AXAF Data Analysis Challenges

Aneta Siemiginowska[1], Martin Elvis[1], Alanna Connors[2], Peter Freeman[3], Vinay Kashyap[3], and Eric Feigelson[4]

ABSTRACT The high quality of the AXAF X-ray data provides new challenges for the X-ray data analysis. It is clear that an "old" approach is not enough to fully exploit the capabilities of the AXAF instruments. We describe a few of the statistical and computational problems that we have so far identified. Some of them appear to be theoretically solvable but computationally challenging, while others state problems for theoretical statistics which, so far as we know, are unsolved.The problems divide, from an astronomical point of view, into: Modeling the Data (e.g. nonlinear parameter estimation, uncertainties in the model, weighting the data, correlated residuals), Source Detection (events in N-space, use of wavelets, significance of detected structures) and Instrument Related Issues (pile-up in AXAF ACIS, overlapping orders in grating spectra).

1 Introduction

Study of X-ray emission from stars and galaxies requires placing highly specialized telescopes and detectors on space-based satellites, because X-rays do not penetrate the Earth's atmosphere. Following the initial discoveries of cosmic X-ray sources in the early 1960s, 28 satellite-borne X-ray missions have been launched by several nations (Bradt et al. 1992). NASA's forthcoming "Advanced X-ray Astronomy Facility" (AXAF) mission will provide the highest spatial and spectral resolution yet achieved in X-ray astronomy (see Zombeck 1996, 1982, 1979).

For the first time X-ray astronomy will obtain comparable resolution to that commonly available in the other regions of the spectrum. Detection of very faint point sources ($\sim$10 times fainter than ROSAT) becomes possible because of the reduced background per beam, and accurate locations of sources facilitate their identification with counterparts at other

[1]AXAF Science Center, 60 Garden Street, Cambridge, MA 02138
[2]University of New Hampshire, Durham, NH 03824
[3]University of Chicago, IL 60637
[4]Department of Astronomy & Astrophysics, Penn State University, University Park PA 16802

wavelengths. Moreover one of the imaging detectors records a spectrum in each spatial pixel allowing spatially resolved spectroscopy. Transmission gratings provide wavelength resolution which improves linearly with the reductions in a beam size, so that high resolution spectroscopy (E/ΔE ∼ several hundred), particularly at low energies, becomes feasible. However, these instrumental advances will generate new computational and conceptual challenges for X-ray data analysis. The statistical methodology traditionally used in X-ray astronomy may not prove adequate for AXAF.

Table 1 compares the basic characteristics of AXAF and the two X-ray astronomy satellites now operating: the German Röntgen Satellite (ROSAT) and the Japanese Advanced Satellite for Cosmology and Astrophysics (ASCA). Both carry US instrumentation. ROSAT has high spatial resolution and low spectral resolution, while the reverse applies to ASCA. AXAF will outperform these satellites in both respects although its field of view is more limited. Though the X-ray data is transmitted by the satellite to ground stations in a linear telemetry stream, each observation can be considered to be a four-dimensional multivariate database where each photon is characterized by its position in the detector (representing two- dimensional location in the sky or wavelength along a grating spectrum); its energy in units of kilo-electron Volts (keV); and its arrival time. Different analysis problems can thus be viewed as challenges in image restoration, interpretation of spectra, and time series analysis.

TABLE 1.1. Instrument characteristic

Instrument	Mirror PSF[a] [arcsec]	ΔE[b] [keV]
ROSAT PSPC	5	0.4
ASCA SIS	150	0.1
AXAF ACIS	0.5	0.1
AXAF ACIS/HETG	0.5 (1-D)	0.005

[a] Width of the Point Spread Function; [b] Energy resolution at 1 keV.

2 Example of X-ray Data Analysis Problems

2.1 Spectral Analysis of ROSAT, ASCA and AXAF data

A quasar spectrum, or plot of X-ray photon flux F_x against energy E, observed with ROSAT PSPC (Position Sensitive Proportional Counter) is shown in Figure 1.

