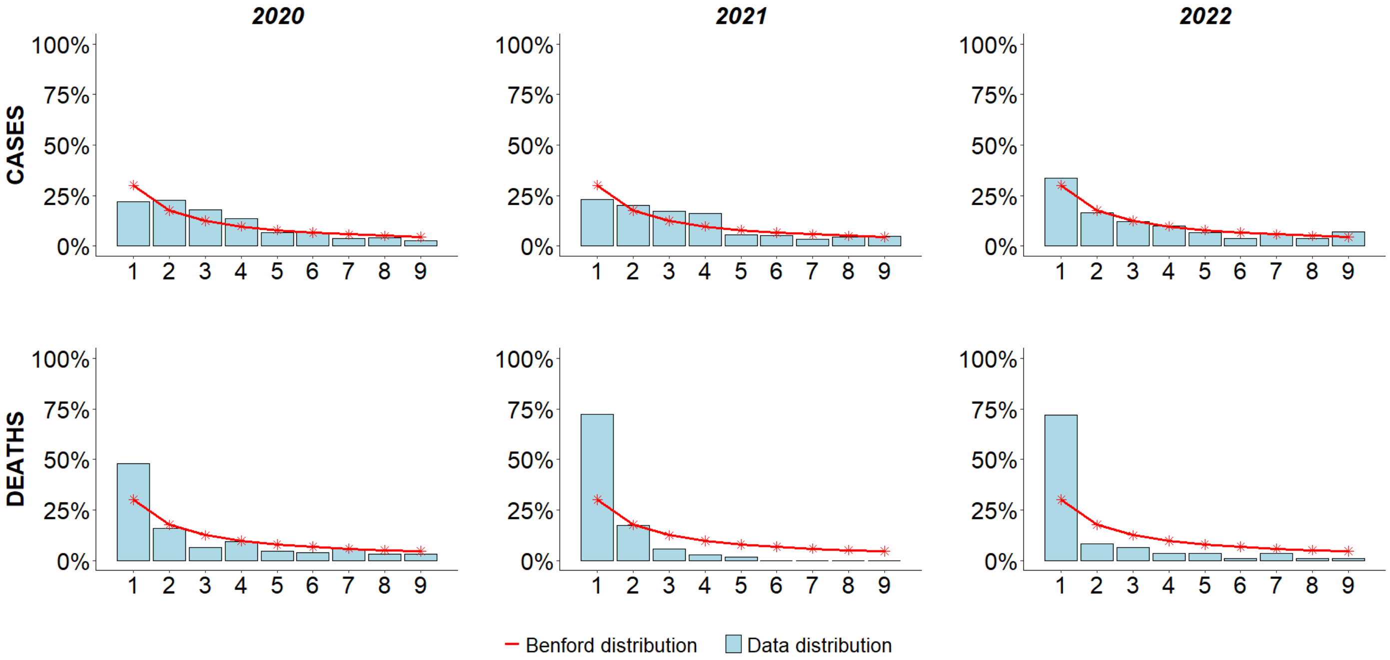

Figure 1. NBL distribution of the first digit for COVID-19 new cases and deaths in China (2020-2022).

In order to strengthen the validity of our findings, we performed several supplementary tests. One of these tests involved assessing the distribution of the second significant digit in the number of new cases and deaths recorded in China. The outcomes of this analysis are presented in Figure 2.

