First Digit Distributions of Gamma-Ray Bursts (Postprint) Abstract: This study investigates the first-digit distributions of gamma-ray bursts (GRBs) to test their conformity with Benford's Law. We analyze observational data from three satellites—BATSE, Swift, and Fermi—including physical quantiti...

Authors: Hou-Yu Lai and Jun-Jie We

Date: 2024-05-24T00:00:00+00:00

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 T90 > 2 s and T90 ≤ 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.

Full Text

Preamble

Research in Astronomy and Astrophysics, 24:055007 (6pp), 2024 May

Ltd. Printed in China and the U.K.
https://doi.org/10.1088/1674-4527/ad386c

ChinaXiv 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 first significant digits from real-world data sources is typically not uniformly distributed, but instead follows a logarithmic distribution known as Benford's law. In this work, we perform the first comprehensive investigation of the first-digit distributions of the duration, fluence, and energy flux of gamma-ray bursts (GRBs). For a complete GRB sample detected by the Fermi satellite, we find that the first digits of duration and fluence adhere to Benford's law. However, the energy flux shows a significant departure from this law, likely because a considerable fraction of the energy flux measurements is restricted by insufficient spectral information. Based on the conventional duration classification scheme, we also examine whether the durations and fluences of long and short GRBs (with duration T90 $>$ 2 s and T90 $\leq$ 2 s, respectively) obey Benford's law. We find that while the fluences of both long and short GRBs still agree with the Benford distribution, their durations do not follow Benford's law. Our results suggest 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 assume that the first significant digits (i.e., 1, 2, ..., 9) of any randomly chosen dataset would be uniformly distributed, but this is not true for natural phenomena. As early as 1881, Simon Newcomb observed an unexpected pattern in the first digits of logarithm tables: the number 1 appears more frequently than 2, 2 more frequently than 3, and so on (Newcomb 1881).

More than half a century later, Frank Benford rediscovered this logarithmic distribution of first digits in numerous data tables, a phenomenon now known as the first-digit law or Benford's law (Benford 1938). This law states that for a given real dataset, the probability of numbers having first digit k is expressed as (Benford 1938):

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

Empirically, Benford's law has been verified across diverse research fields, including geography (e.g., river lengths and lake areas; 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 bacterial gene data lengths; Cáceres et al. 2008), seismology (e.g., recurrence times of seismic events; Sottili et al. 2012), and 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 lepton branching fractions; Dantuluri & Desai 2018). Practically, this peculiar law has been effectively used to detect fraud in taxation and accounting (Nigrini 1996; Nigrini & Mittermaier 1997; Geyer & Williamson 2004) and to minimize storage space and accelerate calculations in computer science (Barlow & Bareiss 1985; Schatte 1988; Berger & Hill 2007). Theoretically, the law has been well explained using a central-limit-like theorem for first digits (Hill 1995) and a simple Markov process (Burgos & Santos 2021). Mathematically, Benford's law is scale-invariant (Berger et al. 2008), meaning it is independent of any particular choice of units (Pinkham 1961).

In astronomy, Benford's law has been extensively applied to various astrophysical datasets, including 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), and Gaia Data Release 2 (DR2) parallaxes (de Jong et al. 2020). Nevertheless, some types of data, such as pulsar and fast radio burst dispersion measures, do not obey Benford's law (Mamidipaka & Desai 2023).

In this work, we investigate for the first time the first-digit distributions of the duration, fluence, and energy flux of gamma-ray bursts (GRBs) and test whether these digits conform to Benford's law. GRBs are flashes of high-energy radiation originating from the most energetic explosions in the universe. According to their duration T90 (the time interval containing 90% of the prompt emission), GRBs are classified into long GRBs (T90 > 2 s) and short GRBs (T90 ≤ 2 s) (Kouveliotou et al. 1993). Generally, long GRBs are believed to be powered by the core collapse of massive stars (Paczyński 1998; Woosley & Bloom 2006), while short GRBs are thought to originate from mergers of binary compact objects (Eichler et al. 1989; Narayan et al. 1992).

2.1. Dataset

We downloaded the durations T90 (in units of s), fluences F (in units of erg cm^{-2}), and energy fluxes $P\gamma$ (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). The Fermi-GBM burst catalog comprises 3665 cosmic GRBs that occurred between 2008 July 12 and 2023 December 6. We removed one GRB for which no relevant data were available, leaving 3664 GRBs with T90 and fluence measurements.

The energy flux of each burst in the observer frame is calculated by integrating the spectral model over the 10–1000 keV band:

$$P_\gamma = \int_{E_{\min}}^{E_{\max}} N(E)\,dE,$$

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 and E are 10 keV and 1000 keV, respectively. Spectral parameters are required to calculate the energy flux, but not every burst has the requisite spectral information. Consequently, only 2298 GRBs in the catalog have energy flux measurements.

