Is a data set distributed as a power law? A test, with application to gamma-ray burst brightnesses

Wijers, R.A.M.J.; Lubin, L.M.

DOI
10.1086/174561

Publication date
1994

Published in
Astrophysical Journal

THE ASTROPHYSICAL JOURNAL, 432:207–212, 1994 September 1

IS A DATA SET DISTRIBUTED AS A POWER LAW? A TEST, WITH APPLICATION TO GAMMA-RAY BURST BRIGHTNESSES

RALPH A. M. J. WIJERS[1] AND LORI M. LUBIN[2]
Princeton University Observatory, Peyton Hall, Princeton, NJ 08544-1001
Received 1993 November 17; accepted 1994 March 7

ABSTRACT

We present a method to determine whether an observed sample of data is drawn from a parent distribution that is a pure power law. The method starts from a class of statistics which have zero expectation value under the null hypothesis, H_0, that the distribution is a pure power law: $f(x) \propto x^{-\alpha}$. We study one simple member of the class, named the "bending statistic" B, in detail. It is most effective for detection a type of deviation from a power law where the power-law slope varies slowly and monotonically as a function of x. Our estimator of B has a distribution under H_0 that depends only on the size of the sample, not on the parameters of the parent population, and is approximated well by a normal distribution even for modest sample sizes. The bending statistic can therefore be used to test whether a set of numbers is drawn from any power-law parent population.

Since many measurable quantities in astrophysics have distributions that are approximately power laws, and since deviations from the ideal power law often provide interesting information about the object of study (e.g., a "bend" or "break" in a luminosity function, a line in an X- or gamma-ray spectrum), we believe that a test of this type will be useful in many different contexts. In the present paper, we apply our test to various subsamples of gamma-ray burst brightnesses from the first-year BATSE catalog and show that we can only marginally detect the expected steepening of the log $N(>C_{\max})$–log $C_{\max}$ distribution.

Subject headings: gamma rays: bursts — methods: data analysis — methods: statistical

1. INTRODUCTION

There are many situations in astrophysics where to first approximation a quantity has a power-law distribution. Examples are galaxy luminosities, correlation functions, energies of X-ray photons from hot, optically thin plasmas, fluxes of gamma-ray bursts as detected by any single instrument, energies of cosmic rays, and spectra of many types of radio source. In some cases the reason is known (X-ray bremsstrahlung, cosmic rays, radio spectra), whereas in others we do not understand why the distribution has this shape. Unfortunately, a power law is scale-free and one of the most featureless of all possible distributions, usually revealing little about the underlying process.

Because of this, it is often quite helpful to find deviations from pure power-law behavior, as these define a characteristic value for the quantity of interest; such a value in turn may provide clues to the underlying physics. Examples are a break in an X-ray spectrum indicating the temperature of the emitting medium, a turnover in a radio synchrotron spectrum indicating (among others) a characteristic electron energy or self-absorption, a "maximum energy" for cosmic rays, slope changes in the angular correlation function of distant objects, etc. In cases where the underlying reason for the power-law distribution is not known, a bend in the distribution may provide a hint of what the reason is. A case in point, to which we shall return in § 4, is the distribution of gamma-ray burst brightnesses. First we introduce the statistics (§ 2) and examine the properties of one, the bending statistic (§ 3).

[1] E-mail: rw@astro.princeton.edu.
[2] E-mail: lml@astro.princeton.edu.

2. A CLASS OF SPECIAL STATISTICS

Consider a random variable x that is distributed as a power law, with a cumulative distribution

$$F(x) = 1 - \left(\frac{x}{x_-}\right)^{1-\alpha} \quad (x \geq x_-) , \tag{1}$$

with the restrictions that $x_- > 0$ and $\alpha > 1$. Let x_q be the location of the qth quantile, i.e., $F(x_q) = q$. Then we find

$$1 - \alpha = \frac{\ln(1-q)}{\ln(x_q/x_-)} \equiv g(x_q) . \tag{2}$$

Now we write the last relation for a set of q's and multiply each relation by a weight w_j. If the weights sum to zero, it follows that

$$P(\boldsymbol{q}, \boldsymbol{w}) \equiv K \sum_{j=1}^{m} w_j g(x_{q_j}) = (1-\alpha)K \sum_{j=1}^{m} w_j = 0 , \tag{3}$$

