LOGN-LOGS SLOPE DETERMINATION IN IMAGING X-RAY ASTRONOMY

Tommaso Maccacaro[1,2], Suzanne Romaine[1] and Jurgen H.M.M. Schmitt[3]

1) Harvard-Smithsonian Center for Astrophysics, Cambridge, MA.
2) CNR - Istituto di Radioastronomia, Bologna, Italy.
3) Max Planck Institute fur Physik und Astrophysik, Institute fur Extraterrestrische Physik, Garching bei Munchen, FRG.

ABSTRACT

We briefly discuss the problem of estimating the slope of the number-counts relations for the specific case of imaging X-ray surveys. Results have been obtained from extensive simulations of Einstein Observatory imaging X-ray data. We conclude that the bias which affects the X-ray number-counts slope determination is much smaller than that which affects the radio number-counts slope.

1. INTRODUCTION

The analysis of the number count relationship for a given class of astronomical objects is not merely an academic or strictly mathematical matter but carries important information on the nature of the sources under study and on the geometry of the Universe.

Objects which do not evolve with cosmic time and are uniformly distributed in the Universe, are characterized by a logN-logS of integral slope 1.5 as long as they are at small redshifts where the geometry of the Universe is well approximated by the Euclidean geometry. Under these hypotheses, any class of objects has the same logN-logS, regardless of its luminosity function. In this case, the amount of information derived from the study of the number-count relation is rather limited. It has allowed, however, the establishment of the extragalactic nature of the unidentified high galactic latitude sources discovered with early X-ray satellites. As soon as objects at high redshifts are sampled, a flattening of the logN-logS is expected. This flattening carries information on the geometry of the Universe and on the luminosity function of the sources. Furthermore, if the objects under consideration exhibit some form of evolution with cosmic time, then their number count relationship will again deviate from the 1.5 slope. Indeed, any departure from this slope is the signature, to be interpreted, of some effect.

2. LOGN-LOGS SLOPE DETERMINATION

It is customary to describe the flux distribution function of the underlying source population with a power law of slope α. The task is to obtain the best estimate for α. Techniques to estimate the number-counts slope have been developed mostly by radioastronomers. Crawford, Jauncey and Murdoch (1970) and Murdoch, Crawford and Jauncey (1973) (MCJ) have shown that the Maximum Likelihood Method (MLM) provides the minimum variance best estimate of the slope α and that in the presence of measurement errors, this estimate is biased and systematically overestimates the slope of the true distribution. The bias is due to the presence of a lower flux limit S_{th} in conjunction with the measured errors. Sources with fluxes just above the threshold

A. Hewitt et al. (eds.), Observational Cosmology, 597–600.

https://doi.org/10.1017/S0074180900159601 Published online by Cambridge University Press

S_{th} may be lost because of "negative" error fluctuation, while sources with fluxes just below the threshold may enter the survey because of "positive" error fluctuations. In general the two effects do not cancel out. Their relative importance depends on the number of sources present above and below the adopted threshold (i.e. on the slope α), as well as on the detailed shape of the distribution of the measurement errors. MCJ have tabulated correction factors to remove the bias in the source counts slope determination and have shown that these corrections can be determined only if sources with flux density at least 5 times the rms error (i.e. SNR$\geq$5) are accepted in a survey.

3. THE CASE OF IMAGING X-RAY SURVEYS

X-ray surveys differ in their statistical properties from surveys at longer wavelengths in several respects. First the measurement errors are determined by the integrated flux, i.e. the total number of counts recorded, rather than by the flux itself. Next and more important, the measurement error distribution is an asymmetrical Poisson distribution rather than a symmetrical Gaussian distribution. Further, it is necessary to subtract background counts since they often constitute a significant portion of the observed signal. Finally, most of the Medium Sensitivity Survey (MSS) exposures as well as large parts of the planned ROSAT all-sky survey are or will be photon limited rather than background or confusion limited. For the case of the MSS the statistical error on α exceeds at the moment the applied bias correction. However, as discussed by Gioia, Maccacaro and Wolter at this Symposium, the MSS has been extended and now contains more than 800 sources. Furthermore, the planned ROSAT all-sky survey will yield a 100 fold increase in the number of X-ray sources. The statistical error on the determination of the next generation of LogN-logS curves will be significantly smaller. It is thus worthwhile to reconsider the problem of the logN-logS slope estimation in the light of the features intrinsic to an imaging X-ray survey. We have approached the problem both analytically, analogously to MCJ, and empirically through Monte Carlo simulations. The two methods are complementary in the sense that analytical calculations provide checks of the Monte Carlo simulations under simplifying assumptions, whereas instrumental effects (i.e., non uniformity of the detector response, background inhomogeneities) can be far more easily incorporated into a Monte Carlo computation. The analytical work is presented by Schmitt and Maccacaro (1986). Here we discuss the results of our Monte Carlo simulations.

