Research in Astronomy and Astrophysics, 24:055007 (6pp), 2024 May https://doi.org/10.1088/1674-4527/ad386c

First Digit Distributions of Gamma-Ray Bursts

Hou-Yu Lai[1,2] and Jun-Jie Wei[1,2]
[1] Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210023, China; jjwei@pmo.ac.cn
[2] School of Astronomy and Space Sciences, University of Science and Technology of China, Hefei 230026, China
Received 2024 January 11; revised 2024 March 11; accepted 2024 March 26; published 2024 May 2

Abstract

The occurrence of the first significant digits from real world sources is usually not equally distributed, but is consistent with a logarithmic distribution instead, known as Benford's law. In this work, we perform a comprehensive investigation on the first digit distributions of the duration, fluence, and energy flux of gamma-ray bursts (GRBs) for the first time. For a complete GRB sample detected by the Fermi satellite, we find that the first digits of the duration and fluence adhere to Benford's law. However, the energy flux shows a significant departure from this law, which may be due to the fact that a considerable part of the energy flux measurements is restricted by lack of spectral information. Based on the conventional duration classification scheme, we also check if the durations and fluences of long and short GRBs (with duration $T_{90} > 2$ s and $T_{90} \leqslant 2$ s, respectively) obey Benford's law. We find that the fluences of both long and short GRBs still agree with the Benford distribution, but their durations do not follow Benford's law. Our results hint that the long–short GRB classification scheme does not directly represent the intrinsic physical classification scheme.

Key words: (stars:) gamma-ray burst: general – methods: statistical – astronomical databases: miscellaneous

1. Introduction

People might think that the first significant digits (i.e., 1, 2, …, 9) of any randomly chosen data set tend to be uniformly distributed, but it is not true in the natural world. As early as in 1881, Simon Newcomb had observed an unanticipated pattern in the first digits of logarithm tables: the number 1 appears more frequently than 2, 2 than 3, and so on (Newcomb 1881). More than a half century later, Frank Benford rediscovered that there is a logarithmic distribution of first digits in numerous data tables, which is often called the first digit law or Benford's law (Benford 1938). This law states that for a given real database, the probability of numbers with the first digit k is expressed as (Benford 1938)

$$P(k) = \log_{10}\left(1 + \frac{1}{k}\right), \quad k = 1, 2,\ldots, 9. \tag{1}$$

Empirically, Benford's law has been verified in various research fields, including geography (e.g., the lengths of rivers and the areas of lakes; Benford 1938), finance (e.g., stock market indices; Ley 1996; De Ceuster et al. 1998), biology (e.g., pre-vaccination measles incidence data, absolute values from human magnetoencephalography recordings, and gene data lengths of bacteria; Cáceres et al. 2008), seismology (e.g., the recurrence times of seismic events; Sottili et al. 2012), statistical and nuclear physics (e.g., physical constants and distributions; Burke & Kincanon 1991; Shao & Ma 2010b, half-lives of unstable nuclei; Buck et al. 1993; Ni & Ren 2008; Ni et al. 2009, widths of hadrons; Shao & Ma 2009, and the lepton branching fractions; Dantuluri & Desai 2018), etc. In practice, this peculiar law has been effectively used to distinguish and diagnose frauds in taxing and accounting (Nigrini 1996; Nigrini & Mittermaier 1997; Geyer & Williamson 2004), and to minimize storage space and speed up calculation in computer science (Barlow & Bareiss 1985; Schatte 1988; Berger & Hill 2007). Theoretically, this law has been well explained by using a central-limit-like theorem for first digits (Hill 1995) and a simple Markov process (Burgos & Santos 2021). In mathematics, Benford's law is scale invariant (Berger et al. 2008), which indicates that it is independent of any particular choice of units (Pinkham 1961).

