Benford’s Law Analysis to Evaluate the Quality Data of Covid-19 Epidemiological Surveillance in Indonesia

by Evi Sinaga

Benford's Law Analysis to Evaluate the Quality Data of Covid-19 Epidemiological Surveillance in Indonesia

Evi Susanti Sinaga[1], Novia Indriani Sudharma[2]
[1,2]Department of Public Health, Faculty of Medicine, Universitas Trisakti, Indonesia

Article Info

Article history:

Received month dd, yyyy
Revised month dd, yyyy
Accepted month dd, yyyy

Keywords:

Benford's Law
Covid-19
Epidemiological
Surveillance

ABSTRACT

Countries worldwide, including Indonesia, grappled with the unprecedented challenges brought about by the Corona Virus Disease (COVID-19) pandemic. Surveillance data vividly illustrates the profound effect of the COVID-19 pandemic on Indonesia. Both daily cases and deaths were raised, revealing the rapid transmission of the virus within communities. A quantitative study was accomplished with secondary data to evaluate the quality of COVID-19 epidemiological surveillance data in Indonesia during the period between March 2020 to January 2021. The data was sourced from the World Health Organization (WHO) website using data reports on COVID-19 confirmed cases and deaths. A rapid tool called the first digit law or the fulfillment of Benford's Law was used to suggest good data quality for epidemiological surveillance. Data analysis used the chi-squared test and the log-likelihood ratio test. Also, it displayed the difference in mean absolute deviation (MAD) to identify the proximity of the data and Benford's Law distribution. The results showed that both confirmed and death case distributions were statistically non-conformity with Benford's Law distribution. The next phase of this study would be to conduct a complete evaluation suitably, especially in post-pandemic COVID-19.

***Corresponding Author*:**

Evi Susanti Sinaga
Department of Public Health, Faculty of Medicine, Universitas Trisakti
Jl. Kyai Tapa No.260, Grogol, Jakarta Barat, 11440, Indonesia.
Email: sinaga.evisusanti@trisakti.ac.id

1. INTRODUCTION

The transmission of Covid-19 (Corona Virus Disease-19) impacted many countries, including Indonesia. Indonesia has experienced a rise in the number of cases over a short period. As an infectious disease, Covid-19 transmission happens via droplets and contact with the virus. After that, through the open mucosa, the virus can enter [1]. An outbreak of pneumonia of obscure origin started December 2019 and was reported in China, at Wuhan, Hubei Province. Furthermore, the global spread of SARS-CoV-2 was reported by the World Health Organization (WHO) on March 12, 2020, and because of Covid-19, thousands of deaths occurred [2].

Based on the national trend, Covid-19 in Indonesia still shows occurrence in 2023, with the total number of cases in May of 6,798,097 confirmed cases and 161,630 deaths. In Indonesia, DKI Jakarta is the region with the highest cases of the spread of Covid-19 followed by Central Java and then East Java [3]. In March 2020, according to the study and investigation of responses to Covid-19 in Indonesia, surveillance and epidemiological analyses were emphasized to comprehend the scope of the Covid-19 condition in the Indonesia [4].

Surveillance systems have a fundamental function in controlling the pandemic of Covid-19 [5]. In practice, the surveillance system experiences challenges such as different data collection platforms, poor interoperability, data duplication problems, data integration, data completeness, and data analysis which ultimately have an impact on countermeasures responses [6]. A complete evaluation of a surveillance system should be assessed including the features of the surveillance characteristics of flexibility, sensitivity, simplicity, acceptability, representativeness, stability, timeliness, and positive predictive value (PPV) [5], [7]. Epidemiological surveillance systems are commonly evaluated after an epidemic. It has occurred due to a need for rapid evaluation techniques to specify whether cases meet expectations during an occurrence [8], [9]. Finally, the method of Benford's Law, known as the Law of First Digits, Newcomb-Benford's Law, or the Law of Anomaly Numbers, was utilized to evaluate the quality data of the surveillance system, especially in public health surveillance [10].

In mathematics, there are methods to determine the authenticity of data. One of these methods is based on the frequency of occurrence of the first digit, which is called Benford's Law. Frank Benford, a physicist, 1938 found that number one appears in the first digit of random data more often than number two, number two more often than number three, and so on. The frequency of occurrence of a number will decrease as the number in the first digit increases [11]–[13].

In more detail, Benford's Law can estimate the frequency of occurrence of a number in a series of numerical data. If the numerical data is generated without intention, then the number's occurrence frequency will be by the expected frequency in Benford's Law. Conversely, suppose there is an element of human intentionality to create and include a number combination in a data set. In that case, the Benford's Law analysis results will show that specific numbers appear more or less than expected. Benford's Law is widely used in various fields because it detects anomalous data in a data set. If further explained, such data anomalies can help detect fraud [14], [15].