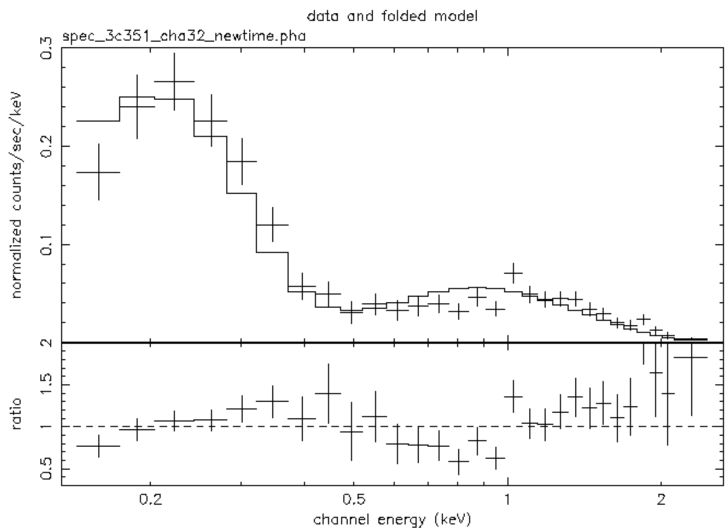

FIGURE 1. Upper panel shows the best fit power law emission model to the ROSAT PSPC quasar spectrum (3C 351). Lower panel shows the residuals. Large deviations from the model indicate more complicated structure present in the data. Thanks to Fabrizio Nicastro.

Detection of the X–ray photons is an intrinsically Poisson process. When binned, and with high enough source counts per bin, a Gauss–Normal distribution is a reasonable approximation. Further, a "smearing" of the true energy and angular position of each photon by an instrument response function is fundamental to the X–ray measurement process. The wider the "smearing", the lower the spatial or energy resolution. The ROSAT spectrum contains just a few independent energy channels (3-4) binned on a finer scale (32 bins) between 0.1 and 2.5 keV; this is a low resolution spectrum with resolution $\Delta E \simeq 0.5$ keV). In a simple analysis of this source, the model spectrum of a power law ($F_x \propto E^{-\alpha}$) with the galactic absorption (a complicated nonlinear function) has been assumed. The χ^2 statistic has been used to find the best-fit model parameters: spectral index α of the power law emission and a column density of the absorber. Note that the overall shape of the spectrum is not a simple power law because the nonlinear spectral reflectivity of the focussing mirrors and of the detector have been included in the model. Additional features in the spectrum are identified by comparing the model prediction to the observed data and searching through the residuals (Figure 1, lower panel). For example, absorption edges from ionized oxygen are present in this spectrum around 0.6-0.8 keV, but this low resolution spectrum does not allow us to distin-

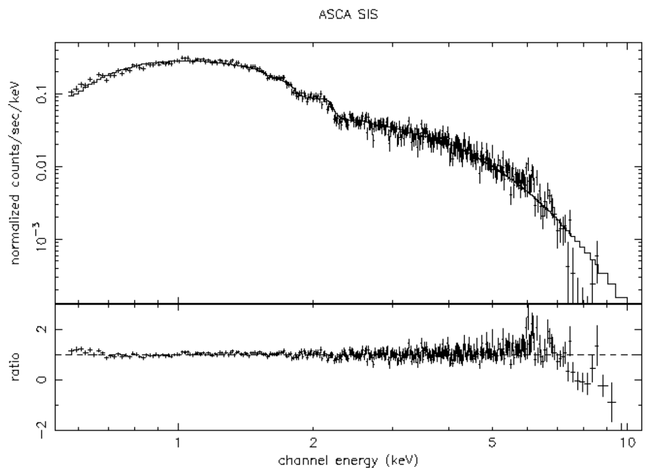

FIGURE 2. Upper panel shows the best fit power law model to the ASCA SIS spectrum of the Seyfert 1 galaxy (NGC 7469, observed in 1993). The complicated structure (emission lines and edges) above 5 keV is clearly present in this data.

guish between different ionization states (O VII and O VIII) or to decide whether both edges or only one are present.