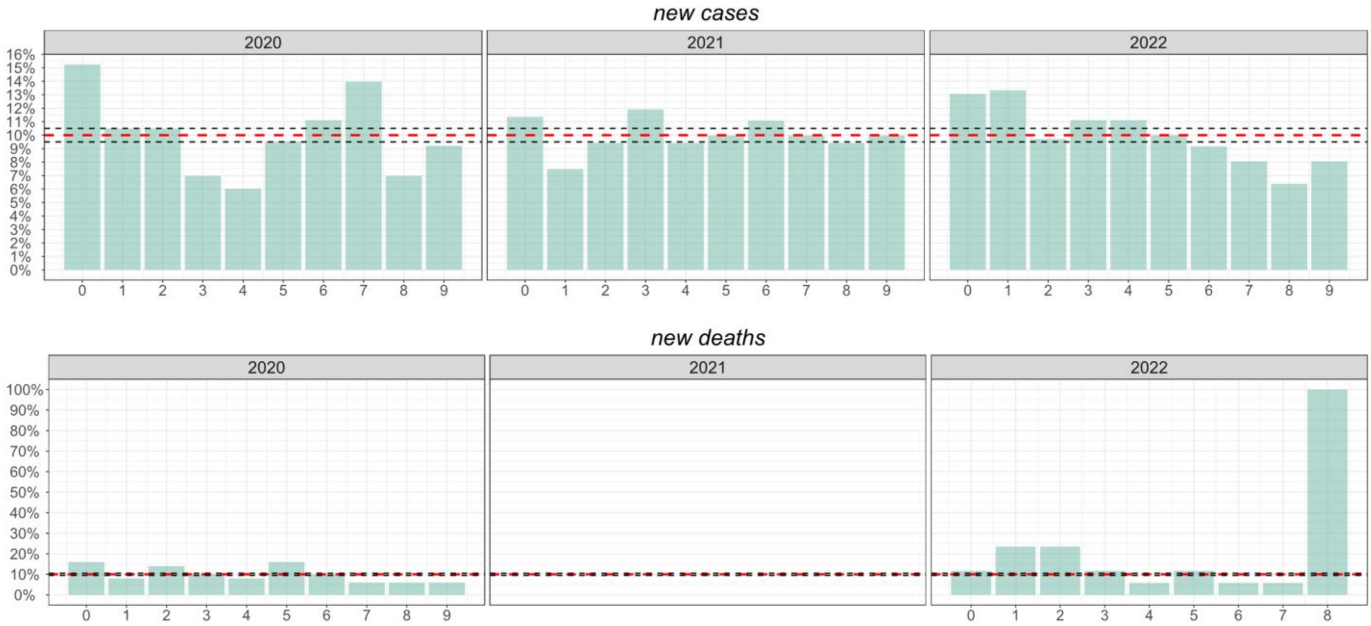

Figure 2. Second digit distribution in COVID-19 records for new cases and deaths in China (2020-2022).

Both distributions exhibit noteworthy deviations from what is theoretically expected. Hicken & Mebane (2015) state that the anticipated mean for the second digit is 4.187. However, in 2020, the mean for new COVID-19 cases in China was 4.273. In 2021, the average rose to 4.510, and in 2002, it dropped to 3.956, indicating substantial variability. Concerning deaths, the observed averages were 3.820 in 2020, zero in 2021, and 3.235 in 2022. The direct comparison between cases and deaths indicate that the integrity of COVID-19 mortality data is poor.

The next step involved fitting a linear regression model to estimate the number of deaths using the number of new cases and the population size as explanatory variables. Subsequently, we examined the residuals in Figure 3.

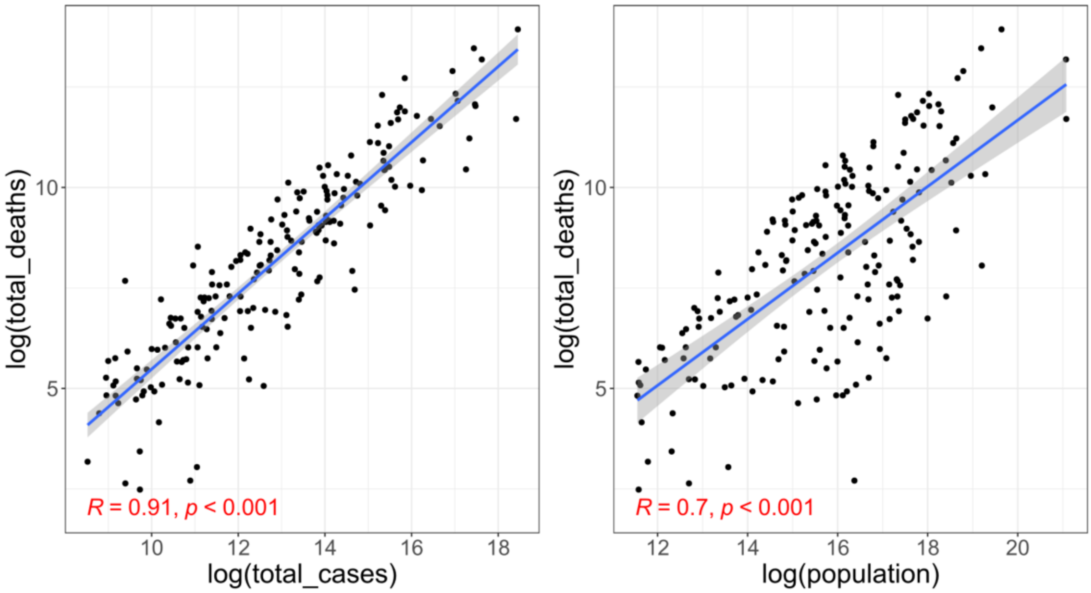

Figure 3. COVID-19 cases and deaths in China (on May 3, 2023).

