

# Origin, Alternative Expressions of Newcomb-Benford Law and Deviations of Digit Frequencies

Gene Whyman

Physics Department, Faculty of Natural Sciences, Ariel University, Ariel, Israel  
Email: gevaiman@gmail.com

**How to cite this paper:** Whyman, G. (2021) Origin, Alternative Expressions of Newcomb-Benford Law and Deviations of Digit Frequencies. *Applied Mathematics*, 12, 576-586.  
<https://doi.org/10.4236/am.2021.127041>

**Received:** May 14, 2021

**Accepted:** July 9, 2021

**Published:** July 12, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

The Newcomb-Benford law, which describes the uneven distribution of the frequencies of digits in data sets, is by its nature probabilistic. Therefore, the main goal of this work was to derive formulas for the permissible deviations of the above frequencies (confidence intervals). For this, a previously developed method was used, which represents an alternative to the traditional approach. The alternative formula expressing the Newcomb-Benford law is re-derived. As shown in general form, it is numerically equivalent to the original Benford formula. The obtained formulas for confidence intervals for Benford's law are shown to be useful for checking arrays of numerical data. Consequences for numeral systems with different bases are analyzed. The alternative expression for the frequencies of digits at the second decimal place is deduced together with the corresponding deviation intervals. In general, in this approach, all the presented results are a consequence of the positionality property of digital systems such as decimal, binary, etc.

## Keywords

Benford's Law, Confidence Intervals in Benford's Law, Alternative Expression of Benford's Law, Benford's Law for Different Numeral Systems, Frequencies of Digits at the Second Decimal Place

## 1. Introduction

The surprising fact of the uneven distribution of decimal digits over decimal places was noticed by Newcomb [1] in 1881 and then rediscovered in 1938 by Benford [2], who gave the corresponding mathematical expression

$$F(n) = \log\left(1 + \frac{1}{n}\right), \quad (1)$$

where  $F(n)$  is the frequency of numbers having the first digit  $n$  (the base of the logarithm is 10). Evidently, the sum of all 9 frequencies equals 1.

Since then, the law has been repeatedly tested and applied in a wide variety of areas [3]-[10] and continues to attract the attention of researchers [11]-[22]. Meanwhile, efforts have been made to justify or derive the above equation [23] [24] [25] [26]. In Ref [26] it was connected with the scaling invariance of physical laws. It was shown that the Benford law is valid for numbers distributed exponentially [27]. Also, there exists the geometrical explanation of the Benford law [28]. Such a variety of explanations for the same law is somewhat unusual in physics and mathematics where, as a rule, there is a single main reason explaining its origin.

The simplest explanation of the Newcomb-Benford law has been given by the author of the present communication in cooperation with Ed. Bormashenko and E. Shulzinger [29]. We have shown that Benford's law follows as a consequence of the "positionality" of numeral systems like the decimal one.

People unacquainted with the literature on the subject refuse to believe that in arrays of unfalsified data almost a third of decimal numbers begin with the digit 1. At the same time, no one is surprised by the fact that in the binary system all the numbers begin with the digit 1.

The main assumptions of the present work are as follows. 1) Any natural numerical array, which is being analyzed, is bounded from above either essentially or by the number of presented digits (the position of the decimal point is irrelevant). 2) Within this array, the probability of encountering any number is the same for all numbers, but not for digits in each position.

First, I will give a brief overview of the cited work [29] presenting its results in a more convenient form. An alternative to Equation (1) expression will be deduced in a new way; its relation to Equation (1) will be clarified. The importance of inequalities for frequencies of digits will be emphasized and illustrated by the distribution of the population in Israeli cities and by the results by state of the 2020 presidential elections in the United States. Finally, the same method of extremal frequencies will be applied to the determination of digits frequencies at the second decimal place.

## 2. Frequencies of Digits as a Consequence of the Structure of the Positional Numeral System

Benford's law is often used to check the reliability of various arrays of numerical data. In any case, these arrays are bounded above and below, and the role of these restrictions can be played simply by the given number of digits. For example, the number of votes cast for a particular candidate in some place is limited by the number of those who have the right to vote; the distribution of the population by city is limited by the population of the country, etc. As will be clear from what follows, only the upper limit is significant, while the lower one is unimportant.