A higher resolution quasar spectrum obtained with the ASCA CCD detectors (the "SIS", Solid-state Imaging Spectrometer) is shown in Figure 2. Compared to the ROSAT PSPC detection, the SIS spectrum contains more independent channels over a wider range of energies (0.5-10 keV) and provides higher spectral resolution ($\Delta E \sim$0.1 keV). The data are binned into 256 energy bins. More features, emission lines and edges, can be found in such a spectrum, such as the likely iron line complex around 6-7 keV. These data were analyzed in exactly the same way as in the ROSAT PSPC example. First a power law emission model is assumed and plot of the residuals to this fit is made. Strong deviations are identified. The significance of the additional features is estimated by comparing the χ^2 values of different models (e.g. power law, power law + emission lines, blackbody emission, plasma emission). This spectrum is similar to these we will get from each pixel of ACIS, the CCD spectral imager on AXAF.

AXAF spectra at much higher resolutions can be obtained with the help of grating elements. An example of a simulated High Energy Transmission Grating (HETG) spectrum is presented in Figure 3. The energy covered with this spectrum ranges from 0.4-10 keV similar to the ASCA SIS spec-

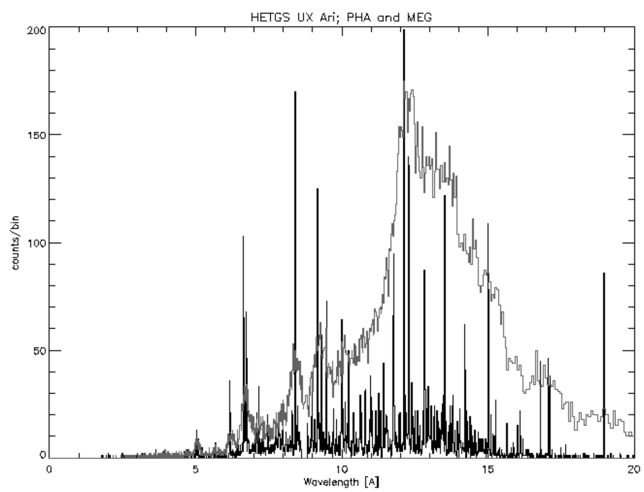

FIGURE 3. Simulated high resolution spectra of a stellar X-ray source (UX Ari) observed with the AXAF ACIS/HETG. The spectra obtained with the AXAF ACIS – dotted line, from zero order photons – and the AXAF/HETG (with medium energy grating of the HETG instrument) – solid line, are plotted on the same scale for the comparison. The assumed source model was a continuous emission measure distribution with a peak near logT=7.0. Courtesy David Huenemoerder.

trum, but with much higher spectral resolution ($\Delta E \sim 0.005$ keV) for a point source. A forest of emission lines is clearly visible in this simulation, though each may be represented by only a few photons. Using the HETG, for the first time, the quality of X-ray spectra will be comparable to those obtained in the optical band. Global statistics like χ^2 are unlikely to be effective in modeling such complex and low- signal (e.g. no longer Gauss–Normal) spectra.

2.2 Spatial Analysis and Imaging with ROSAT and ASCA

The spatial resolution of ROSAT allows us to distinguish individual point sources or extended emission regions. However, a point source always spreads over a finite region of the detector mostly because of non-ideal optics. A "Point Spread Function" (PSF) is used to characterize the way the photons are spread around the central position of a point source. The half power diameter of the AXAF PSF will be less than 0.5 arcsec (Table 1). The pixel size in ACIS is 0.5 arcsec so the PSF is undersampled, although the random jitter of the spacecraft and analysis of 'split events' may allow sub-pixel imaging.

The high imaging resolution allows us to identify and separate sources in the image. The improved energy resolution of ACIS will simultaneously provide spatially resolved spectra of particular regions in the image, which was not possible with previous missions. Combined spectral-spatial analysis, with complicated nonlinear parametric models in both domains, will be important in studying supernova remnants, galaxies and clusters of galaxies, and in general for any extended X-ray source. The high quality of the AXAF X-ray data provides new challenges for the X-ray data analysis. It is clear that the current approaches can not fully exploit the capabilities of the AXAF instruments. Here, we describe a few of the statistical and computational problems that we have so far identified. Some of them appear to be theoretically solvable but computationally challenging, while others state problems for mathematical statistics which, so far as we know, are unsolved. From an astronomical point of view, the problems can be classified as follows: Modeling the Data, Source Detection, and Instrument Related Issues.