In astronomy, Benford's law also has extensive application in all sorts of astrophysical data sets, such as light curves of variable stars and other X-ray sources (Moret et al. 2006), pulsar properties (Shao & Ma 2010a), distances of galaxies and stars (Alexopoulos & Leontsinis 2014), exoplanetary and asteroid data (Shukla et al. 2017; Melita & Miraglia 2021), Gaia Data Release 2 (DR2) parallaxes (de Jong et al. 2020), and so on. Nevertheless, there are also some types of data, e.g., pulsar and fast radio burst dispersion measures, that do not obey Benford's law (Mamidipaka & Desai 2023).

In this work, we investigate the first digit distributions of the duration, fluence, and energy flux of gamma-ray bursts (GRBs), and check if these first digits conform to Benford's law for the first time. GRBs are flashes of high-energy radiation originating from energetic explosions in the Universe. According to their duration time T_{90} (the time interval observed

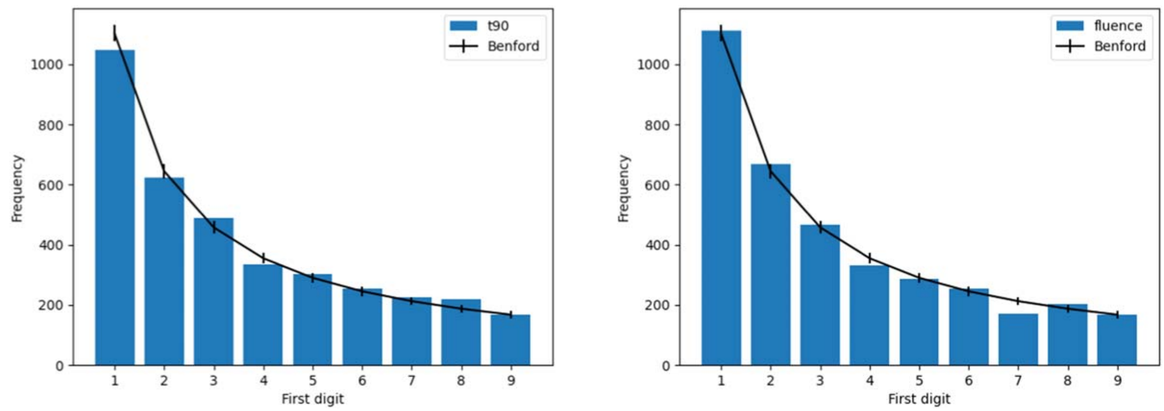

Figure 1. Distributions of the first digit of the duration T_{90} and fluence of all 3664 GRBs from the Fermi-GBM burst catalog. The theoretical predictions from Benford's law (solid lines) along with associated binomial errors are also shown for comparisons.

to contain 90% of the prompt emission), GRBs can be classified into long GRBs ($T_{90} > 2$ s) and short GRBs ($T_{90} \leqslant 2$ s) (Kouveliotou et al. 1993). Generally, long GRBs are suggested to be powered by the core collapses of massive stars (Paczyński 1998; Woosley & Bloom 2006), and short GRBs by the mergers of binary compact objects (Eichler et al. 1989; Narayan et al. 1992).

2. Observational Data and Statistical Results

2.1. Dataset

We download the durations T_{90} (in units of s), fluences F (in units of erg cm^{-2}), and energy fluxes P_γ (in units of erg cm^{-2} s^{-1}) in the 10–1000 keV energy range from the online catalog of GRBs observed by Fermi's Gamma-ray Burst Monitor (Fermi-GBM) (Gruber et al. 2014; von Kienlin et al. 2014; Narayana Bhat et al. 2016; von Kienlin et al. 2020).[3] The Fermi-GBM burst catalog comprises a list of 3665 cosmic GRBs that occurred between 2008 July 12 and 2023 December 6. We remove one GRB for which no relative data were available. We carry out the first digit analysis for all the remaining 3664 GRBs with T_{90} and fluence measurements. The energy flux of each burst in the observer frame is calculated as

$$P_\gamma = p_{64} \times \frac{\int_{E_{\rm min}}^{E_{\rm max}} E \times N(E) {\rm d}E}{\int_{E_{\rm min}}^{E_{\rm max}} N(E) {\rm d}E}, \tag{2}$$

where p_{64} is the peak flux on the 64 ms timescale (in units of photon cm^{-2} s^{-1}), the spectral model $N(E)$ is the Band function (Band et al. 1993), and $E_{\rm min}$ and $E_{\rm max}$ are 10 keV and 1000 keV, respectively. It is obvious that the spectral parameters are required to calculate the energy flux, but not every burst has the requisite spectral information. There are only 2298 GRBs with energy flux measurements in the catalog.