Now let us consider the frequency  $F(n)$  of numbers with a certain digit  $n = 1, 2, \dots, 9$  at first place in the set of natural numbers  $\{1, 2, \dots, m\} \equiv \{m\}$ . How this frequency changes with increasing  $m$ ? In Ref. [29] it has been shown that the frequency  $F(n)$  to pass a sequence of alternating local minimums,  $F_{\min,k}(n)$ , and local maximums  $F_{\max,k}(n)$ . For example, for  $n = 1$ , the minimums of  $F(1)$  are achieved for values of  $m$  equal to 9, 99, 999, ... due to the maximum of the denominator in the frequency definition while the numerator remains constant. Starting from these values, the frequency begins to rise since the growth of the numerator in percentage is greater than the growth of the denominator. Maximums are attained at values of  $m$  equal to 19, 199, 1999, and so on. Another example is  $n = 7$  where local minimums and maximums of the frequency are at  $m$  values of 69, 699, 6999, ..., and 79, 799, 7999, ..., respectively. In general form, these statements are expressed by the formulas for  $k = 1, 2, 3, \dots$

$$m_{\min,k} = n \cdot 10^k - 1, \tag{2}$$

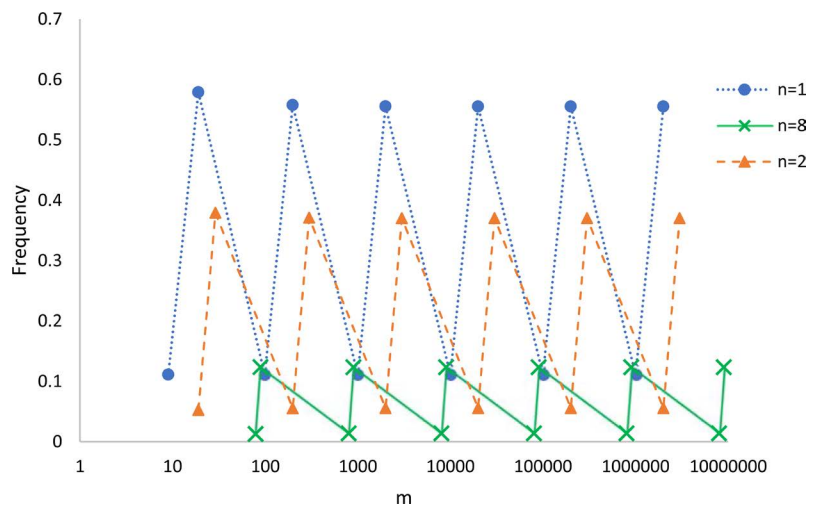
$$m_{\max,k} = (n + 1) \cdot 10^k - 1. \tag{3}$$

Amount of numbers starting with the digit  $n$  up to  $m_{\min,k}$  equals to  $10^{k-1} + 10^{k-2} + \dots + 1 = (10^k - 1)/9$ ; analogously, amount of numbers starting with the digit  $n$  up to  $m_{\max,k}$  equals  $10^k + 10^{k-1} + \dots + 1 = (10^{k+1} - 1)/9$ . Once more taking into account Equations ((2), (3)), one has for the frequencies at local minimums and maximums:

$$F_{\min,k}(n) = \frac{10^k - 1}{9(n \cdot 10^k - 1)}, \tag{4}$$

$$F_{\max,k}(n) = \frac{10^{k+1} - 1}{9[(n + 1) \cdot 10^k - 1]}. \tag{5}$$

The dependences of these quantities on  $m$  and the digit value  $n$  are shown in **Figure 1**. With the increase of  $k$ , both quantities in Equations ((4), (5)) converge



**Figure 1.** Maximal and minimal frequencies of decimal digits at the first place in a set of natural numbers restricted by  $m$  (logarithmic scale).