As expected, we find a positive and strong association between population, number of COVID-19 cases and deaths. In particular, the correlation between total number of cases and deaths is .91 (p-value < 0.01) and the relationship between population size and mortality is .7 (p-value < .01).

Upon conducting an analysis of the residuals, we observed the presence of abnormal cases. To further investigate this, we computed the Mahalanobis distance for COVID-19 cases, deaths and population variables across a sample of 194 countries. This multivariate estimate was combined with Cook's distance and leverage values, resulting in the creation of a new dataset. Figure 4 presents the final results of this analysis.

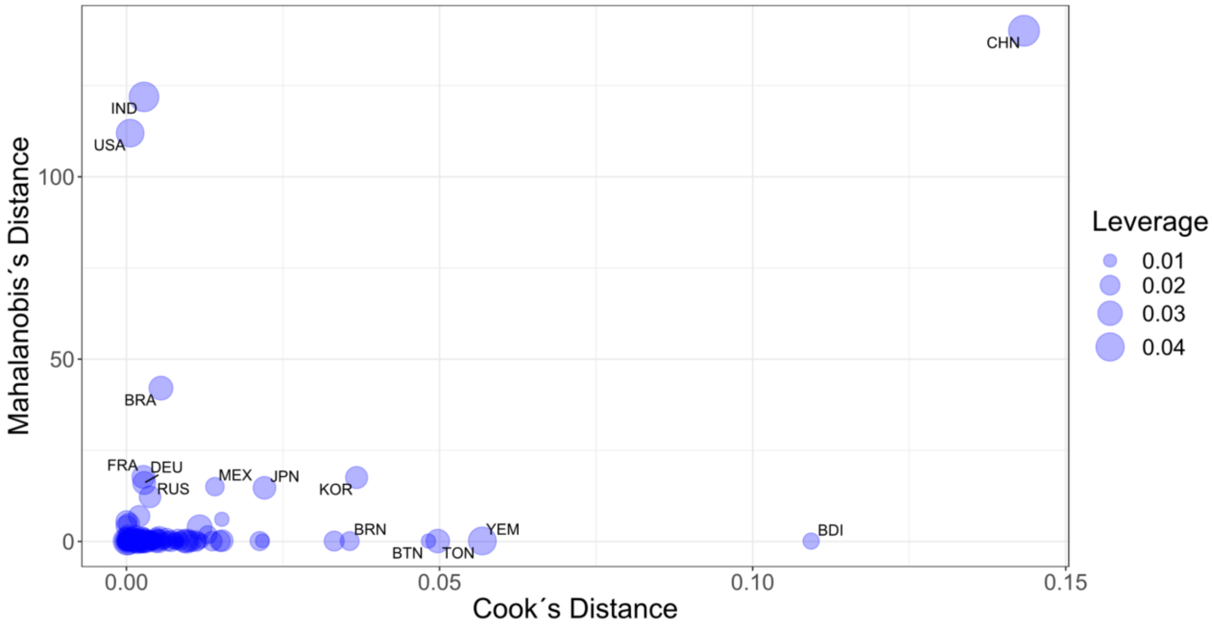

Figure 4. Mahalanobis distance, Cook distance and Leverage for COVID-19 data in China.

The results consistently demonstrate that China stands out as an abnormal observation within the dataset, regardless of the statistical criteria employed to measure this distinction. With an exceptionally high score in Mahalanobis distance, China emerges as a multivariate outlier, displaying unusual values in terms of population, COVID-19 cases, and deaths. Furthermore, the magnitude of the Cook's distance indicates a stronger influence of individual data points on the regression model, suggesting that the inclusion of China would lead to biased inferences. Finally, the high leverage further confirms that China is an observation that deviates significantly from the rest of the sample, indicating that it is unlikely to be generated by the same process. In summary, we have strong reservations regarding the reliability of epidemiological data in China, to say the least.