3 https://heasarc.gsfc.nasa.gov/W3Browse/fermi/fermigbrst.html

Table 1
Summary of our Benford Analyses on the First Digit Distributions of the Duration, Fluence, and Energy Flux of GRBs Observed with Fermi-GBM

Dataset	Physical Quantity	Number	χ^2/dof	p-value
All GRBs	Duration	3664	13.0/8	0.11
	Fluence	3664	12.5/8	0.13
Long GRBs	Duration	3061	19.9/8	0.01
	Fluence	3061	11.5/8	0.18
Short GRBs	Duration	603	74.2/8	7×10^{-13}
	Fluence	603	7.1/8	0.53
All GRBs	Energy flux	2298	67.4/8	2×10^{-11}

2.2. Results

The first digit distributions of the duration and fluence for the complete GRB sample are presented in Figure 1 and Table 1. As described above, there are totally $N_{\rm tot} = 3664$ available GRBs. The expected number according to Benford's law, $N_{\rm Ben} = N_{\rm tot}P(k)$, along with the root mean square error estimated by the binomial distribution, $\Delta N = \sqrt{N_{\rm tot} P(k)(1 - P(k))}$, are also shown in the figure. From Figure 1, we can see that the observed distributions are well consistent with the theoretical predictions from Benford's law. In order to quantify the goodness of fit of Benford's law, we adopt the Pearson χ^2

$$\chi^2 = \sum_{k=1}^{9} \frac{[N_{\rm obs}(k) - N_{\rm Ben}(k)]^2}{N_{\rm Ben}(k)}, \tag{3}$$

where $N_{\rm obs}$ and $N_{\rm Ben}$ are the observed number and the expected Benford number for a single digit k, respectively. For the first digit distribution of the duration, we obtain a Pearson χ^2 value of 13.0 for 8 degrees of freedom (dof). For the fluence distribution, we obtain a χ^2 value of 12.5. These two χ^2 values correspond to p-values of 0.11 and 0.13, respectively, which strongly support the

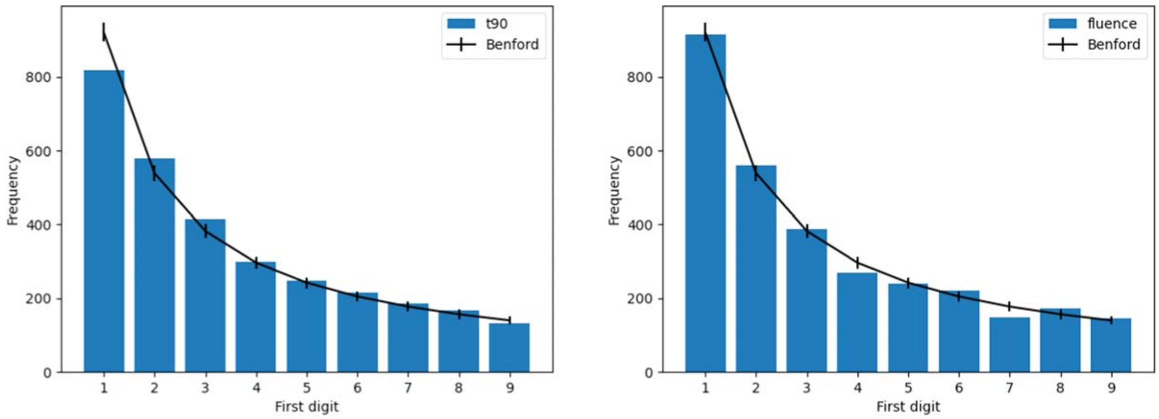

Figure 2. Same as Figure 1, but now for 3061 long GRBs (with $T_{90} > 2$ s) from the Fermi-GBM burst catalog.