very rapidly to

$$F_{\min}(n) = \lim_{k \rightarrow \infty} F_{\min,k}(n) = \frac{1}{9n}, \tag{6}$$

$$F_{\max}(n) = \lim_{k \rightarrow \infty} F_{\max,k}(n) = \frac{10}{9(n+1)}. \tag{7}$$

It is seen that the frequencies vary inversely with a digit value that is the qualitative formulation of the Newcomb-Benford law. The limiting values of frequencies (6) and (7) also presented in **Table 1** are very useful in analyzing real numerical arrays because in reality the exact upper bond,  $m$ , is unknown. The lowest bond is also unknown, but this is unessential since contribution of the first terms of the set  $\{1, 2, 3, \dots, m\}$  to both numerator and denominator of the frequency definition is negligible compared to contribution of the last terms.

From the minimal and maximal asymptotical frequencies, some mean values may be constructed like arithmetic, geometric, logarithmic or harmonic one. The geometric mean turned out to be closest to the original Benford distribution in Equation (1). One has from Equations ((6), (7))

$$F(n) = \frac{0.43077}{\sqrt{n(n+1)}} \tag{8}$$

where a numerical multiplier is the normalizing factor  $A = \left[ \sum_{i=1}^9 1/\sqrt{i(i+1)} \right]^{-1}$ . The quantity  $F(n)$  is interpreted as the probability for the digit  $n$  to occupy the first place in a number of the decimal numeral system. As is seen from **Table 1**, the differences between the results of Benford's law (1) and Equation (8) are negligible, at least from a practical point of view.

### 3. Equivalence of Two Formulations of Benford's Law

It is possible to prove the equivalence of Equation (8) to the Benford law (1). After the change of variable  $n$

$$n = (e^x - 1)^{-1}, \quad \ln(10/9) \leq x \leq \ln 2 \tag{9}$$

Equation (8) turns into

$$F(n) = A \frac{1}{\sqrt{n(n+1)}} = 2A \sinh\left(\frac{x}{2}\right) = 2A \left[ \frac{x}{2} + \frac{1}{6} \left(\frac{x}{2}\right)^3 + \dots \right]. \tag{10}$$

Passing to approximate formulas we have

**Table 1.** Frequencies of digits in numerical arrays.

Digit, $n$	1	2	3	4	5	6	7	8	9
Benford's Equation (1)	0.3010	0.1761	0.1249	0.09691	0.07918	0.06695	0.05799	0.05115	0.04576
$F(n)$ , Equation (8)	0.3046	0.1759	0.1244	0.09632	0.07865	0.06647	0.05756	0.05077	0.04541
$F_{\min}(n)$ , Equation (6)	0.111	0.0556	0.037	0.028	0.022	0.018	0.016	0.014	0.012
$F_{\max}(n)$ , Equation (7)	0.556	0.370	0.278	0.222	0.185	0.159	0.139	0.123	0.111

$$F(n) = A \left[ x + O\left(\frac{x^3}{24}\right) \right] \approx Ax. \tag{11}$$

The inverse to (9) transformation is

$$x = \ln \frac{n+1}{n} \tag{12}$$

that gives after substitution of  $A = 0.43077$  from Equation (8) and transition to logarithms with a base of 10

