Benford's Law and Music Note Frequencies

Sybil Prince Nelson(✉), Brian Wickman, Jack Null, and Eric Gazin

Washington and Lee University, Lexington, VA 24450, USA
sprincenelson@wlu.edu

Abstract. We considered musical note frequencies of the 88 keys of the piano and found that they are Benford distributed. We extended our focus beyond the 88 keys and found the connection to Benford holds for the lowest note measuring at 16.35 Hertz (Hz) to the highest note of 7902.13 Hz. We next investigated whether the distribution holds within specific types of music. We found that classical music such as a random sample of songs from the Romantic period adhere to the Benford distribution while modern music such as a random sampling of songs from the 2000s do not. We also coined a term called "Naturalness" to assess how well a song adheres to the Benford distribution.

Keywords: Benford distribution · Logarithmic distribution of first digits · Classical music

1 Introduction

Greek mathematician and scientist Pythagoras said: "There is geometry in the humming of strings, there is music in the spacing of the strings". Ever since Pythagoras demonstrated that there were simple numerical ratios that produced all the intervals necessary to create a musical scale, there have been countless connections made between mathematics and music [1]. Counting, rhythm, keys, intervals, patterns, symbols, harmonies, time signatures, overtones, tone, pitch and even the notations of composers are all connected to mathematics [2]. This paper introduces yet another connection: musical note frequencies are Benford distributed.

For this paper, we first considered the formula for frequency of music notes and showed that it is a sequence whose limit is Benford. Next, we considered the 88 keys of the piano and their note frequencies in Hertz. Figure 5 provides the 88 frequencies. We also compared a collection of classical songs from the Romantic period of music and compared them to a random selection of songs from the 2000s. We downloaded the classical music MIDI files from http://www.kunstderfuge.com and the modern songs from bitmidi.com and freemidi.com. Each of these websites provide a resource for music files. We used R to extract the notes used in each song. We found that the 88 keys of the piano are Benford

M. Montiel et al. (Eds.): MCM 2022, LNAI 13267, pp. 383–389, 2022.
https://doi.org/10.1007/978-3-031-07015-0_34

distributed. When we extend our focus and look at notes beyond those found on a piano, we find that those frequencies are even closer to the Benford distribution. When comparing the works from the Romantic period to the songs of the 2000s, we find that Romantic period works are much more Benford. This paper will proceed as follows: first, we will provide a summary of Benford's Law. Then we will present our analysis and results of the distribution of note frequencies. Next, we will discuss what conclusion can be made based on our analysis as well as possible future steps.

1.1 What is Benford's Law

Benford's Law began as the empirical observation from Newcomb in 1881 that distribution for first digits in any set of numbers is not uniform or symmetric [3]. Specifically, the frequency of 1s is higher than the frequency of 2s and so on. This observation was later supported by Benford in 1938 with a data set of over 20,000 numbers from naturally occurring data sets such as the atomic weights of elements [4]. He was able to show that the probability of a digit being $d (d = 1, 2, \cdots, 9)$ was equal to

$$Prob(d) = \log_{10}(1 + \frac{1}{d})$$

This was then named Benford's Law. The first digit probabilities follow the specific distribution shown in Fig. 1.

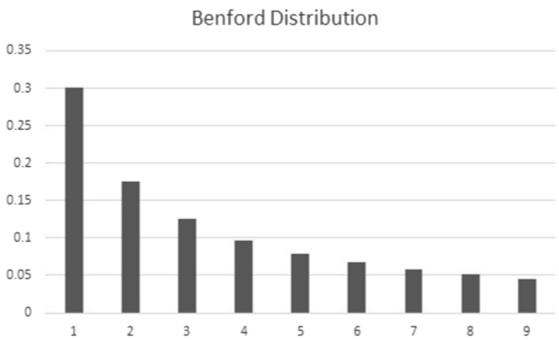

Fig. 1. Distribution of Benford's law

The Benford distribution has been found in many collections of numerical data across various fields of natural and social sciences [5] and even in music [6]. This paper further explores its connection to music. To measure how far a song deviates from Benford, we use a measure for Delta found in Berger and Hill's *An Introduction to Benford's Law.* [7]

$$\varDelta = 100 \cdot (max)_{d=1}^{9} |Prob(D_1 = d) - \log_{10}(1 + \frac{1}{d})|$$

where D_1 denotes the first digit of a number, D_2 denotes the second and so on. For example, $D_1(2568) = 2$ while $D_2(2568) = 5$

The Δ values simply measure how well a data set adheres to the Benford distribution in that if $\Delta = 0$, it is perfectly Benford. Essentially, it compares the frequency of each digit with that of the corresponding digit in the Benford distribution. The largest deviation becomes the Δ value for the set. One limitation of this method is that the digit $d = 1, ..., 9$ that determines the Δ significantly affects whether a χ^2 test of Goodness of Fit is rejected. For example, a Δ of 4.72 creates a p-value of 0.69 if the Δ results from the first digit. A Δ of 4.72 creates a p-value of 0.007 if it results from the seventh digit. This means that we would fail to reject that our observed distribution is Benford if the Δ is from the first digit while we would reject if the same Δ was found in the seventh digit. A future paper address this issue creating a classification system for determining the "Benfordness" of a song.