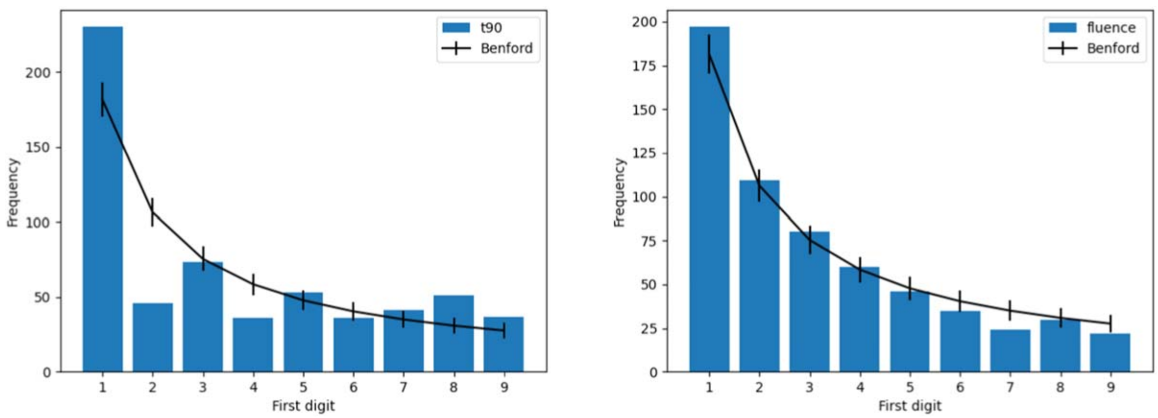

Figure 3. Same as Figure 1, but now for 603 short GRBs (with $T_{90} \leqslant 2$ s) from the Fermi-GBM burst catalog.

null hypothesis that the durations and fluences of the complete GRB sample follow Benford's law. It is worth emphasizing that the higher the p-value, the more likely the null hypothesis. In our present study, we exclude a null hypothesis if $p < 0.05$ (equivalent to under the 95% confidence level).

GRBs can be divided into two classes, long GRBs and short GRBs, with a division line at $T_{90} = 2$ s (Kouveliotou et al. 1993). Using the conventional division between the long and short GRB groups ($T_{90} > 2$ s and $T_{90} \leqslant 2$ s, respectively), we find that there were 3061 long GRBs and 603 short GRBs in the Fermi-GBM burst catalog. In this work, we also check if the durations and fluences of long and short GRBs obey Benford's law. The first digit distributions of the duration and fluence of long and short GRBs are depicted in Figures 2 and 3, respectively. As listed in Table 1, the χ^2 tests and the corresponding p-values for the fluences are extremely supportive to the null hypothesis that the fluences of both long and short GRBs conform to Benford's law. However, the durations of both long and short GRBs obviously deviate from Benford's law. We find that the Pearson χ^2 for the durations of 3061 long GRBs and 603 short GRBs are 19.9 and 74.2, which correspond to p-values of 0.01 and 7×10^{-13}, respectively. On the basis of the p-values, we can safely reject the null hypothesis.

Similarly, in Figure 4 and Table 1, we show the first digit distribution of the energy flux of 2298 available GRBs. We can see that the relative rank of the probability of occurrence of leading digits roughly agrees with Benford's law, but the Benford distribution is not scrupulously obeyed. The rather large χ^2 value and the extremely low p value suggest that the energy flux of GRBs does not adhere to Benford's law.

3. Summary

In this work, we have performed a systematic investigation on the first digit distributions of the duration, fluence, and energy flux of GRBs. For the complete GRB sample detected by Fermi-GBM, our results show that the first digits of the duration and fluence are not uniformly distributed, but that small digits are more common than big ones according to a logarithmic distribution as expected by Benford's law.