For the field of epidemic control, a reliable epidemiological surveillance system is essential. Providing high-quality data so that decisions can be made using evidence is one of its responsibilities [16]. In this study, Benford's Law analysis was used to evaluate the first significant digit distribution of daily confirmed cases and Covid-19-related deaths in Indonesia.

2. RESEARCH METHOD

A quantitative method was used in this research. A quantitative study is a systematic scientific analysis of the components and phenomena and the cause of their associations [17]. An method based on Benford's Law was presented in this work [18]. In order to assess the data quality for an epidemiological monitoring system, Benford's distribution of confirmed cases and deaths of Covid-19 was examined. This research used Covid-19 epidemiological surveillance data. The collected data were obtained from WHO website *https://covid19.who.int/* by using data daily reports on confirmed cases and death caused by Covid-19 and covering all subjects in Indonesia. The cases included in this study were from March 2020 to January 2021. The period used adjusts before mass vaccination interventions are carried out in Indonesia [19]. Covid-19 epidemiological surveillance data in Indonesia will be evaluated for data quality by analyzing the confirmed cases and deaths and conformance to Benford's distribution. Data analysis used visual analysis through pictures and statistical analysis using the Chi-squared test and the log-likelihood ratio, on top of that, displayed diff Mean Absolute Deviation (MAD). Data analysis used statistical software STATA 17.

A mathematical phenomenon called Benford's Law, commonly referred to as the First-Digit Law shows how leading digits are distributed across numerous real-world datasets. This analysis uses physics-related assumptions regarding the distribution of naturally occurring data. Using integral calculus, Benford calculates how often a digit or a combination of digits will likely appear. In the law regarding the probability of occurrence of numbers, Benford formulates the expectation of the appearance of a series of numbers as the first digit by assuming that the occurrence will follow a logarithmic distribution, with a pattern that can be summed up in the following formula [11]:

$$P(d) = \log\,[1+(1/d)]$$

$P(d)$ = probability of a digit will be the leading number
d = 1,2…….,9, series of numbers

According to Benford's Law, data whose initial number is one (30.103%) appears more often than data that begins with another number, following the order from 2 to 9 (respectively 17.609%; 12.494%; 9.691%; 7.918%; 6.695%; 5.799%; 5.115%; and 4.576%). In simpler terms, this means that smaller digits (1, 2, 3) have a higher probability of being the leading digit, while larger digits (7, 8, 9) have a lower probability

[20]. Benford’s law was utilized in this study to analyze the conformity of the epidemiological data by using the formula and replacing *(d)* with Covid-19 daily cases and death.

3. RESULTS AND DISCUSSIONS

The quality of epidemiological surveillance system data can be assessed quickly using Benford’s Law analysis. As several studies demonstrated throughout the AI pandemic (Avian Influenza/H1N1) and the epidemic Dengue incidence in Paraguay assessed surveillance systems performance in each country using Benford’s Law [21], [22]. Surveillance systems work well if the data and distribution follow the first-digit distribution. Benfod's Law said the digit that appears most frequently in surveillance reports is the first digit. It supposes the first digit (30.103%), then pursued by the different numbers from digits two to nine. (respectively, 17.609%, 12.494%, 9.691%, 7.918%, 6.695%, 5.799 %, 5.115%, 4.576%) [20].

Based on the undertaking of the International Health Regulation, it is known that there are challenges faced by developing countries in terms of disease control and response. One of them is that governments are aware of gaps in disease surveillance capacity which will impact the ability to monitor and respond to disease. The pandemic of Covid-19 has demonstrated a failure to respond to the emergence of approvingly infectious and deadly microbes [23], [24]. It is necessary to strengthen the existing health system and build an adequate surveillance system to prevent the next pandemic [25], [26].

In this study, the first step is to evaluate the quality of surveillance data using Benford’s Law by evaluating confirmed cases and Covid-19 deaths, whether following the algorithm (distribution) or not. The following is an overview of confirmed cases and death cases started from March 2020 to January 2021. Observations ended on January 2021, before the intervention, the mass vaccination of Covid-19 in Indonesia. Figure 1 illustrates that confirmed cases and deaths in Indonesia from March 2020 to January 2021 are based on daily cases. Covid-19 showed an increasing trend. However, there was a decrease at a certain point, but it continued to increase both confirmation cases and deaths.