$$F(n) \approx 0.992 \log \frac{n+1}{n}. \tag{13}$$

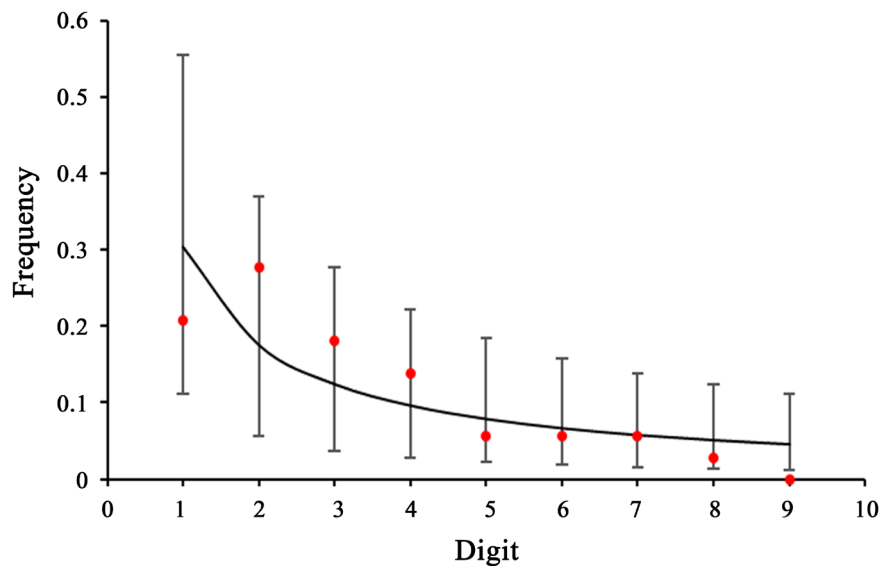
The error  $Ax^3/24$  associated with neglecting higher-order terms reaches a maximum of +0.006 for  $n = 1$ , but for  $n = 2$  it is already equal to +0.001. The validity of the proof of the equivalence of Benford’s law and Equation (8) is confirmed by numerical results in the table. In conclusion, the question arises whether to consider the formula (8) as an excellent approximation to Benford’s law (1), or, conversely, to consider Benford’s formula as an excellent approximation to the law (8) deduced from the properties of the decimal numeral system?

#### 4. Population of Israeli Cities and by State Results of the 2020 Presidential Elections in the United States

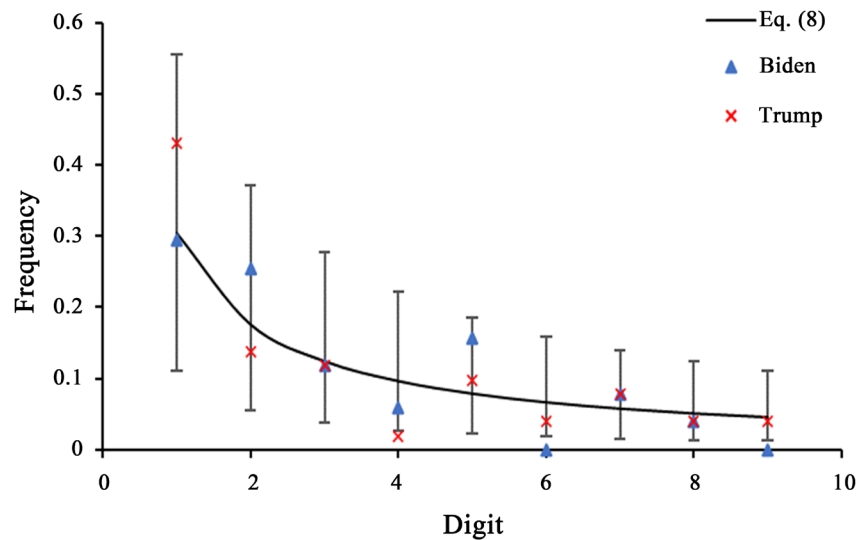
From Equations ((6) and (7)) the inequalities for  $F(n)$  follow

$$\frac{1}{9n} \leq F(n) \leq \frac{10}{9(n+1)}, \tag{14}$$

which is more useful than Equations ((1), (8)) in applications because the exact upper bound of an array considered is unknown. **Figure 2** shows an analysis



**Figure 2.** The frequencies of digits at the first place of the numbers from data on population of 72 Israeli cities [30]. Solid line corresponds to Equation (8). Deviation intervals according to inequality (14).



**Figure 3.** The frequencies of digits at the first place of the numbers from the results by state of the 2020 US presidential election [31].

using (8) and (14) of the population distribution over 72 cities of Israel in accordance with the results of the 2008 census [30]. Lacking the right software and programming skills, I choose data with a small number of items and performed the calculation manually. For the same reason, as another example, a sampling of 51 items was considered, which represents the results by state of the 2020 US presidential election [31] (Figure 3).