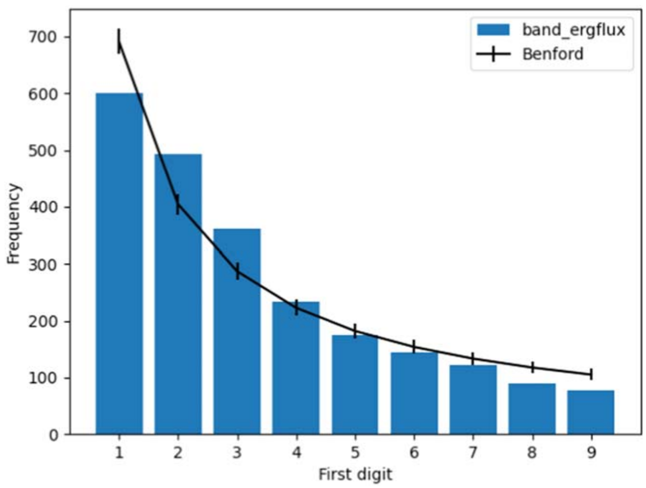

Figure 4. Distribution of the first digit of the energy flux of 2298 available GRBs from the Fermi-GBM burst catalog. The theoretical prediction from Benford's law (solid line) along with associated binomial errors is also shown for comparisons.

However, not all quantities obey Benford's law. Artificial and restricted quantities often deviate from the law, such as the energy fluxes of GRBs in our study. The main reason for the deviation may be that the energy flux measurements are restricted by spectral fits for each GRB, and only 63% of bursts have the required spectral information. Thus, there is not enough dynamic range to make the energy fluxes be fully compliant with Benford's law.

The Fermi-GBM burst catalog also revealed a bimodal T_{90} distribution, and two classes of GRBs, i.e., long versus short GRBs with a separation at about 2 s, were confirmed (Kouveliotou et al. 1993; von Kienlin et al. 2020). Using the conventional long–short GRB classification scheme, we found that there were 3061 long GRBs (with $T_{90} > 2$ s) and 603 short GRBs (with $T_{90} \leqslant 2$ s). Here, the first digits of the duration and fluence of long and short GRBs were also examined for adherence to Benford's distribution. We found that the fluences of both long and short GRBs still follow Benford's law, but their durations are no longer consistent with this law. Our results indicate that the data on fluence seem to be very natural and believable, but that T_{90} is not always a good quantity to conduct GRB classification. That is, the long–short GRB classification scheme does not directly represent the intrinsic physical classification scheme (Zhang et al. 2007; Lü et al. 2010; Qin et al. 2013).

To cross-check the results produced on Fermi-GBM, we also analyzed the first digit distributions for the GRB data observed with the Burst Alert Telescope (BAT) onboard the Swift satellite (please see Appendix for more details). We found that (i) the derived p-values change quantitatively, though the qualitative results and conclusions remain the same for the duration and fluence distributions of the overall GRB data, independent of what kind of GRB mission is considered; (ii) Benford's law is still followed by the fluence distributions of both long and short GRBs, for each GRB mission; (iii) the duration distributions of the long and short GRB groups observed with Swift-BAT are generally consistent with Benford's law, but those of the Fermi-GBM sample deviate from this law. Since the duration distributions of long and short GRBs for samples observed with different missions do not always apply to Benford's law, we emphasize again that the duration classification scheme does not always match the intrinsic physical classification scheme.

Acknowledgments

We are grateful to the anonymous referee for helpful comments. This work is partially supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (grant No. XDB0550400), the National Natural Science Foundation of China (NSFC, Grant Nos. 12373053 and 12321003), the Key Research Program of Frontier Sciences (grant No. ZDBS-LY-7014) of Chinese Academy of Sciences, and the Natural Science Foundation of Jiangsu Province (grant No. BK20221562).

Appendix
First Digit Distributions for Swift-BAT GRBs

To cross-check the results produced by Fermi-GBM, we also analyze the first digit distributions for the GRB data observed with Swift-BAT. For the Swift-BAT sample, the durations (in units of s) and fluences (in units of erg cm^{-2}) in the 15–150 keV energy range are taken from the online burst catalog.[4] The data set contains 1627 GRBs up to 2024 January. We remove 143 bursts for which no duration or fluence measurements were available. In total, there are 1484 GRBs (including 1354 long bursts with $T_{90} > 2$ s and 130 short ones with $T_{90} \leqslant 2$ s) for us to perform the first digit analysis.