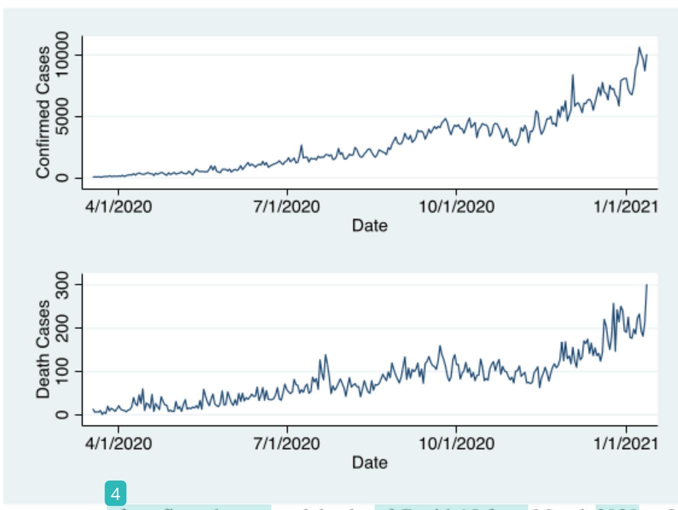

Figure 1. Daily occurrence of confirmed cases and deaths of Covid-19 from March 2020 to January 2021

Evaluation of surveillance data of Covid-19 uses the first digit test (Table 1 and Table 2), which shows whether the first digit of each number in the observed distribution of numbers has data anomalies or otherwise conforms to the expected distribution. It can be proved by calculating the goodness of fit test and using the P-value as significance. The P-value was obtained through the chi-square test and the likelihood ratio. Besides that, it also calculates the MAD value to see how far the data match Benford’s Law distribution. For this test using the statistical hypothesis H_0 which is the two distributions (observation cases and Benford’s Law distribution) are the same. It means that the distribution follows that predicted by Benford’s Law. In contrast, H_1 is the two different distributions. Hence, the bigger the P-value, the higher the confidence to accept H0 that the observed distribution is following the expectations of Benford’s Law theory. In addition, visualization in graphs is also presented, showing the observed data distribution compared with Benford’s Law distribution. Figures 2 and 3 give a summary of the distribution of the daily number of confirmed cases and deaths caused by Covid-19 and compare it to the distribution determined by Benford's algorithm. Tables 1 and Table 2 display the results of the first-digit test analysis for compliance with Benford's Law.

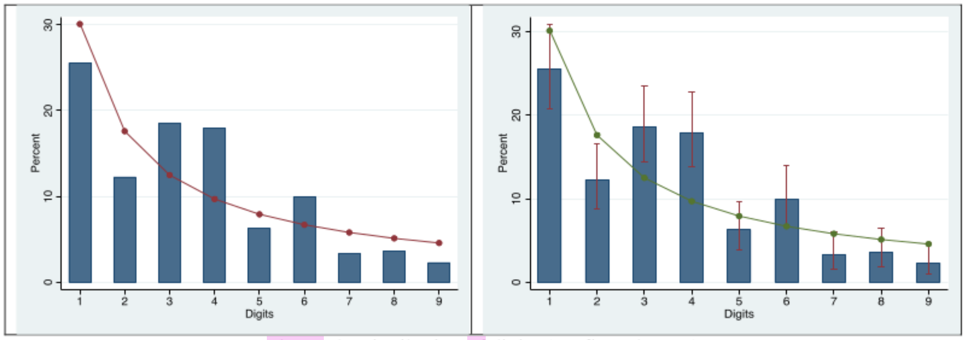

Figure 2. Distribution of digits (confirmed cases)

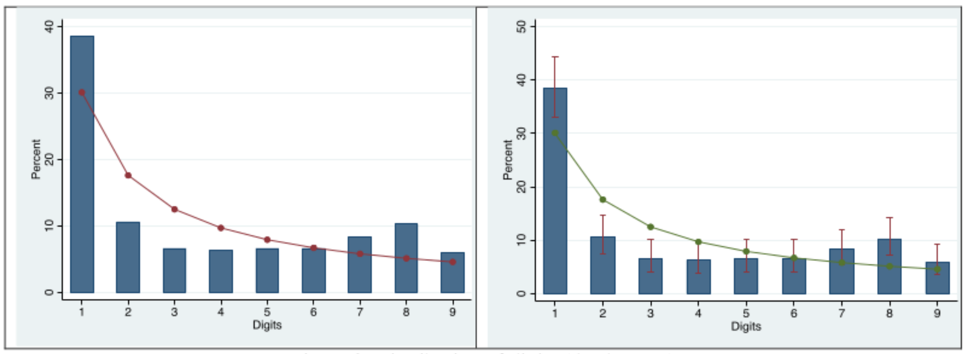

Figure 3. Distribution of digits (death cases)

Figure 2 illustrates the Covid-19 confirmed case distribution to the predicted Benford's Law distribution using the first digit distribution. The graph of the daily reports of confirmed cases of Covid-19 shows that the first digits of the numbers two, three, four, and six do not follow the predicted distribution according to Benford's Law algorithm and confidence intervals. Furthermore, Figure 3 illustrates the death case of Covid-19 distribution to the Benford's Law expected. The first digits one, two, three, seven, and eight do not follow the expected distribution seen from Benford's frequency distribution and confidence intervals.