where $\boldsymbol{q} = (q_1, q_2, \ldots, q_m)$, and K is an arbitrary statistic. We have now defined a whole class of quantities that are identically zero for a power-law distribution but that can only be computed using exact knowledge of the properties of the parent population. We can define analogous quantities for a sample drawn from the parent population by simply replacing the quantiles x_q in equation (3) by estimators $\tilde{x}_q$ and K by an estimator k. Depending on the application, x_- may have to be estimated from the data as well. We then obtain estimators

$$p(\boldsymbol{q}, \boldsymbol{w}) = k \sum_{j=1}^{m} w_j g(\tilde{x}_{q_j}) \tag{4}$$

of P, which will have approximately zero expectation value for samples drawn from a power-law distribution, and nonzero expectation values for samples from other distributions. (Note: strictly speaking, for any choice of $\boldsymbol{q}$, one could find nonpower-law distributions that happen to have $P(\boldsymbol{q}, \boldsymbol{w}) = 0$, but only for a power-law distribution would this hold regardless of the choice of $\boldsymbol{q}$.) We can now construct tests of this null hypothesis H_0: "this sample is drawn from that power-law parent distribution." This is done by investigating the distribution of some particular $p(\boldsymbol{q}, \boldsymbol{w})$ assuming that H_0 is valid. We can then determine within what range the value of $p(\boldsymbol{q}, \boldsymbol{w})$ should fall under H_0 with some specified confidence. This range can then be used as a confidence interval for accepting H_0, i.e., if and only if the value of p measured for the real data falls within the specified range will we accept H_0.

Which of the infinite number of possible P's is the best test statistic clearly depends on the circumstances. For example, if we have a specific alternative hypothesis H_1, the obvious criterion for choosing $(\boldsymbol{q}, \boldsymbol{w})$ is to maximize the power of the test, i.e., to minimize the probability of accepting H_0 if H_1 is true (while keeping the probability of unjustly rejecting H_0 fixed, of course).

3. A SPECIFIC CASE: THE BENDING STATISTIC B

3.1. *Properties*

In the present paper we have no well-defined H_1. We expect that the deviation of the brightness distribution of gamma-ray bursts from a power law (if any) will be smooth, i.e., it will take the form of an α that varies montonically with brightness. We should therefore choose values of q that are wide apart in order to sample as much as possible of the slowly accumulating curvature. But in absence of a definite H_1, we want to minimize the variance of p to improve our chances of any deviation of p from 0 being signficant. This implies that the q's should not be close to 0 or 1 because the corresponding $\tilde{x}_q$ would have a large variance (Lupton 1993, p. 43).

As a compromise between the requirements, we choose $m = 2$ and $\boldsymbol{q} = (1/3, 2/3)$. For two elements, we can set $\boldsymbol{w} = (-1, 1)$, since a multiplicative factor in p is of no consequence. We choose $K = \ln(x_{1/3}/x_-)/\ln(1/3)$ and thus define the bending statistic to be

$$B \equiv \frac{\ln(x_{1/3}/x_-)}{\ln(x_{2/3}/x_-)} - \frac{\ln(2/3)}{\ln(1/3)} . \tag{5}$$

A more intuitive (and our initial) way for arriving at this result is to consider equation (1) for $x = x_{1/3}$ and solve for $1 - \alpha$. This means that α is fixed once the location of $x_{1/3}$ (and x_-) is known. We can now use this α with equation (1) to predict the location of $x_{2/3}$. Requiring this prediction to coincide with the actual value of $x_{2/3}$ amounts to requiring $B = 0$ after a little rewriting. In this way, it is also easily seen that $B > 0$ corresponds to a steepening of the power law with x, because in that case the predicted value of $x_{2/3}$ exceeds the actual value. It is also clear from this reasoning that two is the minimum number of quantiles needed to test the hypothesis, H_0, that a data set is distributed as a power law.

To compute b, the estimator of B, we must derive estimators for $x_{1/3}$, $x_{2/3}$, and x_- from a sample $\{x_i\}_{i=1}^n$. We estimate $\tilde{x}_{1/3}$ and $\tilde{x}_{2/3}$ by interpolating linearly between the two elements in the sample that bracket them. The quantity $\tilde{x}_-$ is estimated by solving the relation $g(x_{1/n}) - g(x_{1/3}) = 0$ for x_- and then replacing all quantities by their estimators in the resulting expression. The estimator for $x_{1/n}$ is $x_{(1)}$, the smallest element in the sample, and $\tilde{x}_-$ is essentially equal to $x_{(1)}$ with a small correction (of order $1/n$) because one expects the smallest element in the sample to be slightly greater than the minimum possible value, x_-.