2.2. Results

The first-digit distributions of duration and fluence for the complete GRB sample are presented in Figure 1 and Table 1. As noted above, there are N = 3664 available GRBs. The expected number according to Benford's law, NBen = N $P(k)$, along with the root-mean-square error estimated from the binomial distribution, $\sigma = \sqrt{}(N\ \ P(k)[1-P(k)])$, are also shown in the figure. Figure 1 demonstrates that the observed distributions are well consistent with theoretical predictions from Benford's law.

To quantify the goodness of fit to Benford's law, we adopt the Pearson ^{2} statistic:

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

where N bs and NBen are the observed and expected Benford numbers for a single digit k, respectively. For the first-digit distribution of duration, we obtain a Pearson ^{2} value of 13.0 for 8 degrees of freedom (dof). For the fluence distribution, we obtain a ^{2} value of 12.5. These two ^{2} values correspond to p-values of 0.11 and 0.13, respectively, which strongly support the null hypothesis that the durations and fluences of the complete GRB sample follow Benford's law. It is worth emphasizing that higher p-values indicate greater support for the null hypothesis. In this study, we reject the null hypothesis if $p < 0.05$ (equivalent to the 95% confidence level).

GRBs are divided into two classes—long GRBs and short GRBs—with a division at T90 = 2 s (Kouveliotou et al. 1993). Using this conventional division (T90 > 2 s for long GRBs and T90 ≤ 2 s for short GRBs), we find 3061 long GRBs and 603 short GRBs in the Fermi-GBM burst catalog. We also test whether the durations and fluences of long and short GRBs obey Benford's law. The first-digit distributions of duration and fluence for long and short GRBs are depicted in Figures 2 and 3, respectively. As listed in Table 1, the ^{2} tests and corresponding *p*-values for fluences are extremely supportive of 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 clearly deviate from Benford's law. We find Pearson ^{2} values of 19.9 and 74.2 for the durations of 3061 long GRBs and 603 short GRBs, respectively, corresponding to *p*-values of 0.01 and 7×10^{-13}. Based on these *p*-values, we can safely reject the null hypothesis.

Similarly, Figure 4 and Table 1 show the first-digit distribution of energy flux for the 2298 available GRBs. While the relative ranking of digit probabilities roughly agrees with Benford's law, the distribution is not strictly obeyed. The large ^{2} value and 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 of the first-digit distributions of duration, fluence, and energy flux of GRBs. For the complete GRB sample detected by Fermi-GBM, our results show that the first digits of duration and fluence are not uniformly distributed; instead, smaller digits appear more frequently than larger ones according to the logarithmic distribution expected from Benford's law. 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 this 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 fully compliant with Benford's law.

The Fermi-GBM burst catalog reveals a bimodal T90 distribution, confirming the existence of two GRB classes—long versus short GRBs with a separation at about 2 s (Kouveliotou et al. 1993; von Kienlin et al. 2020). Using the conventional long–short GRB classification scheme, we found 3061 long GRBs (with T90 > 2 s) and 603 short GRBs (with T90 ≤ 2 s). We also examined whether the first digits of duration and fluence for long and short GRBs adhere to Benford's distribution. We found that while the fluences of both long and short GRBs still follow Benford's law, their durations are no longer consistent with this law. Our results indicate that the fluence data appear very natural and believable, but T90 is not always a good quantity for GRB classification. 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).

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 the 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 from Fermi-GBM, we also analyzed the first-digit distributions for GRB data observed with the Burst Alert Telescope (BAT) onboard the Swift satellite. 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 were taken from the online burst catalog. The dataset contains 1627 GRBs up to 2024 January. We removed 143 bursts for which no duration or fluence measurements were available, leaving 1484 GRBs (including 1354 long bursts with T90 $>$ 2 s and 130 short bursts with T90 $\leq$ 2 s) for our first-digit analysis.

The first-digit distributions of 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 is provided in Table A1. The Pearson ^{2} tests and corresponding p-values indicate that the durations and fluences for the overall sample and for the long and short GRB subsamples 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. The duration and fluence of short GRBs do not appear to fit Benford's law by eye, but their corresponding p-values support that they do. This optical illusion may be caused by the relatively small sample size of short GRBs.

The comparison between the Fermi-GBM and Swift-BAT samples can be summarized as follows: (i) the first-digit distributions of duration and fluence for the overall GRB data conform to Benford's law, independent of the GRB mission considered; (ii) Benford's law is still followed by the fluence distributions of both long and short GRBs for each 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 at T90 = 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 T90 distribution is instrument-dependent. Again, our results suggest that the duration classification scheme does not always match the intrinsic physical classification scheme. If it did, the duration distributions of both long and short GRBs for samples observed with different missions would 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,

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,
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

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.