In both Figures, there are only a few cases of going beyond the deviation intervals (14), and these goings are small. It should not be forgotten that all consideration is probabilistic. In addition, for the above reason, small samplings were chosen. The Benford law is often used to detect violations and fraud in datasets. From this point of view, the adequacy of the census and the vote count in the presidential 2020 elections successfully pass this test.

## 5. Other Positional Numeral Systems

Generalization to other than decimal numeral systems is straightforward. Instead of Equations ((8) and (14)) one has Equations ((15) and (16)) [29]

$$F_N(n) = \frac{A_N}{\sqrt{n(n+1)}} \quad (15)$$

where  $N$  is the base of a numeral system,  $1 \leq n \leq N-1$ , and  $A_N = \left[ \sum_{i=1}^{N-1} 1/\sqrt{i(i+1)} \right]^{-1}$  is the normalizing factor and

$$\frac{1}{(N-1)n} \leq F_N(n) \leq \frac{N}{(N-1)(n+1)}. \quad (16)$$

In particular, in the binary system ( $N=2$ ), all the last equations turn into 1 for  $n=1$  (all the numbers in the binary system begin with 1).

Normalizing factors for the most popular numeral systems are as follows:  $A_2 = 1$ ,  $A_8 = 0.467469$ ,  $A_{10} = 0.430773$ ,  $A_{16} = 0.353036$ . For systems with a

low base, the probability of finding the digit 1 at the first place of number is high; for example,  $F_4(1) = 0.503626$ . This makes them undesirable for digital encryption since the chances are high that the encoded number starts with one. In this regard, coding using high base numeral systems is preferred; for example, in hexadecimal numeral system this probability is two times lower,  $F_{16}(1) = 0.249634$ .

## 6. Frequencies of Digits at the Second Decimal Place

The alternation of local minimums and maxima of digit frequencies when expanding a limited set of natural numbers  $\{m\}$  can also be used for the second decimal position. Let us show this by the example of the digits 0 and 2. For  $n = 0$ , maxima are attained for the following values of  $m$ : 10, 20, ..., 90, 109, 209, ..., 909, 1099, 2099, ... 9099, 10,999, 20,999, ..., 90,999, .... Wherein, amount of numbers contained in  $\{m\}$  with 0 at the second place changes, respectively, as follows: 1, 2, ..., 9, 19, 29, ..., 99, 199, 299, ..., 999, 1999, 2999, ... 9999, .... In general form, this is expressed with a help of two indices:  $i \cdot 10^k + 10^{k-1} - 1$  and  $(i+1) \cdot 10^{k-1} - 1$  where  $i$  runs from 1 to 9. Thus, the corresponding frequency is

$$G_{i,k}^{\max}(0) = \frac{(i+1) \cdot 10^{k-1} - 1}{i \cdot 10^k + 10^{k-1} - 1}. \quad (17)$$

Analogously, for the minimal frequencies the values of  $m$  and amount of the required numbers in  $\{m\}$  are, respectively, 9, 19, 29, ..., 99, 199, 299, ..., 999, 1999, 2999, ..., 9999, 19,999, 29,999, ..., 99,999, ... and 0, 1, 2, ..., 9, 19, 29, ..., 99, 199, 299, ... 999, 1999, 2999, ..., 9999. The minimal frequencies are

$$G_{i,k}^{\min}(0) = \frac{i \cdot 10^{k-1} - 1}{i \cdot 10^k - 1}. \quad (18)$$

For  $n = 2$ , the maximal frequencies are attained for  $m$  values: 12, 22, ..., 92, 129, 229, ... 929, 1299, 2299, ..., 9299, 12,999, 22,999, ..., 92,999, .... Amount of numbers with 2 at the second place in  $\{m\}$  is as follows: 1, 2, ..., 9, 19, 29, ..., 99, 199, 299, ..., 999, 1999, 2999, ..., 9999, .... The formula for the frequency:

$$G_{i,k}^{\max}(2) = \frac{(i+1) \cdot 10^{k-1} - 1}{i \cdot 10^k + 3 \cdot 10^{k-1} - 1}. \quad (19)$$