It is clear that we can scale all the values in the sample by a constant factor without affecting the value of b. Therefore, the distribution of b will not be affected by the value of x_- except that the use of an estimator $\tilde{x}_-$ will increase the variance somewhat relative to the case where the exact value is known and possibly affect the bias. We will therefore concentrate on investigating the dependence of b on α and n. We chose the 13 α-values 1.2, 1.25, 1.3, 1.4, 1.5, 1.6, 1.7, 1.75, 1.8, 1.9, 2.0, 2.25, and 2.5 and the five n-values 30, 100, 300, 1000, and 3000, and set $x_- = 1$. For each combination (α, n) we generated 10^5 Monte Carlo samples of n elements from a power-law distribution with exponent α and then computed the value of b for each sample in the above manner.

In Figure 1 we show the estimated mean m_b and standard deviation s_b of the $f(n, \alpha; b)$ as a function of n. The expectation value of b indeed tends to zero for large samples, but for finite samples there is a small bias that decreases with sample size roughly as $1/n$; it varies in magnitude from 8% to 1% of the standard deviation as n runs from 30 to 3000 and is therefore negligible. The standard deviation decreases with sample size as $n^{1/2}$, as expected, except that there are small deviations from that trend for sample sizes $n \lesssim 100$. If we include all n in the fit to $s_b(n)$, we find almost the same slope ($c_1 = -0.5013 \pm 0.0002$, see Fig. 1) but the fit is poor ($\chi^2/\nu = 6.96$).

The most striking result is that m_b and s_b are *independent of* α. This is very useful because it means we need not specify α to set confidence intervals on b for rejecting H_0. In other words, we can specify one test for the hypothesis "this sample is drawn from a power-law distribution with unknown slope" rather than only for the much more restricted hypothesis of the type "this sample is drawn from a power-law distribution with slope 1.23."

It would be convenient if the distribution $f(b)$ resembled some well-known distribution, so that confidence intervals for H_0 could easily be determined from standard tables rather than from extensive Monte Carlo simulations. In Figure 2 we show histograms of $f(n, \alpha; b)$ for the four smallest n as derived from the 10^5 Monte Carlo samples. We scale out the already known variations of the distribution with sample size by plotting the distribution as a function of $x = (b - m_b)/s_b$. Also plotted are the relative differences between the scaled $f(x)$ and a standard normal distribution. Deviations of $f(x)$ from a standard normal distribution are noticeable for the smaller sample sizes and in the far wings of the distribution. Furthermore, each panel actually contains three plots of $f(x)$ and its deviation from normality, for $\alpha = 1.25$, 1.75, and 2.25. It is clear that, where $f(x)$ differs from a standard normal distribution, the difference is the same for all α, i.e., the independence of $f(n, \alpha; b)$ on α extends beyond the mean and variance and holds true for the distribution in detail.

To better determine when one can safely set confidence regions for H_0 by approximating $f(b)$ with a normal distribution, we show in Figure 3 what the positions of $\tilde{x}_q$ are for a special set of nine q-values, namely those for which x_q would equal $(-4, -3, \ldots, 4)$ if $f(b)$ were exactly a standard normal distribution. The $\tilde{x}_q$ are again shifted by m_b and scaled to s_b. The points corresponding to identical values of q and α are connected, and it is clear that the quantiles are essentially equal to those of a standard normal distribution for sample sizes $n \gtrsim 100$. A final formal test of the absence of correlation