A large number of IPC images were simulated, containing point-like sources with fluxes distributed according to a logN-logS of known slope α_t. These images were analyzed with the same software used for the real IPC images (see Harnden et al. 1984). We have simulated IPC images of about 5000 s exposure (~13 background counts per detection cell). The threshold for source detection has been set to SNR = 3 (SNR is defined as the ratio between the source net counts and the square root of the source plus background counts. For a 5000 s IPC image this threshold corresponds to a flux of $\sim 2\text{x}10^{-13}$ ergs/cm^2/s in the 0.3 - 3.5 keV band. Sources were created with fluxes 5 times smaller than this value and according to 3 different flux distributions, simulating logN-logS with (integral) slopes of 1.5, 1.8, and 2.0. We are still in the process of creating simulated data to reduce the uncertainties due to the finite sampling statistics. The preliminary results we report here are based on the analysis of 20586 (α_t = 1.5), 23592 (α_t = 1.8), and 8948 (α_t = 2.0) sources. For these three cases we have derived, using the MLM, the best estimate of α using all the detected sources (SNR $\geq$ 3) or only those sources detected above SNR = 4 and SNR = 5. The results are shown in Figure 1 where the bias (α/α_t) is plotted as a function of the threshold SNR. For comparison, the bias as derived by MCJ for the case of α_t = 1.5 is also indicated. It is evident that in the case of imaging X-ray surveys the bias in the determination of the number-counts slope is much smaller than that which has been determined for radio surveys. This is mainly due to the fact that in the former case the measurement error distribution is an asymmetrical Poisson distribution while in the latter case it is a symmetrical Gaussian distribution.

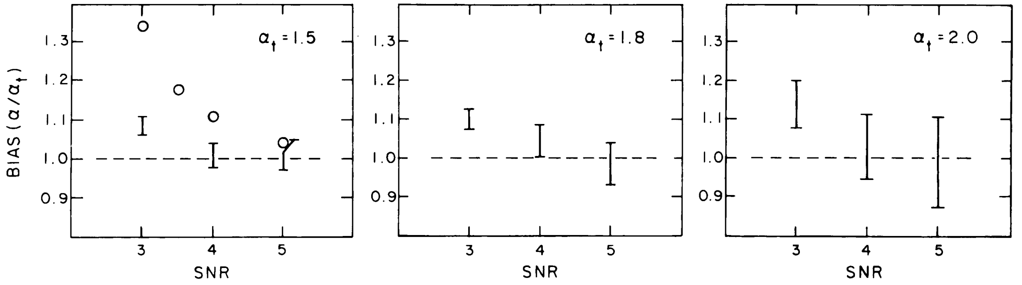

Figure 1. Bias in the slope determination versus the threshold SNR used. In the α_t = 1.5 case the bias derived by MCJ for radio surveys is also indicated (open circles).

This leads to a very small number of spurious sources detected above the threshold SNR = 3. In fact, at least 17 net counts have to be recorded on top of the 13 average background counts in order to satisfy the requirement that SNR = 3. But the probability to observe 30 counts or more, when 13 are expected is 3.8×10^{-5}, much smaller than the probability associated to the 3σ level in the Gaussian statistics. Furthermore, the algorithm used to detect sources discriminates "real" sources from "spurious" sources on the basis of the shape as well as the amplitude of the candidate source. In other words, a fluctuation which satisfies the requirement of SNR $\geq$ 3 can be rejected because the spatial distribution of the counts may be too dissimilar from the expected distribution (point response function).

4. CONCLUSION

X-ray surveys differ in their statistical properties from surveys at longer wavelengths. This requires a different treatment of the data when determining the slope of the number-counts relationship. For X-ray surveys, the bias in the determination of the logN-logS slope is much smaller than that which has been determined for radio surveys. It is possible to lower the threshold for inclusion of sources in the sample to SNR = 4 and still be able to make valid corrections for the bias. This results in an increase of 50% or more in the sample size, with a significant reduction in the errors due to finite sampling statistics. We thank Donna Irwin for her care in preparing this manuscript for publication. This work has received partial financial support from NASA contract NAS8-30751.

REFERENCES

Crawford, D.G., Jauncey, D.L., and Murdoch, H.S., 1970, Ap. J., 162, 405.

Gioia, I.M., Maccacaro, T., Wolter, A., this volume.

Harnden, F.R., Jr., Fabricant, D.g., Harris, D.E., and Schwarz, J., 1984, SAO Report 393.

Murdoch, H.S., Crawford, D.G., and Jauncey, D.L., 1973, Ap. J., 183, 1.

Schmitt, J.H.M.M., and Maccacaro, T., 1986, Ap. J., 310, in press.

DISCUSSION

SCHMIDT: Dr. Setti stated that an extrapolation of the MSS towards higher fluxes with the MSS shape yields two or three times fewer sources than Piccinotti's HEAO1-A2 counts. Can you comment on the significance of this?

MACCACARO: I am not too concerned about this disagreement. First there is a difference in the energy range of the two surveys and this contributes to uncertainty in the comparison. Second, it may not be correct to extrapolate the MSS AGN log N - log S (which has a slope steeper than the Euclidean slope) to the flux range of the Piccinotti survey. At these fluxes in fact, only very, very low redshifts are sampled and therefore the log N - log S has to have the Euclidean slope.

WAMPLER: You said that in your model the probability of getting an extra 18 counts on top of the background was 10^{-5}. In the real data some 3σ sources were removed because their "shape" was wrong. Have the real data been shown to be governed by counting statistics? If the probability of getting a false event was really 10^{-5} I would not have expected many false events.

MACCACARO: The real data are governed by what I would call "modified Poisson" statistics - the underlying statistics is obviously Poisson statistics (we are dealing with a small number of integer counts) but is "modified" in the sense that after the amplitude of a fluctuation is checked, its shape is also checked. This is likely to remove more "spurious" sources than "real" sources.

https://doi.org/10.1017/S0074180900159601 Published online by Cambridge University Press