The minimal frequencies are for  $m$ : 11, 21, ..., 91, 119, 219, ..., 919, 1199, 2199, ..., 9199, 11,999, 21,999, ..., 91,999, ... with the same amount of numbers in numerator as in Equation (18):

$$G_{i,k}^{\min}(2) = \frac{i \cdot 10^{k-1} - 1}{i \cdot 10^k + 2 \cdot 10^{k-1} - 1}. \quad (20)$$

For  $0 \leq n \leq 9$ :

$$G_{i,k}^{\max}(n) = \frac{(i+1) \cdot 10^{k-1} - 1}{i \cdot 10^k + (n+1) \cdot 10^{k-1} - 1} \quad (21)$$

$$G_{i,k}^{\min}(n) = \frac{i \cdot 10^{k-1} - 1}{i \cdot 10^k + n \cdot 10^{k-1} - 1}. \quad (22)$$

Both the minimum of Equation (22) (minimum minimorum) and the maxi-

imum of Equation (21) are attained at  $i = 1$ . Further, the limits for  $k$  going to infinity:

$$G^{\max}(n) = \frac{2}{11+n}, \tag{23}$$

$$G^{\min}(n) = \frac{1}{10+n}. \tag{24}$$

One of possible estimations of the probability to find the digit  $n$  at the second place of a number may be the normalized geometrical mean of the maximal and minimal frequencies (23) and (24):

$$G(n) = \frac{1.44237}{\sqrt{(10+n)(10+n+1)}} \tag{25}$$

(compare to Equation (15)).Corresponding numerical data are presented in **Table 2**.

It is seen from **Table 2** that the confidence intervals strongly overlap that may prevent the second digit statistics from analyzing arrays of numbers. Also, the distribution of probabilities is too smooth.

For the numeral systems with the arbitrary base  $N$ , one has from Equation (25):

$$G_N(n) = \frac{B_N}{\sqrt{(N+n)(N+n+1)}} \tag{26}$$

where  $B_N = \left[ \sum_0^{N-1} 1/\sqrt{(N+i)(N+i+1)} \right]^{-1}$  (compare to Equation (25)). For low bases, the probabilities may differ more sharply. Thus, for the binary system ( $N=2$ ):  $G_2(0) = 2 - \sqrt{2} \approx 0.59$  while  $G_2(1) = \sqrt{2} - 1 \approx 0.41$ .

### 7. Conclusions

With the expansion of a bounded set of natural numbers, the density of numbers starting with a certain digit experiences quasiperiodic oscillation (see **Figure 1**). The maxima and minima of this oscillation quickly stabilize and determine possible deviations from Benford’s law. Formulas for these deviations are useful when analyzing numeric arrays for fraud.

The geometric mean of the above minimum and maximum decreases with the increase in values of the initial digits of numbers, giving an alternative quantitative

**Table 2.** Probabilities and deviation intervals for digits at the second decimal place.

Digit, $n$	0	1	2	3	4	5	6	7	8	9
$G(n)$ , Equation (25)	0.1375	0.1255	0.1155	0.1069	0.0995	0.0931	0.0875	0.0825	0.0780	0.0740
$G^{\min}(n)$ , Equation (24)	0.1000	0.0909	0.0833	0.0769	0.0714	0.0667	0.0625	0.0588	0.0556	0.0526
$G^{\max}(n)$ , Equation (23)	0.1818	0.1667	0.1538	0.1429	0.1333	0.1250	0.1176	0.1111	0.1053	0.1000



expression (8) for the Newcomb-Benford law. This expression approximately coincides with Benford's formula (1) up to the second order in some small parameter, which is less than 1.

The results are generalized to the arbitrary base of positional numeral systems. The systems with the higher base are preferable in digital encryption because for them digit frequencies are close to each other and maximal possible deviations overlap.

The elaborated method of extremal digital frequencies applies to the second decimal place. In this case, a smooth dependence on the digit value may prevent the method from applying to the check of numerical arrays for fraud.