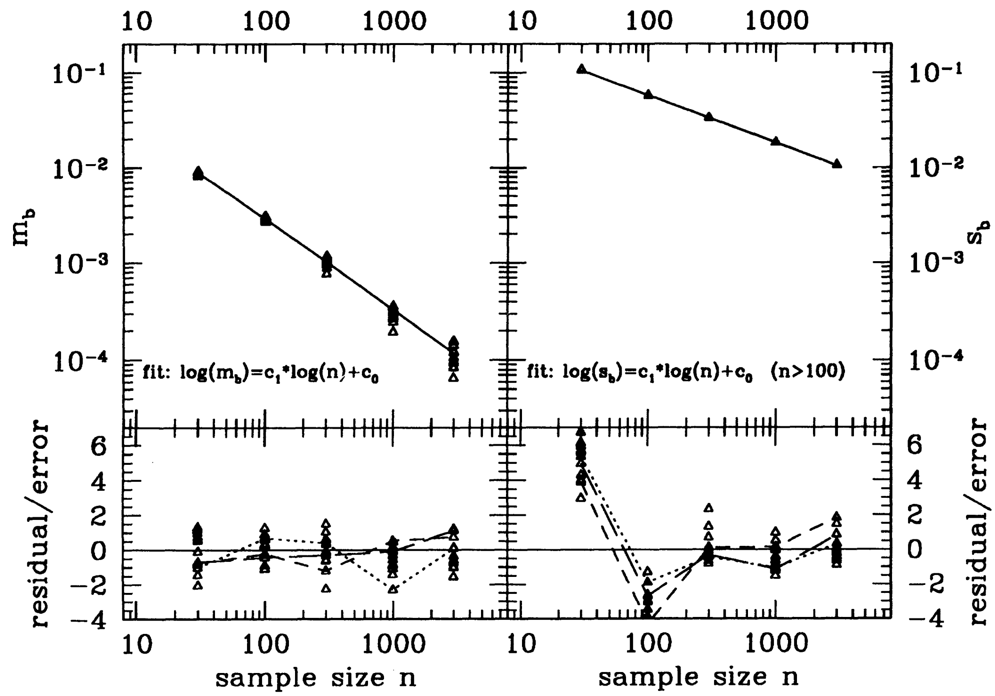

FIG. 1.—The mean (*left*) and standard deviation (*right*) of b as a function of sample size n, and the residuals of each relative to the fitted line. The fit in both panels treats each combination (n, α) as a separate data point, but points for $n \leq 100$ are excluded in the right panel (see text). The fit for the means has $c_1 = -0.936 \pm 0.009$, $c_0 = -0.674 \pm 0.018$, $\chi^2/\nu = 0.78$; for the standard deviation $c_1 = -0.4999 \pm 0.0004$, $c_0 = -0.2385 \pm 0.0011$, $\chi^2/\nu = 0.72$. The lack of dependence on α is illustrated in the bottom panels by connecting the fit residuals for the α-values 1.25 (*dotted*), 1.75 (*dashed*), and 2.25 (*long-dashed*). Note that the residuals have been divided by the estimated errors in the values of m_b and s_b (which are so small for s_b that all points coincide on the top plot).

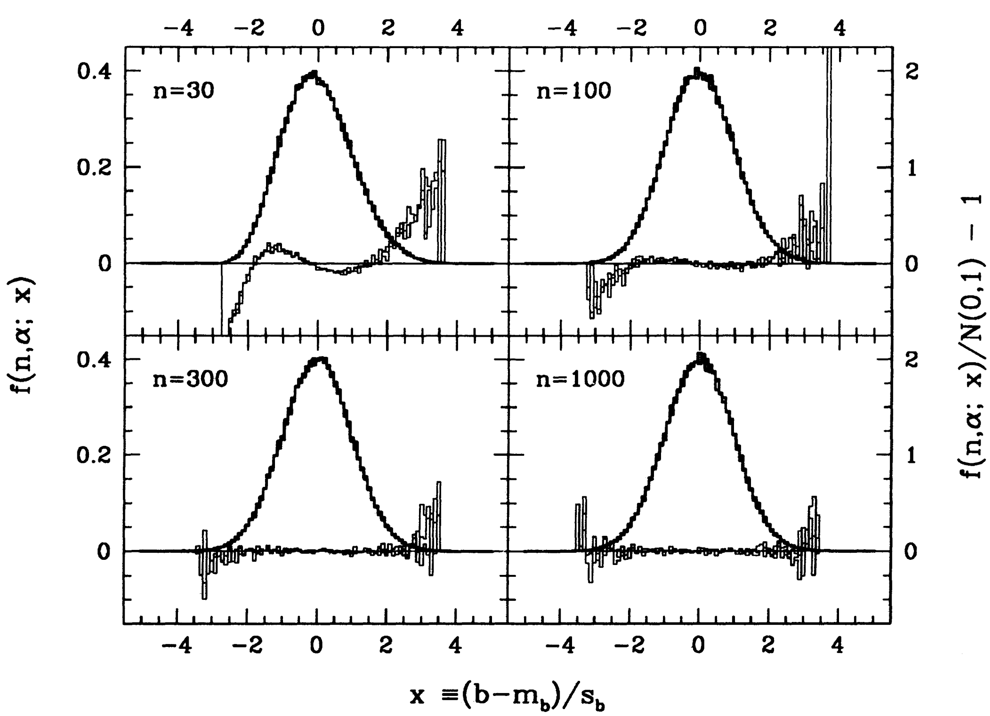

FIG. 2.—A comparison between $f(n, \alpha; b)$, shifted to zero mean and scaled to unit variance, and a standard normal distribution $N(0, 1)$ for all sample sizes [except 3000, as that case is indistinguishable from $N(0, 1)$]. Thick lines are the $f(n, \alpha; x)$; thin lines are the relative differences with $N(0, 1)$ (*right scale*). Three curves of each type are drawn in each panel, corresponding to $\alpha = 1.25$, 1.75, and 2.25; notice that the three are identical up to noise fluctuations.