Table 1. Table of first-digit distribution of confirmed cases and tests of significance

Digit	*Count*	*Observed*	*Expected (Benford)*	*Diff. MAD*	*P-value*	*Pearson's X^2 (p-value)*	*Log likelihood ratio (p-value)*
1	77	25.581	30.103	-4.522	0.0900	0.0000	0.0000
2	37	12.292	17.609	-5.317	0.0152		
3	56	18.605	12.494	6.111	0.0022		
4	54	17.940	9.691	8.249	0.0000		
5	19	6.312	7.918	-1.606	0.3376		
6	30	9.967	6.695	3.272	0.0281		
7	10	3.322	5.799	-2.477	0.0644		
8	11	3.654	5.115	-1.461	0.2952		
9	7	2.326	4.576	-2.250	0.0708		
Total	301	100.000	100.000	3.918			

Table 2. Table of first-digit distribution of death cases and tests of significance

Digit	*Count*	*Observed*	*Expected (Benford)*	*Diff. (MAD)*	*P-value*	*Pearson's X^2 (p-value)*	*Log likelihood ratio (p-value)*
1	116	38.538	30.103	8.435	0.0020	0.0000	0.0000
2	32	10.631	17.609	-6.978	0.0011		

3	20	6.645	12.494	-5.849	0.0012
4	19	6.312	9.691	-3.379	0.0506
5	20	6.645	7.918	-1.274	0.5208
6	20	6.645	6.695	-0.050	1.0000
7	25	8.306	5.799	2.506	0.0823
8	31	10.299	5.115	5.184	0.0003
9	18	5.980	4.576	1.404	0.2668
Total	301	100.000	100.000	3.895	

In the analysis of fulfillment with Benford's Law, both the confirmed case and death case variables show the result of rejecting the null hypothesis. The likelihood ratio and the chi-square statistical test results in Tables 1 and 2 have a P-value of 0.000 (P-value<0.05), indicating that the alternative hypothesis is accepted. It denotes a statistical difference between the observed and expected distributions (Benford distribution). In addition, the analysis using Mean Absolute Deviation (MAD) was performed to determine whether the observed data and the Benford distribution were similar. The outcome indicated that the two variables, confirmed cases and deaths brought on by Covid-19, both display a MAD value of >0.015, suggesting that they do not follow the Benford distribution (3.918 and 3.895, respectively).

The results of this analysis were likewise found to be the same as an investigation performed in India. The data is from daily data on Covid-19 patients and dyings in India and Kerala. Based on the analysis, the distribution of Covid-19 cases complies with the first digit of Benford's Law. Still, for death reports, India's national data and the state of Kerala do not match Benford's Law distribution. It is shown through The MAD value of Covid-19 deaths for India's national data (0.0171) and the MAD value for the state of Kerala (0.0415), which means it does not match the Benford distribution [27]. These results are also in line with research conducted by Kilani and Georgiou, showing that many Covid-19 report data were found to be inconsistent with the distribution of Benford's Law, especially in developing countries [28].

Previously were some examples of studies conducted in developing countries. In developed countries like America, the same thing has also been found, and it is possible that there is unreported data related to Covid-19. The findings of this investigation, which are based on a study done in the United States (US), showed that Covid-19 deaths were not reported in several US states based on an analysis using Benford's Law to judge the accuracy of epidemiological data. The testing procedure determines the degree of compliance with Benford's Law. Still, the studies with the most substantial and most apparent evidence using the MAD criteria show evidence of deaths from Covid-19 that were not reported in the US [29]. However, it is slightly different from studies conducted in China. Based on Benford's Law, the results show that there is no proof of the unconformity of Covid-19 data in the China [30].

When Benford's Law is not fulfilled, there are several potential scenarios. Suppose the observed distribution of death data does not meet Benford's Law, but the number is greater than the average death rate. In that case, the response to the pandemic was likely inadequate. When Benford's Law is not fulfilled, mortality rates are descending, potential reasons are insufficient scope or coverage, or the country is in the early phase of an epidemic/pandemic. Insufficient surveillance systems are indicated by a shortage of diagnostic tests, restricted scope, or existing in the early stages of an epidemic where the Benford distribution has yet to be observed. However, after an epidemic or pandemic ends, complementary studies can be carried out related to evaluating a surveillance system that is more stringent than other attributes of a surveillance system [22] [31].