The first digit distributions of the duration and fluence for all 1484 GRBs, 1354 long GRBs, and 130 short GRBs are illustrated in Figures A1, A2, and A3, respectively. A tabular summary of our Benford analyses for the Swift-BAT sample can be found in Table A1. The Pearson χ^2 tests and the corresponding p-values indicate that the durations and fluences for the overall sample and the subsamples of long and short GRBs are all roughly consistent with Benford's law. As shown in Figure A3, the digits 2, 5, and 9 are smaller than the expected Benford distributions. That is, the duration and fluence of short GRBs do not seem to fit Benford's law by eye, but their corresponding p-values support they do. The optical illusion may be caused by the relatively small sample size of short GRBs.

The comparison between the Fermi-GBM and Swift-BAT samples may be summarized as follows: (i) the first digit

[4] https://swift.gsfc.nasa.gov/archive/grb_table/

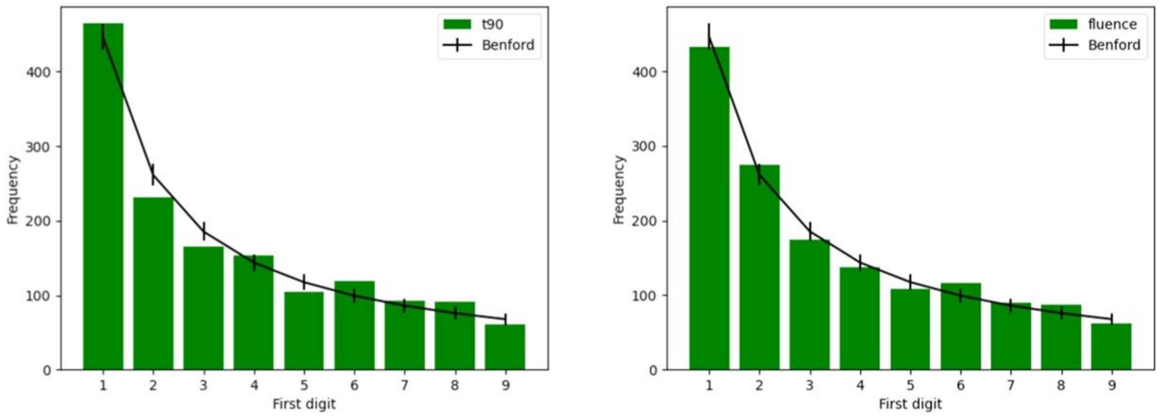

Figure A1. Distributions of the first digit of the duration T_{90} and fluence of all 1484 GRBs from the Swift-BAT burst catalog. The theoretical predictions from Benford's law (solid lines) along with associated binomial errors are also shown for comparisons.