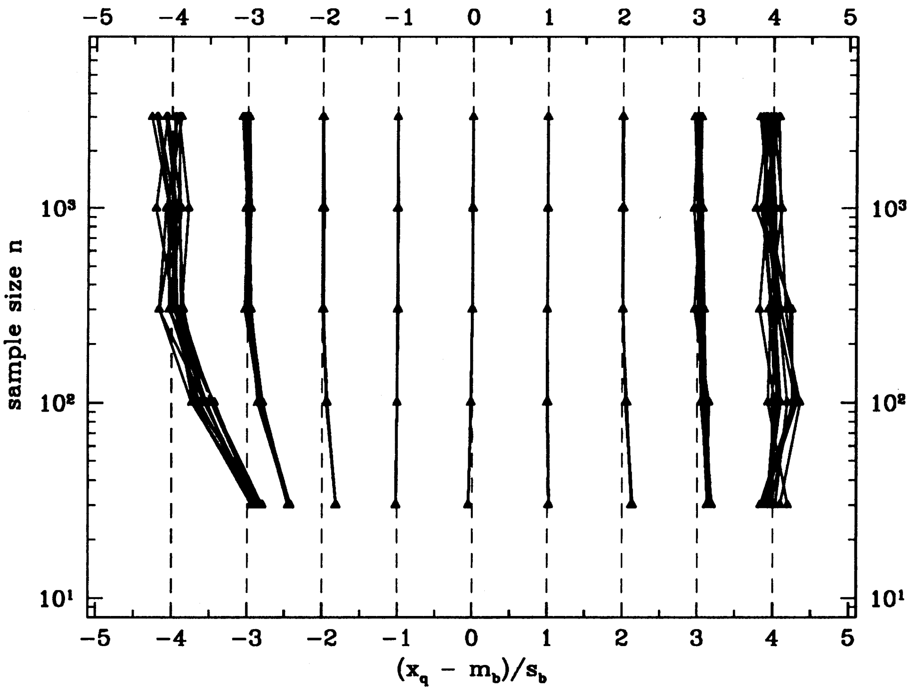

FIG. 3.—Scaled values of the quantiles $\tilde{x}_q$ as a function of sample size. The lines connect data with equal values of q and α. Dashed lines are drawn for reference at the position where x_q would be for a standard normal distribution. It is seen that the quantiles are quite consistent with a standard normal distribution for $n \gtrsim 100$.

between α and $\tilde{x}_q$ was performed: for each α and n, 23 quantiles were computed (of which only nine are shown in Fig. 3). We calculated Spearman's rank correlation coefficient (see, e.g., Lupton 1993, p. 107) between α and $\tilde{x}_q$ of each set of 13 pairs $(\alpha, \tilde{x}_q)$ that have the same n and q. There are 115 such sets, and the distribution of the 115 rank correlation coefficients obtained was compared with the theoretical distribution of rank correlation coefficients expected for sets of 13 uncorrelated data points. A Kolmogorov-Smirnov test yielded a 23% probability that the 115 coefficients were drawn from the theoretical distribution of coefficients for uncorrelated data; therefore, we can safely state that the $\tilde{x}_q$ are indeed independent of α.

3.2. *Comparison with Other Statistics*

One may well ask why we should go to such length in exploring a new statistic when there are already many familiar methods of testing hypotheses and comparing model distributions to data, such as maximum-likelihood estimation and Kolmogorov-Smirnov tests.

Maximum-likelihood estimation is a good way of finding the best-fit slope of a distribution, assuming it is of power-law type. We found that it is not a good way of detecting a deviation from power-law behavior: we fitted samples drawn from a pure power-law distribution with a power-law model using maximum-likelihood estimation (see also Crawford, Jauncey, & Murdoch 1970). While this is the most accurate (i.e., lowest variance) way of estimating the power-slope from the data, the likelihood value at the best-fit slope, α_{ML}, does not reveal deviations from a power-law distribution: the best-fit likelihood is a function of the fitted α_{ML} alone for a given sample size and parent distribution, regardless of whether or not a small bend was present in the parent distribution.