Although the results of Benford's Law fulfillment analysis concluded that the data on confirmed cases and deaths caused by Covid-19 was not the same as the expected distribution, this did not mean that these results were able to complete as a whole regarding the evaluation of the surveillance system. As a preliminary study that can quickly analyze the implementation of a surveillance system, Benford's Law has been accepted and recognized in several studies [31]–[33]. A surveillance system's overall evaluation should consider its attributes, such as its adaptability, sensitivity, acceptance, simplicity, representativeness, stability, timeliness, and positive predictive value (PPV) but are commonly evaluated after an epidemic. Benford's Law was an effective method for evaluating the accuracy or reliability of the data produced by various nations or even distinct regions within one nation [16]. The authors tested the reliability of COVID-19 death-case reporting in nations with authoritarian regimes using Benford's Law, and their results are presented. They concluded that when it comes to the reported COVID-19 death-case numbers, democratically run nations adhere to Benford's Law more closely than authoritarian ones [34]. Since Benford's Law can be utilized to test data set reliability, a study in another field study used Benford's to assess the quality of widely employed survey data groups. The result concludes that nearly all the data groups indicate substantial dissimilarities, significantly suggesting reliability issues in the survey data [35].

However, what needs to be noted is that this evaluation only includes some elements of the surveillance system involved in managing the Covid-19 pandemic. Further, a complete evaluation of a surveillance system is required, which consists of the attributes of the surveillance system [36]. A quick evaluation using Benford's Law provides feedback to relevant stakeholders in Indonesia. Ongoing evaluation allows responsible governments to create appropriate determinations to enhance epidemiological surveillance systems.

4. CONCLUSION

In fulfillment with Benford's Law, the confirmed case and death caused by Covid-19 show the result of non-conformity to Benford distribution. The conclusion has been drawn based on the chi-square statistical test and the likelihood ratio. The next stage of this study would be to conduct a complete evaluation of a surveillance system that includes the features of the system characteristics that are undertaken suitably, especially in the post-pandemic of Covid-19.

ACKNOWLEDGEMENTS

We thank the World Health Organization that has provided access to the data. The authors also thank the Research Institute of the Universitas Trisakti as a research funder with number 165/A.1/LPPM-P/USAKTI/IX/2022.

REFERENCES

[1] M. Lotfi, M. R. Hamblin, and N. Rezaei, "COVID-19: Transmission, prevention, and potential therapeutic opportunities," *Clin. Chim. Acta*, vol. 508, pp. 254–266, 2020, doi: https://doi.org/10.1016/j.cca.2020.05.044.

[2] M. Ciotti, M. Ciccozzi, A. Terrinoni, W.-C. Jiang, C.-B. Wang, and S. Bernardini, "The COVID-19 pandemic," *Crit. Rev. Clin. Lab. Sci.*, vol. 57, no. 6, pp. 365–388, Aug. 2020, doi: 10.1080/10408363.2020.1783198.

[3] The Task Force for COVID-19, "Map of Spread of Covid 19," 2021. https://covid19.go.id/peta-sebaran (accessed Mar. 21, 2022).

[4] R. Djalante *et al.*, "Review and analysis of current responses to COVID-19 in Indonesia: Period of January to March 2020," *Prog. Disaster Sci.*, vol. 6, p. 100091, 2020, doi: https://doi.org/10.1016/j.pdisas.2020.100091.

[5] N. K. Ibrahim, "Epidemiologic surveillance for controlling Covid-19 pandemic: types, challenges and implications," *J. Infect. Public Health*, vol. 13, no. 11, pp. 1630–1638, 2020, doi: https://doi.org/10.1016/j.jiph.2020.07.019.

[6] Z. Allam and D. S. Jones, "On the coronavirus (COVID-19) outbreak and the smart city network: universal data sharing standards coupled with artificial intelligence (AI) to benefit urban health monitoring and management," in *Healthcare*, 2020, vol. 8, no. 1, p. 46, [Online]. Available: https://doi.org/10.3390/healthcare8010046.

[7] M. Peyre *et al.*, "The RISKSUR EVA tool (Survtool): A tool for the integrated evaluation of animal health surveillance systems," *Prev. Vet. Med.*, vol. 173, p. 104777, 2019, doi: https://doi.org/10.1016/j.prevetmed.2019.104777.

[8] P. Michelozzi *et al.*, "Mortality impacts of the coronavirus disease (COVID-19) outbreak by sex and age: Rapid mortality surveillance system, Italy, 1 February to 18 April 2020," *Eurosurveillance*, vol. 25, no. 19, pp. 1–5, 2020, doi: 10.2807/1560-7917.ES.2020.25.19.2000620.

[9] N. W. Sari *et al.*, "Teori dan Aplikasi Epidemiologi Kesehatan (Theory and Application of Health Epidemiology)," 1st Editio., Yogyakarta: Zahir Publishing, 2021, pp. 109–132.

[10] V. S. Balashov, Y. Yan, and X. Zhu, "Using the Newcomb–Benford law to study the association between a country's COVID-19 reporting accuracy and its development," *Sci. Rep.*, vol. 11, no. 1, p. 22914, 2021, doi: 10.1038/s41598-021-02367-z.

[11] A. Berger and T. P. Hill, "The mathematics of Benford's law: a primer," *Stat. Methods Appt.*, vol. 30, no. 3, pp. 779–795, 2021, doi: 10.1007/s10260-020-00532-8.