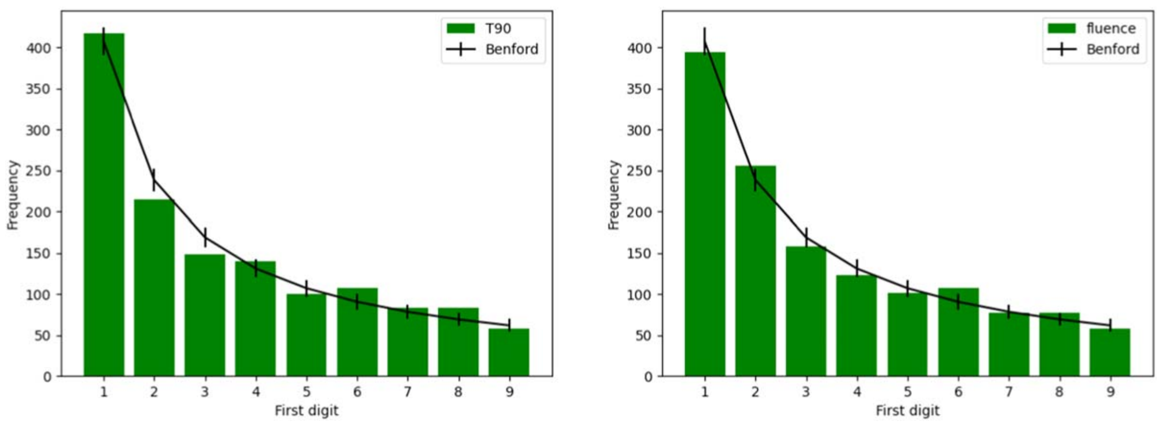

Figure A2. Same as Figure A1, but now for 1354 long GRBs (with $T_{90} > 2$ s) from the Swift-BAT burst catalog.

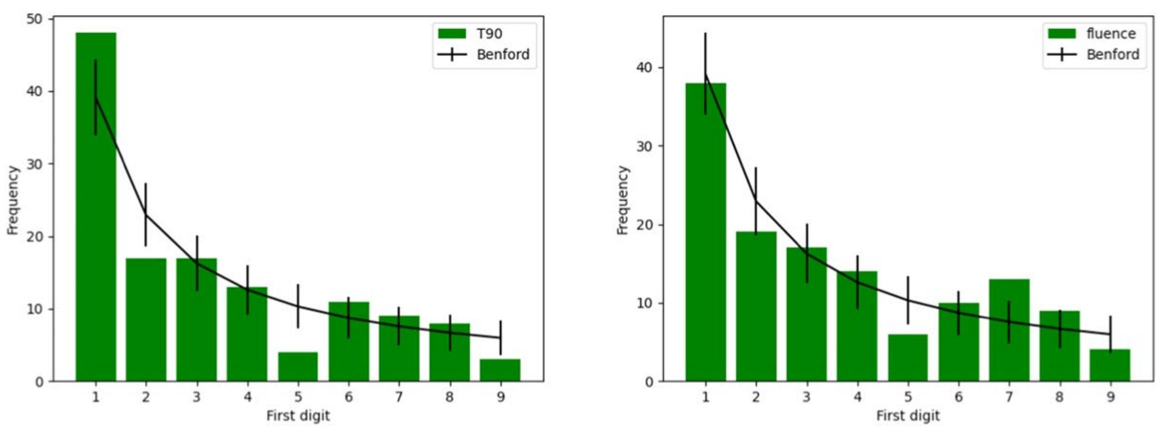

Figure A3. Same as Figure A1, but now for 130 short GRBs (with $T_{90} \leqslant 2$ s) from the Swift-BAT burst catalog.

Table A1
Summary of our Benford Analyses on the First Digit Distributions of the Duration and Fluence of GRBs Detected by Swift-BAT

Dataset	Physical Quantity	Number	χ^2/dof	p-value
All GRBs	Duration	1484	16.4/8	0.04
	Fluence	1484	8.3/8	0.41
Long GRBs	Duration	1354	12.7/8	0.12
	Fluence	1354	7.5/8	0.48
Short GRBs	Duration	130	10.0/8	0.26
	Fluence	130	8.3/8	0.41

distributions of the duration and fluence of the overall GRB data conform to Benford's law, independent of what kind of GRB mission is considered; (ii) Benford's law is still followed by the fluence distributions of both long and short GRBs, for each GRB mission; (iii) the duration distributions of the long and short GRB groups observed with Fermi-GBM do not obey Benford's law, but those of the Swift-BAT sample are generally consistent with this law. Note that the ratios of short-to-long GRB numbers with a division of $T_{90} = 2$ s for the Fermi-GBM and Swift-BAT samples are 603:3061 (1:5.1) and 130:1354 (1:10.4), respectively. Obviously, the T_{90} distribution is instrument dependent. Again, our results suggest that the duration classification scheme does not always match the intrinsic physical classification scheme. If it does, then the duration distributions of both long and short GRBs for samples observed with different missions should always adhere to Benford's law.