Kolmogorov-Smirnov tests cannot be used directly to answer the question we pose in this paper. This is because the K-S distance between a model and data can only yield the probability that the data are drawn from the model distribution in a simple manner if no parameters of the model have been estimated from the data (e.g., if the model has free parameters which have been adjusted by fitting to the data). Since we

TABLE 1

RESULTS OF BENDING STATISTIC FOR THE THREE COMPLETE SUBSAMPLES FROM THE FIRST-YEAR BATSE CATALOG

t_{trig} (ms)	n_{tot}	C_* (counts per bin)	n_{sam}	b_{sam}	$P_{H_0}(b > b_{sam})$	$P_{H_0}(d_{KS} > d_{KS,sam})$
64.........	135	72.0	128	−0.0164	0.640(5)	0.400(5)
256.........	148	147.8	133	0.0951	0.032(2)	0.027(2)
1024.........	193	342.9	165	0.0609	0.096(3)	0.184(4)

NOTE.—t_{trig} is the duration of the time bins on which triggering for that sample is based. n_{tot} is the total number of bursts triggered on that timescale, i.e., bursts with $C_{max}/C_{min} > 1$, excluding overwrites. C_* is the uniform threshold adopted to create a complete subsample. n_{sam} is the number of bursts left in the complete subsample. b_{sam} is the value of the bending statistic, and the next column gives the probability of getting a value of $b > b_{sam}$ by chance from a pure power-law distribution. The last column is discussed in § 3.2. Figures in parentheses indicate uncertainties in the last digits of the number.

want to search for a bend independent of the best-fit slope, we cannot directly use a K-S test. What we can try is the following: find the slope, α_{KS}, that results in the smallest K-S distance to the data among all possible power-law models and record this distance d_{KS}. Now find the distribution of this minimum K-S distance under the null hypothesis that the distribution is a power law via Monte Carlo simulations in which the above fit procedure is applied to many independent samples (of the same size as the data set) drawn from a pure power-law distribution. Using this distribution we can infer what is the probability $P(d_{KS} > d_{KS,sam})$ that value of d_{KS} greater than that of the actual data sample would arise by chance under H_0. For small enough P we would reject H_0.

We applied this test to the same data as the bending test (see § 4 and Table 1). In the cases studied, both tests are on average about equally sensitive to detecting a smooth deviation from a power law. However, the outcome of the bending test contains additional information about the nature of the deviation, such as whether the bend is up or down. This is an advantage if we are seeking a specific type of deviation (as we are here). If we are interested only in whether or not there is any kind of deviation, we might prefer the K-S method. The advantage is probably small, because the lowest order smooth deviation from a pure power law is always a bend.

4. AN APPLICATION: GAMMA-RAY BURST BRIGHTNESSES

The origin of gamma-ray bursts is completely unknown, as is their distance scale. However, we do know that bright gamma-ray bursts, such as those seen by the *Pioneer Venus Orbiter* (*PVO*, Chuang et al. 1992), have $d \log N(>C_{max})/d \log C_{max} = -1.5$, and fainter bursts, such as those seen by SIGNE (Atteia et al. 1992) and BATSE (Meegan et al. 1992), have slopes that are progressively shallower. (C_{max} is the peak raw count rate of a burst at a specified time resolution.) The most natural interpretation of this behavior is that up to some characteristic distance the density of gamma-ray bursts is constant. Beyond that either the density decreases and/or effects of non-Euclidean space are visible. Comparison of peak intensity distributions between different instruments is tricky because of their different sensitivities and energy bands (see, e.g., Tamblyn & Melia 1993 and Fenimore et al. 1993 for attempts to do so). We therefore try to detect the effect of a steepening slope in the log $N(>C_{max})$–log C_{max} distribution of one data set, the first-year BATSE catalog (BATSE Team 1993), using the bending statistic discussed in § 3.1.