4. Discussion

Wu *et al*. (2020) have highlighted several challenges associated with big data in public health, including lack of quality, delays in data sharing and concerns regarding privacy protection. In particular, the WHO has raised concerns about China's lack of transparency and failure to share essential data regarding the origin of COVID-19 (Cohen, 2023). The absence of crucial information poses significant challenges to scientific research and hampers the implementation of effective measures for pandemic prevention and control (Cohen, 2023). Alvarez *et al*. (2023) argue that during the initial stages of the pandemic, the recording of cases in the system was subject to approval from local government officials. Additionally, Mi *et al*. (2020) have identified that the fatality rate of COVID-19 in China may have been underestimated just due to this limited inclusion of cases in the early stages.

Situations of this nature give rise to a series of inquiries concerning the veracity of information disseminated by governmental entities to researchers and the general population. Within this context, forensic statistical methods are frequently employed, with the NBL being the predominant procedure used to identify inconsistencies or irregular patterns within the data (Cho

& Gaines, 2007; Figueiredo Filho *et al.*, 2022; Formann, 2010). Investigations into the manipulation of COVID-19 data have been undertaken in both advanced nations such as the United States (Campolieti, 2021) and the European Union (Kolias, 2022), as well as in developing countries like Brazil (Silva & Figueiredo Filho, 2020) and India (Natashekara, 2022). Additional studies have demonstrated that autocratic regimes are more prone to data manipulation when compared to more democratic societies (Annaka, 2021; Balashov *et al.*, 2021; Neumayer & Plümper, 2022).

Contrary to all these technical limitations in data quality provision, Idrovo & Manrique-Hernández (2020) conducted a study that examined the integrity of Chinese data on COVID-19 using the NBL. They gathered information from situation reports and applied first-digit tests to the cumulative cases reported by Chinese provinces, regions, and cities. They found that China's epidemiological surveillance system had an acceptable level of quality and suggested that further systematic analysis could provide a more comprehensive evaluation of China's healthcare services performance. In same view, Ivorra *et al.* (2020) have employed a new θ-SEIHRD model taking into account undetected infections and showed a good agreement between the reported Chinese data and the estimations given by their model.

Despite these findings, Kennedy & Yam (2020) found that Chinese COVID-19 data deviated from the NBL, including both the first and second digits. They proposed three key factors to explain the outcome: a) China being the epicenter of the outbreak which would result in a backlog of reported cases and deaths, b) the lack of reliable tests to detect COVID-19 infection at the beginning of the epidemic, and c) strong intervention by the central government.

5. Conclusions

Our investigation reveals compelling evidence suggesting the compromised integrity of Chinese COVID-19 data. Specifically, the reported cases and death counts exhibit substantial deviations from the expected patterns predicted by the NBL. These results hold robust even when considering alternative conformity estimates, goodness of fit tests and conducting various statistical procedures. Furthermore, residual analysis shows that China does not follow the same data generating process that explains the remainder of the sample. This study contributes to our understanding of the application of forensic tools in evaluating the reliability and validity of epidemiological data.

Acknowledgments

Enivaldo Rocha, retired professor of the Political Science Department of the Federal University of Pernambuco (UFPE), and the anonymous reviewers of the Brazilian Journal of Biometrics for their useful comments.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Alvarez, E., Bielska, I. A., Hopkins, S., Belal, A. A., Goldstein, D. M., Slick, J., Pavalagantharajah, S., Wynfield, A., Dakey, S., Gedeon, M.-C., Alam, E. & Bouzanis, K. Limitations of COVID-19 testing and case data for evidence-informed health policy and practice. *Health Research Policy and Systems* **21** (1), 11 (2023). https://doi.org/10.1186/s12961-023-00963-1

2. Annaka, S. Political regime, data transparency, and COVID-19 death cases. *SSM - Population Health* **15**, 100832 (2021). https://doi.org/10.1016/j.ssmph.2021.100832

3. Balashov, V. S., Yan, Y. & Zhu, X. Using the Newcomb–Benford law to study the association between a country’s COVID-19 reporting accuracy and its development. *Scientific Reports* **11** (1), 22914 (2021).

https://doi.org/10.1038/s41598-021-02367-z

4. Benford, F. The Law of Anomalous Numbers. *Proceedings of the American Philosophical Society* **78** (4), 551–572 (1938).