[12] F. Li, S. Han, H. Zhang, J. Ding, J. Zhang, and J. Wu, "Application of Benford's law in Data Analysis," in *Journal of Physics: Conference Series*, 2019, vol. 1168, no. 3, p. 32133, [Online]. Available: doi: 10.1088/1742-6596/1168/3/032133%0A.

[13] R. Cerqueti and M. Maggi, "Data validity and statistical conformity with Benford's Law," *Chaos, Solitons & Fractals*, vol. 144, p. 110740, 2021, doi: https://doi.org/10.1016/j.chaos.2021.110740.

[14] L. Barabesi, A. Cerioli, and D. Perrotta, "Forum on Benford's law and statistical methods for the detection of frauds," *Stat. Methods Appt.*, vol. 30, no. 3, pp. 767–778, 2021, doi: 10.1007/s10260-021-00588-0.

[15] C. da S. Azevedo, R. F. Gonçalves, V. L. Gava, and M. de M. Spinola, "A Benford's Law based

methodology for fraud detection in social welfare programs: Bolsa Familia analysis," *Phys. A Stat. Mech. its Appl.*, vol. 567, p. 125626, 2021, doi: https://doi.org/10.1016/j.physa.2020.125626.
[16] F. G. Morillas-Jurado, M. Caballer-Tarazona, and V. Caballer-Tarazona, "Applying Benford's Law to Monitor Death Registration Data: A Management Tool for the COVID-19 Pandemic," *Mathematics*, vol. 10, no. 1. 2022, doi: 10.3390/math10010046.
[17] H. K. Mohajan, "Quantitative research: A successful investigation in natural and social sciences," *J. Econ. Dev. Environ. People*, vol. 9, no. 4, pp. 50–79, 2020, [Online]. Available: https://mpra.ub.uni-muenchen.de/105149/.
[18] R. Cerqueti and C. Lupi, "Some new tests of conformity with Benford's law," *Stats*, vol. 4, no. 3, pp. 745–761, 2021, [Online]. Available: https://doi.org/10.3390/stats4030044.
[19] E. S. Sinaga, R. Pou, G. H. Tarigan, B. E. Yuwono, and H. Hartini, "Provision of Covid-19 Booster Vaccination to Accelerate the Herd Immunity in DKI Jakarta Region," *JUARA J. Wahana Abdimas Sejah.*, pp. 228–237, 2022, [Online]. Available: doi: https://doi.org/10.25105/juara.v3i2.13677.
[20] G. Fang and Q. Chen, "Several common probability distributions obey Benford's law," *Phys. A Stat. Mech. its Appl.*, vol. 540, p. 123129, 2020, doi: https://doi.org/10.1016/j.physa.2019.123129.
[21] M. Gómez-Camponovo, J. Moreno, Á. J. Idrovo, M. Páez, and M. Achkar, "Monitoring the Paraguayan epidemiological dengue surveillance system (2009-2011) using Benford's law," *Biomedica*, vol. 36, no. 4, pp. 583–592, 2016, doi: 10.7705/biomedica.v36i4.2731.
[22] A. J. Idrovo, J. A. Fernández-Niño, I. Bojórquez-Chapela, and J. Moreno-Montoya, "Performance of public health surveillance systems during the influenza A(H1N1) pandemic in the Americas: Testing a new method based on Benford's Law," *Epidemiol. Infect.*, vol. 139, no. 12, pp. 1827–1834, 2011, doi: 10.1017/S095026881100015X.
[23] N. Kandel, S. Chungong, A. Omaar, and J. Xing, "Health security capacities in the context of COVID-19 outbreak: an analysis of International Health Regulations annual report data from 182 countries," *Lancet*, vol. 395, no. 10229, pp. 1047–1053, 2020, [Online]. Available: doi.org/10.1016/S0140-6736(20)30553-5.
[24] D. Carroll *et al.*, "Preventing the next pandemic: the power of a global viral surveillance network," *BMJ*, vol. 372, p. n485, Mar. 2021, doi: 10.1136/bmj.n485.
[25] V. Haldane *et al.*, "Health systems resilience in managing the COVID-19 pandemic: lessons from 28 countries," *Nat. Med.*, vol. 27, no. 6, pp. 964–980, 2021, doi: 10.1038/s41591-021-01381-y.
[26] H. Legido-Quigley *et al.*, "Are high-performing health systems resilient against the COVID-19 epidemic?," *Lancet*, vol. 395, no. 10227, pp. 848–850, 2020, [Online]. Available: https://doi.org/10.1016/S0140-6736(20)30551-1.
[27] K. Natashekara, "COVID-19 cases in India and Kerala: a Benford's law analysis," *J. Public Health (Oxf).*, p. fdab199, Jun. 2021, doi: 10.1093/pubmed/fdab199.
[28] A. Kilani and G. P. Georgiou, "Countries with potential data misreport based on Benford's law," *J. Public Health (Bangkok).*, vol. 43, no. 2, pp. e295–e296, Jun. 2021, doi: 10.1093/pubmed/fdab001.
[29] M. Campolieti, "COVID-19 deaths in the USA: Benford's law and under-reporting," *J. Public Health (Bangkok).*, vol. 44, no. 2, pp. e268–e271, Jun. 2022, doi: 10.1093/pubmed/fdab161.
[30] C. Koch and K. Okamura, "Benford's Law and COVID-19 reporting," *Econ. Lett.*, vol. 196, p. 109573, 2020, doi: https://doi.org/10.1016/j.econlet.2020.109573.
[31] E. F. Manrique-Hernández, J. A. Fernández-Niño, and A. J. Idrovo, "Global performance of epidemiologic surveillance of Zika virus: rapid assessment of an ongoing epidemic," *Public Health*, vol. 143, pp. 14–16, 2017, doi: 10.1016/j.puhe.2016.10.023.
[32] A. J. Idrovo and E. F. Manrique-Hernández, "Data Quality of Chinese Surveillance of COVID-19: Objective Analysis Based on WHO's Situation Reports," *Asia-Pacific J. Public Heal.*, vol. 32, no. 4, pp. 165–167, 2020, doi: 10.1177/1010539520927265.
[33] L. Silva and D. Figueiredo Filho, "Using Benford's law to assess the quality of COVID-19 register data in Brazil," *J. Public Health (Oxf).*, vol. 43, no. 1, pp. 107–110, 2021, doi: 10.1093/pubmed/fdaa193.
[34] L. Burlac and N. Giannakis, "Benford's Law: Analysis of the trustworthiness of COVID-19 reporting in the context of different political regimes," Mälardalen University, Västerås, Swedan, 2021.
[35] G. M. Eckhartt and G. D. Ruxton, "Investigating and preventing scientific misconduct using Benford's Law," *Res. Integr. Peer Rev.*, vol. 8, no. 1, p. 1, 2023, doi: 10.1186/s41073-022-00126-w.
[36] A. K. S. Ng'etich, K. Voyi, and C. M. Mutero, "Evaluation of health surveillance system attributes: the case of neglected tropical diseases in Kenya," *BMC Public Health*, vol. 21, no. 1, p. 396, 2021, doi: 10.1186/s12889-021-10443-2.