The first problem in this analysis is that we want to measure the bending of the true brightness distribution, while the observed one is distorted by selection effects. As a measure of burst brightness, we shall choose the peak count rate in the second brightest detector, C_{max}. It is available in the BATSE public catalog and directly determines whether a burst is triggered, so that the dominant selection effects that affect its observed distribution are relatively easy to assess. The main effect we should take into account is the varying sensitivity of the BATSE detectors due to varying backgrounds. It causes the observed log $N(>C_{max})$–log C_{max} distribution to become incomplete at the lowest values of C_{max}, which makes the slope of the distribution shallower. Since the intrinsic effect we are examining is of the same character, using a complete sample is important. We have adopted the following procedure for constructing a complete sample: we choose one of the three BATSE trigger timescales and use only bursts with $C_{max}/C_{min} > 1$ (where C_{min} is the smallest value of C_{max} for which the detection of a burst is deemed significant). The varying C_{min} for the bursts in that list is then replaced with a single uniform value C_*. This means discarding all bursts with $C_{max} < C_*$; however, we should also discard those bursts observed when $C_{min} > C_*$, because at such times only bright bursts could be seen and including them would cause bright bursts to be overrepresented. We now have a complete sample for any C_*, which corresponds to the bursts detected by a fictitious instrument that operated only when the C_{min} of BATSE was less than C_* and had a uniform threshold of C_* at those times. Any value of C_* yields a complete sample, but we fix our choice to be the one that yields the largest sample. Lynden-Bell (1971) and Petrosian (1993) have shown how to reconstruct the distribution of C_{max} without discarding bursts, but the weights of bursts at the faint end become variable through that procedure. This implies that the statistical properties of such a reconstructed sample are different from a simple power law. For the purpose of illustration, we stay with our simpler method and accept the loss of 5%–15% of the bursts. Since there are three trigger timescales, we constructed three complete subsamples in the manner described above. Their log $N(>C_{max})$–log C_{max} distributions are shown in Figure 4.

A second effect that distorts the distribution stems from the fact that C_{max} is the count rate in the second most brightly illuminated of the eight BATSE detectors. It is not corrected for the fact that the angle between the direction of the burst and the normal to that detector varies from burst to burst: for an ideal octahedron configuration of equal detectors, the cosine of this angle varies from 1/3 to $(2/3)^{1/2}$ with a mean of 0.62 and rms deviation of 0.10. This means that the observed C_{max} will vary by the same factor even if the true peak count rate of all bursts were the same. However, if the true peak count rates have a pure power-law distribution, the observed C_{max} distribution will also have one. On the other hand, any deviation from a pure power law will be smeared out somewhat, so by ignoring this effect we are less likely to detect the deviation from a pure power law.

The expectation value of b is positive if the slope of log $N(>C_{max})$–log C_{max} steepens with increasing C_{max} and negative if it flattens. We are looking for a steepening, i.e., for b significantly greater than zero, so we perform one-sided tests on b, quoting the probabilities that $b > b_{sam}$ rather than $|b| > |b_{sam}|$. We used Monte Carlo simulations to determine confidence levels because deviations from normality in the distribution of b are still noticeable with the currently available sample sizes (see § 3.1). The results of applying the bending statistic to the three complete samples described above are somewhat ambiguous (Table 1). There is an indication for a bend in the 256 ms and 1024 ms samples but none at all in the 64 ms sample. This may just be a statistical fluctuation: the evidence for a bend is not overwhelming in any of the samples, and the variation of b over the three samples could reflect the statistical fluctuations expected for an effect that is just barely measurable with the present sample sizes. On the other hand, the 256 ms sample does have the shallowest slope at the faint end, and because gamma-ray bursts have a distribution in duration, it is likely that the different trigger timescales do select slightly different bursts. Some of the differences in the distributions of C_{max} could therefore be real, but we cannot claim such an effect with any confidence on the basis of the present analysis. If the true value of b is actually between the values measured for the current 256 ms and 1024 ms samples, it will take 300–800 bursts to detect the effect with 3 σ signifi-