5. Campolieti, M. COVID-19 deaths in the USA: Benford's law and under-reporting. *Journal of Public Health* (2021) fdab161. https://doi.org/10.1093/pubmed/fdab161

6. Cerioli, A., Barabesi, L., Cerasa, A., Menegatti, M. & Perrotta, D. Newcomb–Benford law and the detection of frauds in international trade. *Proceedings of the National Academy of Sciences* **116**, 201806617 (2018). https://doi.org/10.1073/pnas.1806617115

7. Cho, W. K. T. & Gaines, B. J. Breaking the (Benford) Law. *The American Statistician* **61** (3), 218–223 (2007). https://doi.org/10.1198/000313007X223496

8. Cinelli, C. *Benford.analysis: Benford Analysis for Data Validation and Forensic Analytics* (0.1.5) (2018). https://CRAN.R-project.org/package=benford.analysis

9. Cohen, J. *'It's inexcusable.' WHO blasts China for not disclosing potential data on COVID-19's origin* [Data set] (2023). https://doi.org/10.1126/science.adh8529

10. Deckert, J., Myagkov, M. & Ordeshook, P. C. Benford's Law and the Detection of Election Fraud. *Political Analysis* **19** (3), 245–268 (2011). https://doi.org/10.1093/pan/mpr014

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

12. Durtschi, C., Hillison, W. & Pacini, C. *The effective use of Benford's Law to assist in detecting fraud in accounting data*. (2004) https://digitalcommons.usf.edu/fac_publications/939

13. Filho, D. F., Silva, L., Pires, A., & Malaquias, C. Living with outliers: How to detect extreme observations in data analysis. *BIB - Revista Brasileira de Informação Bibliográfica Em Ciências Sociais*, **99,** 1-24 (2023). https://bibanpocs.emnuvens.com.br/revista/article/view/619

14. Figueiredo Filho, D., Silva, L. & Medeiros, H. "Won't get fooled again": Statistical fault detection in COVID-19 Latin American data. *Globalization and Health* **18** (1), 105 (2022). https://doi.org/10.1186/s12992-022-00899-1

15. Formann, A. K. The Newcomb-Benford Law in Its Relation to Some Common Distributions. *PLOS ONE* **5**(5), e10541 (2010). https://doi.org/10.1371/journal.pone.0010541

16. Hicken, A. & Mebane, W. *A Guide to Election Forensics* (2015). https://pdf.usaid.gov/pdf_docs/PA00MXR7.pdf

17. Hill, T. P. Base-Invariance Implies Benford's Law. *Proceedings of the American Mathematical Society* **123** (3), 887–895 (1995). https://doi.org/10.2307/2160815

18. Hodge, V. & Austin, J. A survey of outlier detection methodologies. *Artificial Intelligence Review* **22** (2), 85–126 (2004).

19. Horton, J., Krishna Kumar, D. & Wood, A. Detecting academic fraud using Benford law: The case of Professor James Hunton. *Research Policy* **49** (8), 104084 (2020). https://doi.org/10.1016/j.respol.2020.104084

20. Idrovo, A. J. & Manrique-Hernández, E. F. Data Quality of Chinese Surveillance of COVID-19: Objective Analysis Based on WHO's Situation Reports. *Asia Pacific Journal of Public Health* **32** (4), 165–167 (2020). https://doi.org/10.1177/1010539520927265

21. Ivorra, B., Ferrández, M. R., Vela-Pérez, M. & Ramos, A. M. Mathematical modeling of the spread of the coronavirus disease 2019 (COVID-19) taking into account the undetected infections. The case of China.

Communications in Nonlinear Science and Numerical Simulation **88**, 105303 (2020). https://doi.org/10.1016/j.cnsns.2020.105303

22. Joenssen, D. W. & Muellerleile, T. *BenfordTests: Statistical Tests for Evaluating Conformity to Benford's Law* (1.2.0) (2015). https://CRAN.R-project.org/package=BenfordTests

23. Kaiser, M. Benford's Law as an Indicator of Survey Reliability—Can We Trust Our Data? *Journal of Economic Surveys* **33** (5), 1602–1618 (2019). https://doi.org/10.1111/joes.12338