References

Alexopoulos, T., & Leontsinis, S. 2014, JApA, 35, 639
Band, D., Matteson, J., Ford, L., et al. 1993, ApJ, 413, 281
Barlow, J. L., & Bareiss, E. H. 1985, Compu, 34, 325347
Benford, F. 1938, PAPhS, 78, 551
Berger, A., & Hill, T. P. 2007, Am. Math. Mon., 114, 588
Berger, A., Hill, T. P., & Morrison, K. E. 2008, J. Theor. Probab., 21, 91
Buck, B., Merchant, A. C., & Perez, S. M. 1993, EJPh, 14, 59
Burgos, A., & Santos, A. 2021, AmJPh, 89, 851
Burke, J., & Kincanon, E. 1991, AmJPh, 59, 952
Cáceres, J. L. H., García, J. L. P., Ortiz, C. M., & Domínguez, L. G. 2008, Biomed., 1, 27
Dantuluri, A., & Desai, S. 2018, PhyA, 506, 919
De Ceuster, M. J. K., Dhaene, G., & Schatteman, T. 1998, J. Empir. Finance, 5, 263
de Jong, J., de Bruijne, J., & De Ridder, J. 2020, A&A, 642, A205
Eichler, D., Livio, M., Piran, T., & Schramm, D. N. 1989, Natur, 340, 126
Geyer, C. L., & Williamson, P. P. 2004, Commun. Stat. - Simul. Comput., 33, 229
Gruber, D., Goldstein, A., Weller von Ahlefeld, V., et al. 2014, ApJS, 211, 12
Hill, T. P. 1995, StaSc, 10, 354
Kouveliotou, C., Meegan, C. A., Fishman, G. J., et al. 1993, ApJL, 413, L101
Ley, E. 1996, The American Statistician, 50, 311
Lü, H.-J., Liang, E.-W., Zhang, B.-B., & Zhang, B. 2010, ApJ, 725, 1965
Mamidipaka, P., & Desai, S. 2023, APh, 144, 102761
Melita, M. D., & Miraglia, J. E. 2021, NewA, 89, 101654
Moret, M. A., de Senna, V., Pereira, M. G., & Zebende, G. F. 2006, IJMPC, 17, 1597
Narayan, R., Paczynski, B., & Piran, T. 1992, ApJL, 395, L83
Narayana Bhat, P., Meegan, C. A., von Kienlin, A., et al. 2016, ApJS, 223, 28
Newcomb, S. 1881, AmJM, 4, 39
Ni, D., & Ren, Z. 2008, EPJA, 38, 251
Ni, D.-D., Wei, L., & Ren, Z.-Z. 2009, CoTPh, 51, 713
Nigrini, M. J. 1996, J. Am. Tax. Assoc., 18, 72
Nigrini, M. J., & Mittermaier, L. J. 1997, Audit.-a J. Pract. Theory, 16, 52
Paczyński, B. 1998, ApJL, 494, L45
Pinkham, R. S. 1961, Ann. Math. Stat., 32, 1223
Qin, Y., Liang, E.-W., Liang, Y.-F., et al. 2013, ApJ, 763, 15
Schatte, P. 1988, J. Inf. Process. Cybern., 24, 443
Shao, L., & Ma, B.-Q. 2009, MPLA, 24, 3275
Shao, L., & Ma, B.-Q. 2010a, APh, 33, 255
Shao, L., & Ma, B.-Q. 2010b, PhyA, 389, 3109
Shukla, A., Pandey, A. K., & Pathak, A. 2017, JApA, 38, 7
Sottili, G., Palladino, D. M., Giaccio, B., & Messina, P. 2012, MaGeo, 44, 619
von Kienlin, A., Meegan, C. A., Paciesas, W. S., et al. 2014, ApJS, 211, 13
von Kienlin, A., Meegan, C. A., Paciesas, W. S., et al. 2020, ApJ, 893, 46
Woosley, S. E., & Bloom, J. S. 2006, ARA&A, 44, 507
Zhang, B., Zhang, B.-B., Liang, E.-W., et al. 2007, ApJL, 655, L25