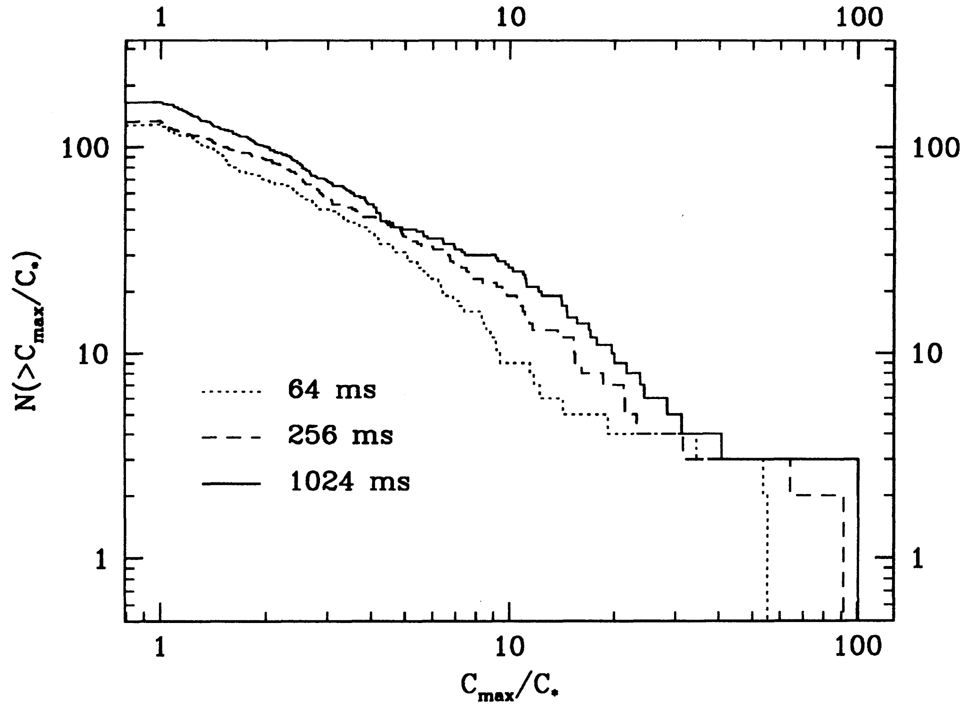

FIG. 4.—The log $N(>C_{max})$–log C_{max} distribution for the three complete subsamples derived from the first-year *BATSE* catalog.

cance. This means that it may already be possible with the complete BATSE sample that is presently available. Eventually, the steepening of the power-law slope must be found because we know from the *PVO* data that very bright bursts have a steeper log $N(>C_{max})$–log C_{max} than the BATSE data.

5. CONCLUSION AND DISCUSSION

We have found a class of statistics that can be used to construct tests to determine whether a set of data is consistent with being drawn from a power-law parent distribution. Investigation of one simple member of this class of statistics, which we named the "bending statistic," showed that its properties are independent of the parameters of the parent population. This means that it can be used to test the general hypothesis "this sample is drawn from a power-law distribution," rather than the more restricted hypothesis "this sample is drawn from a power-law distribution with power-law index α." Since this particular statistic was chosen without knowing that it would have that property, it may turn out that more complicated members of the class also have a distribution that does not depend on the power-law index. The question of which statistic in the class is the optimum choice depends on the problem at hand, especially what type of deviation from a power-law is expected.

With the sample sizes in the available first-year BATSE catalog and the values of the bending statistic measured from them, no conclusive evidence is found for a steepening of the slope of log $N(>C_{max})$–log C_{max} toward larger values of C_{max}. This is in agreement with our earlier work in which we analyzed the data within the context of a specific model in order to determine the dependence of the density of gamma-ray bursts on distance. A sample of 400 bursts was found only marginally sufficient to detect a transition between the slopes of bright and faint bursts, as evidenced by the fact that the fitted value of the core radius of the density distribution had a large error (Lubin & Wijers 1993). The method described here is somewhat more sensitive to slope changes.

We expect that the change in slope that is tentatively indicated by the longer trigger timescale data sets from the first-year BATSE catalog should be measurable with good confidence from the full set of BATSE bursts that has been collected to date.

This work was supported in part by NASA grant NAG5-1901. R. W. is supported by a Compton Fellowship (grant GRO/PFP-91-26). We thank R. Lupton and B. Paczyński for helpful discussions and an anonymous referee for valuable comments on the original manuscript.

REFERENCES

Atteia, J., Barat, C., Jourdain, E., Niel, M., & Vedrenne, G. 1992, in Gamma-Ray Bursts, ed. W. S. Paciesas & G. J. Fishman (AIP Conf. Proc. 265), 82

BATSE Team. 1993, first-year electronic BATSE catalog, GRO Science Support Center

Chuang, K. W., White, R. S., Klebesadel, R. W., & Laros, J. G. 1992, ApJ, 391, 242

Crawford, D. F., Jauncey, D. L., & Murdoch, H. S. 1970, ApJ, 162, 405

Fenimore, E. E., et al. 1993, Nature, 366, 40

Lubin, L. M., & Wijers, R. A. M. J. 1993, ApJ, 418, L9

Lupton, R. 1993, Statistics in Theory and Practice (Princeton: Princeton Univ. Press)

Lynden-Bell, D. 1971, MNRAS, 155, 95

Meegan, C. A., Fishman, G. J., Wilson, R. B., Paciesas, W. S., Pendleton, G. N., Horack, J. M., Brock, M. N., & Kouveliotou, C. 1992, Nature, 355, 143

Petrosian, V. 1993, ApJ, 402, L33

Tamblyn, P., & Melia, F. 1993, ApJ, 417, L21