BIOGRAPHIES OF AUTHORS

	**Evi Susanti Sinaga** is an epidemiologist working as a lecturer in the Department of Public Health at Universitas Trisakti. She has an MPH, especially in the Field Epidemiology Training Program. She can be contacted at email: sinaga.evisusanti@trisakti.ac.id.
	**Novia Indriani Sudharma** is a doctoral candidate in Epidemiology, Public Health. She has been working as a lecturer at Universitas Trisakti since 2008. She is interested in non-communicable diseases. She can be contacted at email: noviaindriani@trisakti.ac.id.

Benford's Law Analysis to Evaluate the Quality Data of Covid-19 Epidemiological Surveillance in Indonesia

ORIGINALITY REPORT

18%	13%	16%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	eprints.uad.ac.id Internet Source	2%
2	www.researchgate.net Internet Source	1%
3	academic.oup.com Internet Source	1%
4	www.nornesk.no Internet Source	1%
5	"Benford's Law", Wiley, 2012 Publication	1%
6	Alvaro J. Idrovo, Edgar F. Manrique-Hernández, Julián A. Fernández Niño. "Report From Bolsonaro's Brazil: The Consequences of Ignoring Science", International Journal of Health Services, 2020 Publication	1%
7	repository.uinsu.ac.id Internet Source	1%

8	doaj.org Internet Source	1 %
9	arxiv.org Internet Source	1 %
10	www.cambridge.org Internet Source	1 %
11	A. J. IDROVO, J. A. FERNÁNDEZ-NIÑO, I. BOJÓRQUEZ-CHAPELA, J. MORENO-MONTOYA. "Performance of public health surveillance systems during the influenza A(H1N1) pandemic in the Americas: testing a new method based on Benford's Law", Epidemiology and Infection, 2011 Publication	<1 %
12	"Principles for Evaluation of One Health Surveillance: The EVA Book", Springer Science and Business Media LLC, 2022 Publication	<1 %
13	Sijia Tao, Keivan Zandi, Leda Bassit, Yee Tsuey Ong et al. "Comparison of anti-SARS-CoV-2 activity and intracellular metabolism of remdesivir and its parent nucleoside", Current Research in Pharmacology and Drug Discovery, 2021 Publication	<1 %
14	journal.ugm.ac.id Internet Source	<1 %