Not all data sets follow a Benford distribution. One of the criteria for those that do is that it is a naturally occurring data set and one that is not subjective or artificially derived. Because of this, we have determined that the closer a dataset or song adheres to Benford the more "naturalness" it has. We do not quantify this quality as good or bad or leading to a more popular or less popular song. It is merely a measure to assess the music.

2 Numeric Proof

We used the formula $f_n = 440 \cdot (2)^{\frac{n-69}{12}}$ to find the frequency in Hertz of piano keys. This formula corresponds to the note values assigned by MIDI. Note 69 is the A above middle C so it has a frequency 440 Hz, while middle C is assigned note 60 and has a frequency of 261.63. The sequence of frequencies is Benford. Its deviation (Δ) from Benford goes to 0 as N$\rightarrow \infty$ Figs. 2, 3, 4 show that as N$\rightarrow \infty$ or the number of notes increases the distribution of the first digit gets more and more Benford. The deviation from Benford measured with Δ gets closer to 0 (Figs. 6 and 7).

Table 1. Individual probabilities for each digit and Δ for each value of N

N	P(d = 1)	P(d = 2)	P(d = 3)	P(d = 4)	P(d = 5)	P(d = 6)	P(d = 7)	P(d = 8)	P(d = 9)	Δ
10	0	0.275	0.45	0.275	0	0	0	0	0	31.51
100	0.24	0.16	0.15	0.12	0.09	0.08	0.07	0.05	0.04	6.11
1000	0.3	0.178	0.125	0.097	0.079	0.066	0.058	0.051	0.046	0.19
10000	0.301	0.1756	0.1251	0.097	0.0791	0.0671	0.058	0.0511	0.0459	0.05

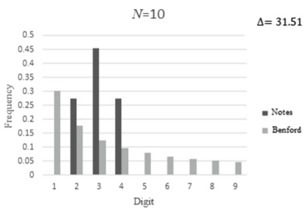

Fig. 2. The first 10 values of the note frequency formula

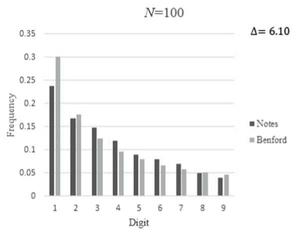

Fig. 3. The first 100 values of the note frequency formula

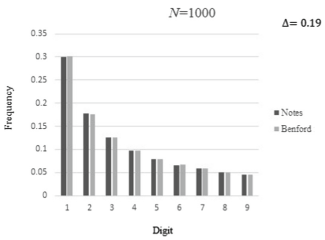

Fig. 4. The first 1000 values of the note frequency formula

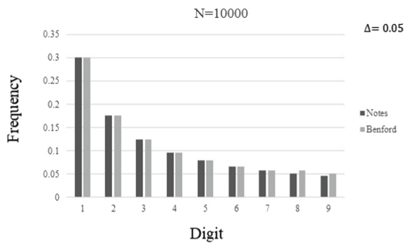

Fig. 5. The first 10000 values of the note frequency formula

As seen in Table 1, for every natural number m, and $d_1 \in 1, 2, ..., 9$ and all $d_j \in 1, 2, ..., 9, j \geq 2$

$$\lim_{N \to \infty} \frac{\text{total}\#1 \leq n \leq N : D_j(f_n) = d_j \text{for} j = 1, 2, ..., 9}{N}$$

$$= \log(1 + (\sum_{j=1}^{m} 10^{m-j} d_j)^{-1})$$

where $f_n = 440 \cdot (2)^{\frac{n-69}{12}}$ and D_j is the first digit of a number corresponding to $j = 1, 2, ..., 9$. Thus, by definition f_n or the sequence of note frequencies is Benford.

3 Piano Note Frequencies

Next, we focused on just the 88 keys of a piano. For piano note frequencies, we also performed a Chi-squared test, Pearson correlation and looked at Euclidean distance. All of these measures provide even more justification for piano note frequencies being Benford distributed. The Chi-Squared p-value for null hypothesis of no difference between the distributions is 0.9755. Thus, we fail to reject the null and conclude that the distribution of the first digit of piano note frequencies

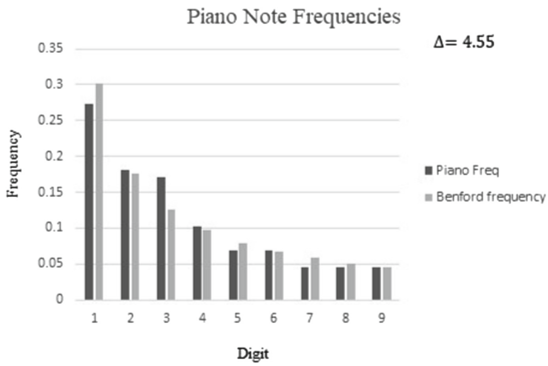

Fig. 6. Piano note frequency distribution

is no different from the Benford distribution. Pearson correlation test indicates that there is a correlation between the distribution of the first digit of piano note frequencies and the Benford distribution. The correlation is 0.97 with a p-value < 0.001. The normalized Euclidean distance is measured as follows:

$$d^* = \frac{\sqrt{\sum_{i=1}^{9}(b_i - e_i)}}{\sqrt{\sum_{i=1}^{8} b_i^2 + (1 - e_9)}}$$

The closer the statistic is to 0 the more similar the distributions are. The Euclidean distance in this instance is 0.055 indicating there is no difference between the distribution of the first digit of piano note frequencies and the Benford distribution.