Perhaps, taking into account information about the boundaries of the numerical arrays under consideration will narrow the confidence intervals and improve the correspondence of the calculated and measured frequencies in the case of truthful data. Work in this direction is expected to be carried out in the near future.

## Acknowledgements

The author is grateful to Professor Edward Bormashenko who brought his attention to this field.

## Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

## References

- [1] Newcomb, S. (1881) Note on the Frequency of Use of Different Digits in Natural Numbers. *American Journal of Mathematics*, **4**, 39-40. <https://doi.org/10.2307/2369148>
- [2] Benford, F. (1938) The Law of Anomalous Numbers. *Proceedings of the American Philosophical Society*, **78**, 551-572.
- [3] Pain, J.-C. (2008) Benford's Law and Complex Atomic Spectra. *Physical Review E*, **77**, Article ID: 012102. <https://doi.org/10.1103/PhysRevE.77.012102>
- [4] Mir, T.A. (2012) The Law of the Leading Digits and the World Religions. *Physica A*, **391**, 792-798. <https://doi.org/10.1016/j.physa.2011.09.001>
- [5] Sambridge, M. (2010) Benford's Law in the Natural Sciences. *Geophysical Research Letters*, **37**, L22301. <https://doi.org/10.1029/2010GL044830>
- [6] Hernandez Caceres, J.L. (2008) First Digit Distribution in Some Biological Data Sets. Possible Explanations for Departures from Benford's Law. *The Electronic Journal of Biomedicine*, **1**, 27-35.
- [7] Friar, J.L, Goldman, T. and Pérez-Mercader, J. (2012) Genome Sizes and the Benford Distribution. *PLoS ONE*, **7**, e36624. <https://doi.org/10.1371/journal.pone.0036624>
- [8] Shao, L. and Ma, B.Q. (2010) Empirical Mantissa Distributions of Pulsars. *Astroparticle Physics*, **33**, 255-262. <https://doi.org/10.1016/j.astropartphys.2010.02.003>
- [9] Mir, T.A., Ausloos, M. and Cerqueti, R. (2014) Benford's Law Predicted Digit Dis-