15	www.scilit.net Internet Source	<1 %
16	Arno Berger, Theodore P. Hill. "An Introduction to Benford's Law", Walter de Gruyter GmbH, 2015 Publication	<1 %
17	download.ei-ie.org Internet Source	<1 %
18	docplayer.net Internet Source	<1 %
19	www.mdpi.com Internet Source	<1 %
20	benfordonline.net Internet Source	<1 %
21	Chang Qing, Zixuan Yan, Yang Xia, Han Su, Shuai Yan, Yitao Gao, Xi Zhang, Yixiao Zhang, Caigang Liu, Yuhong Zhao. "The Global Prevalence of Anxiety Symptoms Among Healthcare Workers Before and During COVID-19: A Systematic Review and Meta-analysis", Research Square Platform LLC, 2023 Publication	<1 %
22	Jared Eutsler, M. Kathleen Harris, L. Tyler Williams, Omar E. Cornejo. "Accounting for Partisanship and Politicization: Employing	<1 %

Benford's Law to Examine Misreporting of COVID-19 Infection Cases and Deaths in the United States", Accounting, Organizations and Society, 2023

Publication

23	LIJING SHAO, BO-QIANG MA. "FIRST DIGIT DISTRIBUTION OF HADRON FULL WIDTH", Modern Physics Letters A, 2011 Publication	<1%
24	Perigrinus Sebong, Dwi Handono Sulistio, Yodi Mahendradhata. "Sustainability Capacity of HIV/AIDS Programmes in Yogyakarta, Indonesia", International Journal of Public Health Science (IJPHS), 2017 Publication	<1%
25	go.gale.com Internet Source	<1%
26	ijphs.iaescore.com Internet Source	<1%
27	Chikodi Modesta Umeozuru, Aishat Bukola Usman, Abdulhakeem Abayomi Olorukooba, Idris Nasir Abdullahi et al. "Performance of COVID-19 case-based surveillance system in FCT, Nigeria, March 2020 –January 2021", PLOS ONE, 2022 Publication	<1%

28 Gabriel Cirac Souza, Robson Moreno, Tales Pimenta. "Benford’s Law and Artificial Intelligence Applied to COVID-19", 2021 International Conference on Microelectronics (ICM), 2021
Publication
<1%

29 Gianluca Sottili, Danilo M. Palladino, Biagio Giaccio, Paolo Messina. "Benford’s Law in Time Series Analysis of Seismic Clusters", Mathematical Geosciences, 2012
Publication
<1%

30 M. Jayasree, C. S. Pavana Jyothi, P. Ramya. "Benford’s Law and Stock Market—The Implications for Investors: The Evidence from India Nifty Fifty", Jindal Journal of Business Research, 2018
Publication
<1%

31 Muhammad Mazhar Iqbal, Irfan Abid, Saddam Hussain, Naeem Shahzad, Muhammad Sohail Waqas, Muhammad Jawed Iqbal. "The effects of regional climatic condition on the spread of COVID-19 at global scale", Science of The Total Environment, 2020
Publication
<1%

32 Roy Cerqueti, Claudio Lupi. "Severe testing of Benford’s law", TEST, 2023
Publication
<1%

33	hdl.handle.net Internet Source	<1 %
34	Adrian Patrick Kennedy, Sheung Chi Phillip Yam. "On the authenticity of COVID-19 case figures", PLOS ONE, 2020 Publication	<1 %
35	Alladoumbaye Ngueilbaye, Joshua Zhexue Huang, Mehak Khan, Hongzhi Wang. "Data quality model for assessing public COVID-19 big datasets", The Journal of Supercomputing, 2023 Publication	<1 %
36	E.F. Manrique-Hernández, J.A. Fernández-Niño, A.J. Idrovo. "Global performance of epidemiologic surveillance of Zika virus: rapid assessment of an ongoing epidemic", Public Health, 2017 Publication	<1 %
37	Mirela Păunescu, Elena-Mirela Nichita, Paula Lazăr, Alexandra Frățilă. "Chapter 4 Applying Benford's Law to Detect Fraud in the Insurance Industry—A Case Study from the Romanian Market", Springer Science and Business Media LLC, 2023 Publication	<1 %
38	Peter Olofsson. "Faking Probabilities: Computer Simulation", Wiley, 2006 Publication	<1 %

39	Sulistyawati Sulistyawati, Syamsu Hidayat, Siwi Pramatama Mars Wijayanti, Tri Wahyuni Sukesi et al. "Knowledge, attitude, and practice towards COVID-19 among university students in Indonesia: A cross-sectional study", International Journal of Public Health Science (IJPHS), 2021 Publication	<1 %
40	Eko Sujadi, Muhammad Fadhli, Muhd. Odha Meditamar, Dairabi Kamil, Ahmad Jamin, Hengki Yandri, Syaiful Indra. "Generalized anxiety disorder associated with individual work performance of Indonesian medical personnel during COVID-19 outbreak", International Journal of Public Health Science (IJPHS), 2021 Publication	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		