24. Kennedy, A. P. & Yam, S. C. P. On the authenticity of COVID-19 case figures. *PLOS ONE* **15**(12), e0243123 (2020). https://doi.org/10.1371/journal.pone.0243123

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

26. Lacasa, L. & Fernández-Gracia, J. Election Forensics: Quantitative methods for electoral fraud detection. *Forensic Science International* **294**, e19–e22 (2019). https://doi.org/10.1016/j.forsciint.2018.11.010

27. Mathieu, E., Ritchie, H., Rodés-Guirao, L., Appel, C., Giattino, C., Hasell, J., Macdonald, B., Dattani, S., Beltekian, D., Ortiz-Ospina, E., & Roser, M. Coronavirus Pandemic (COVID-19). *Our World in Data* (2020). https://ourworldindata.org/coronavirus

28. Mebane, W. R. Comment on "Benford's Law and the Detection of Election Fraud." *Political Analysis* **19** (3), 269–272 (2011). https://doi.org/10.1093/pan/mpr024

29. Mi, Y., Huang, T., Zhang, J., Qin, Q., Gong, Y., Liu, S., Xue, H., Ning, C., Cao, L. & Cao, Y. Estimating the instant case fatality rate of COVID-19 in China. *International Journal of Infectious Diseases* **97**, 1–6 (2020). https://doi.org/10.1016/j.ijid.2020.04.055

30. Mir, T. A. The Benford law behavior of the religious activity data. *Physica A: Statistical Mechanics and Its Applications* **408**, 1–9 (2014). https://doi.org/10.1016/j.physa.2014.03.074

31. Natashekara, K. COVID-19 cases in India and Kerala: A Benford's law analysis. *Journal of Public Health* **44**(2), e287–e288 (2022). https://doi.org/10.1093/pubmed/fdab199

32. Neumayer, E. & Plümper, T. (2022). Does 'Data fudging' explain the autocratic advantage? Evidence from the gap between Official Covid-19 mortality and excess mortality. *SSM - Population Health* **19**, 101247 (2022). https://doi.org/10.1016/j.ssmph.2022.101247

33. Newcomb, S. Note on the Frequency of Use of the Different Digits in Natural Numbers. *American Journal of Mathematics,* **4** (1), 39–40 (1881). https://doi.org/10.2307/2369148

34. Nigrini, M. J. *Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection* (Wiley, New York, 2012).

35. Pericchi, L. & Torres, D. Quick Anomaly Detection by the Newcomb—Benford Law, with Applications to Electoral Processes Data from the USA, Puerto Rico and Venezuela. *Statistical Science,* **26** (4), 502–516 (2011).

36. Power, J. *Why China's COVID figures don't add up* (2023). https://www.aljazeera.com/economy/2022/12/9/why-chinas-covid-figures-are-hard-to-take-at-face-value

37. Sambridge, M. & Jackson, A. National COVID numbers—Benford's law looks for errors. *Nature* **581** (7809), 384–384 (2020). https://doi.org/10.1038/d41586-020-01565-5

38. Sambridge, M., Tkalčić, H. & Jackson, A. Benford's law in the natural sciences. *Geophysical Research Letters* **37** (22) (2010). https://doi.org/10.1029/2010GL044830

39. Shao, L. & Ma, B.-Q. Empirical mantissa distributions of pulsars. *Astroparticle Physics* **33** (4), 255–262 (2010). https://doi.org/10.1016/j.astropartphys.2010.02.003

40. Silva, L. & Figueiredo Filho, D. (2020). Using Benford's law to assess the quality of COVID-19 register data in Brazil. *Journal of Public Health* **43** (1), 107–110 (2021) https://doi.org/10.1093/pubmed/fdaa193.

41. Stanway, D. & Lapid, N. Analysis: How accurate are China's COVID-19 death numbers? (2022). https://www.reuters.com/business/healthcare-pharmaceuticals/how-accurate-are-chinas-covid-19-death-numbers-2022-12-22/

42. WHO. (2023, January). *WHO welcomes data on COVID-19 in China, meeting with Minister*. https://www.who.int/news/item/14-01-2023-who-welcomes-data-on-covid-19-in-china--meeting-with-minister

43. Wu, J., Wang, J., Nicholas, S., Maitland, E., & Fan, Q. Application of Big Data Technology for COVID-19 Prevention and Control in China: Lessons and Recommendations. *J Med Internet Res* **22** (10): e21980 (2020).