- tribution of Aggregated Income Taxes: The Surprising Conformity of Italian Cities and Regions. *The European Physical Journal B*, **87**, 261.  
<https://doi.org/10.1140/epjb/e2014-50525-2>
- [10] Kaiser, M. (2019) Benford's Law as an Indicator of Survey Reliability—Can We Trust Our Data? *Journal of Economic Surveys*, **33**, 1602-1618.  
<https://doi.org/10.1111/joes.12338>
- [11] Alipour, A. and Alipour, S. (2019) Application of Benford's Law in Analyzing Geotechnical Data. *Civil Engineering Infrastructures Journal*, **52**, 323-334.
- [12] da Silva Azevedo, C., Gonçalves, R.F., Gava, V.L. and de Mesquita Spinola, M. (2021) A Benford's Law Based Methodology for Fraud Detection in Social Welfare Programs: Bolsa Familia Analysis. *Physica A: Statistical Mechanics and Its Applications*, **567**, Article ID: 125626. <https://doi.org/10.1016/j.physa.2020.125626>
- [13] da Silva, A.J., Floquet, S., Santos, D.O.C. and Lima, R.F. (2020) On the Validation of the Newcomb-Benford Law and the Weibull Distribution in Neuromuscular Transmission. *Physica A: Statistical Mechanics and Its Applications*, **553**, Article ID: 124606.  
<https://doi.org/10.1016/j.physa.2020.124606>
- [14] Tunalioglu, N. and Erdogan, B. (2019) Usability of the Benford's Law for the Results of Least Square Estimation. *Acta Geodaetica et Geophysica*, **54**, 315-331.  
<https://doi.org/10.1007/s40328-019-00259-3>
- [15] Whyman, G., Ohtori, N., Shulzinger, E. and Bormashenko, Ed. (2016) Revisiting the Benford Law: When the Benford-Like Distribution of Leading Digits in Sets of Numerical Data Is Expectable? *Physica A: Statistical Mechanics and Its Applications*, **461**, 595-601. <https://doi.org/10.1016/j.physa.2016.06.054>
- [16] Istrate, C. (2019) Detecting Earnings Management Using Benford's Law: The Case of Romanian Listed Companies. *Journal of Accounting and Management Information Systems*, **18**, 198-223. <https://doi.org/10.24818/jamis.2019.02003>
- [17] Morag, S. and Salmon-Divon, M. (2019) Characterizing Human Cell Types and Tissue Origin Using the Benford Law. *Cells*, **8**, 1004.  
<https://doi.org/10.3390/cells8091004>
- [18] Yan, X., Yang, S.-G., Kim, B.J. and Minnhagen, P. (2017) Benford's Law and First Letter of Word. *Physica A: Statistical Mechanics and Its Applications*, **512**, 305-315.  
<https://doi.org/10.1016/j.physa.2018.08.133>
- [19] Shulzinger, E. and Bormashenko, E. (2017) On the Universal Quantitative Pattern of the Distribution of Initial Characters in General Dictionaries: The Exponential Distribution Is Valid for Various Languages. *Journal of Quantitative Linguistics*, **24**, 273-288. <https://doi.org/10.1080/09296174.2017.1304620>
- [20] Dantuluri, A. and Desai, S. (2018) Do  $\tau$  Lepton Branching Fractions Obey Benford's Law? *Physica A: Statistical Mechanics and Its Applications*, **506**, 919-928.  
<https://doi.org/10.1016/j.physa.2018.05.013>
- [21] Da Silva, S.B. (2020) Limits of Benford's Law in Experimental Field. *International Journal of Applied Mathematics*, **33**, 685-695.  
<https://doi.org/10.12732/ijam.v33i4.12>
- [22] Branets, S. (2019) Detecting Money Laundering with Benford's Law and Machine Learning. Master's Thesis, University of Tartu, Faculty of Social Sciences, School of Economics and Business Administration, Tartu.
- [23] Hill, T.P. (1995) Base-Invariance Implies Benford's Law. *Proceedings of the American Mathematical Society*, **123**, 887-895. <https://doi.org/10.2307/2160815>
- [24] Hill, T.P. (1995) A Statistical Derivation of the Significant-Digit Law. *Statistical*

- Science*, **10**, 354-363. <https://doi.org/10.1214/ss/1177009869>
- [25] Berger, A. and Hill, T.P. (2011) Benford's Law Strikes Back: No Simple Explanation in Sight for Mathematical Gem. *The Mathematical Intelligencer*, **33**, 85-91. <https://doi.org/10.1007/s00283-010-9182-3>
- [26] Pietronero, L., Tosatti, E., Tosatti, V. and Vespignani, A. (2001) Explaining the Uneven Distribution of Numbers in Nature: The Laws of Benford and Zipf. *Physica A*, **29**, 297-304. [https://doi.org/10.1016/S0378-4371\(00\)00633-6](https://doi.org/10.1016/S0378-4371(00)00633-6)
- [27] Engel, H.A. and Leuenberger, Ch. (2003) Benford's Law for Exponential Random Variables. *Statistics & Probability Letters*, **63**, 361-365. [https://doi.org/10.1016/S0167-7152\(03\)00101-9](https://doi.org/10.1016/S0167-7152(03)00101-9)
- [28] Fewster, R.M. (2009) A Simple Explanation of Benford's Law. *The American Statistician*, **63**, 26-32. <https://doi.org/10.1198/tast.2009.0005>
- [29] Whyman, G., Shulzinger, E. and Bormashenko, Ed. (2016) Intuitive Considerations Clarifying the Origin and Applicability of the Benford Law. *Results in Physics*, **6**, 3-6. <https://doi.org/10.1016/j.rinp.2015.11.010>
- [30] Israel Central Bureau of Statistics (2008) Profiles by Locality. [https://en.wikipedia.org/wiki/List\\_of\\_cities\\_in\\_Israel#cite\\_note-26](https://en.wikipedia.org/wiki/List_of_cities_in_Israel#cite_note-26)
- [31] <https://www.politico.com/2020-election/results>