4 Application

In an attempt to apply what we have noticed about the distribution of piano note frequencies, we compared a sample of music from two different time periods in music history.

Wilcoxon rank sum test shows that there is a significant difference between the median delta value of the Romantic period and that of the 2000 s with a p-value of 0.04. The paper Benford's Law in Music History showed that Western music became more Benford as it moved through history from the Medieval period to the Romantic [8]. One theory as to why this is the case is that Medieval music is known for mostly monophonic chants used in sacred worship [1]. Over time, music became more complex down through the Romantic period which is known for the imagination and virtuosity of its composers. The modern music time period did not maintain this trend as shown in the figure. Our sample of modern songs was less Benford than the songs from the Romantic period. This might be due to the often repetitive nature of modern music, but more research is needed.

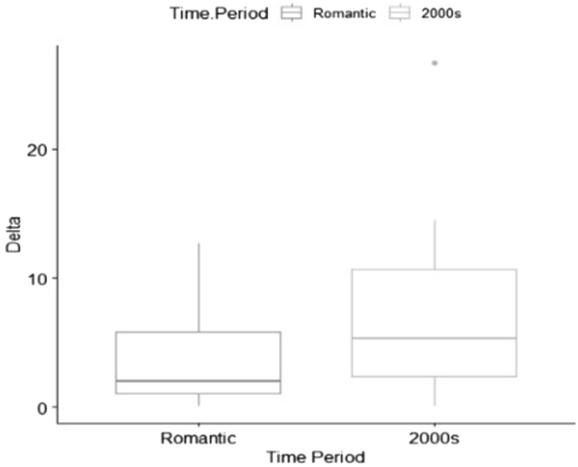

Fig. 7. Comparison of the romantic period to modern music

5 Conclusion

In conclusion, this paper shows that the frequency of musical notes is Benford distributed. Further, classical music such as the works of the Romantic period is more Benford distributed than our collection of modern music. Previous research has shown that music became more Benford as it progressed through history. [8] This trend did not continue through to modern music or music created after World War I. This may be because of the repetition in many modern songs or because of the vast differences found in types of modern music. It may be possible to find trends within specific genres such as Rock or Jazz.

Further research is needed to determine whether how "Benford" a song is will reflect in its popularity. It may even be possible to find some sort of correlation between a song that is Benford distributed and how well or poorly it performs on the charts or in awards programs.

These results cannot be used to quantify what makes a song more popular or more enjoyable. It merely speaks to how well it adheres to the Benford distribution. Since piano note frequencies have a deviation from Benford of $\Delta = 4.55$, we may perhaps conclude that songs near the same Delta value are more "natural". It is possible that it reflects a human component to the music. Thus, we are conducting research in comparing AI created music versus music created by composers. If this holds true, using the Benford distribution may be useful as a first step in identifying digitally manipulated audio recordings just as it is used as a first step in identifying tax fraud or assess the quality of data sets [5].

References

1. Rosenstiel. Leoni: Schirmer History of Music. Schirmer Books (1982)

2. American Mathematical Society: Mathematics and Music. www.ams.org/publicoutreach/math-and-music
3. Newcomb, S.: Note on the frequency of use of the different digits in natural numbers. Am. J. Math. **4**(1), 39–40 (1881). https://doi.org/10.2307/2369148
4. Benford, F.: The law of anomalous numbers. Proc. Am. Philos. Soc. **784**, 551–72 (1938). https://www.jstor.org/stable/984802
5. Li, F., et al.: Application of Benford's law in data analysis. J. Phys. Conf. Ser. **1168**, 032133 (2019). https://doi.org/10.1088/1742-6596/1168/3/032133
6. Khosravani, A., et al.: Emergence of Benford's law in music. J. Math. Sci. Adv. Appl. **54**(1), 11–24 (2018)
7. Berger, A., Hill, T.: An Introduction to Benford's Law. Princeton University Press, New Jersey (2015)
8. Nelson, S.P., et al.: An evaluation of music's adherence to benford's law throughout history. J. Math. Sci. Adv. Appl. **67**(1), 73–84 (2021)
9. Crocetti, E., Randi, G.: "Using Benford's Law as a First Step to Assess the Quality of the Cancer Registry Data" Frontiers in Public Health (2016). https://doi.org/10.3389/fpubh.2016.00225