3 Modeling the Data

The properties of X-ray detectors, combined with low source and background fluxes and fast read-out times, allow the position (x, y), time (t) and energy (E) of each photon to be recorded. Most traditional methods involve binning (grouping) the data so that Gaussian statistics apply and χ^2 can be calculated for each bin. But this results in loss of spatial, temporal or spectral resolution, and unbinned methods are preferable. Tests based on the empirical distribution functions (Kolmogorov-Smirnov, Cramer-von Mises, Anderson-Darling) are available, but are not readily applicable to multivariate datasets. For univariate data, some astronomers use these tests repeatedly for parameter estimation (Fasano et al. 1993), but the validity of this approach has not been evaluated statistically. Bayesian methods based on the Poisson likelihood in four dimensions are possible, but are not fully developed or easy to apply. In general, we have difficulty envisioning a full data analysis system performed in the unbinned "event space".

Another goal is to develop analysis tools that simultaneously treat spatial, temporal and/or spectral information. The complex temporal variability of some X-ray sources is discussed in the chapters by J. Swank and M. van der Klis in this volume. For joint spatial-timing analysis, Giommi et al. (1995) suggest visualizing the image where the value in each pixel represents the Kolmogorov-Smirnov statistic measuring source variability. Kashyap (1996) suggests a source detection algorithm to be applied in the 3D space. More general statistical tools are needed for modeling the multi-dimensional datasets.

3.1 Searching a large parameter space

Complex global models have to be used to describe the new X-ray data. They are derived from astrophysical theory, can have highly nonlinear forms (e.g. with sharp discontinuities due to atomic absorption and emission lines) and include many parameters to be estimated. The main scientific goal of the investigations is to constrain these parameters with the high-quality AXAF data. When fitting complex models, such as fitting multiple plasma temperatures and elemental abundances to the grating spectrum of a star, or fitting a non-equilibrium ionization model to a spatially resolved supernova remnant, there may be dozens of model parameters. In such a complex parameter space, many 'best fit' solutions with similar "goodness of fit" statistics may be present. How can we efficiently search for the minima in the large parameter space? Is it possible to know when the entire parameter space has been adequately explored? Can the statistical probabilities of distant minima be evaluated? How should parameter confidence intervals be determined in cases where the goodness-of-fit statistic is unusually low or high (e.g. reduced χ^2 far from unity)? How should the confidence intervals be represented when there are many model parameters?

In some cases, a solution may be mathematically "best" or acceptable but have physically unreasonable parameters. Can "unphysicality" be included as a constraint on the fitting process in advance of obtaining solution? Perhaps physical priors can be established within a Bayesian approach. The search through a large parameter space is a serious computational challenge: effective and rapid search algorithms are needed. A Euclidean grid search is not good enough; perhaps Metropolis or Markov Chain Monte Carlo algorithms would be helpful.

3.2 Uncertainties in the Model

The models applied during data analysis may contain some intrinsic uncertainties from the astrophysical theory. For example, many X-ray emission lines do not have fully determined atomic physics to predict their strength, or even their wavelength, and the physicist often can estimate the amplitude of these uncertainties. Modeling of the high resolution spectra would require us to include these atomic physics uncertainties with known variances in the model. Are there techniques to assign uncertainties to the predictions, knowing that each wavelength bin may contain many lines? In the Bayesian approach, uncertainties on the model can sometimes be included directly in the priors. Is it possible to include uncertainties on the model using the frequentist (i.e. maximum likelihood) approach? In either case, how do errors propagate through the calculations when both a model and data contain uncertainties?

In cases where the count rate is sufficiently high (or one is willing to bin the data sufficiently) so that Gaussian probability distribution apply, then one might apply a modified χ^2 statistic where errors in both the data and model are used to weight the variance. A best-fit solution in a least-squares sense could then be obtained. But in general AXAF data will lie in the Poisson regime, and specialized likelihood, semiparametric or Bayesian methods must be developed for this problem.

3.3 Weighting Data by its Information Content

In general a spectrum can be divided into continuum and emission/absorption lines components. In a grating spectrum, most of the counts and most of the bins will be due to continuum emission. The continuum is usually fully determined by just a few parameters (e.g. plasma temperature, density and volume).

Overwhelmingly, most of the interesting physical constraints will be made using the emission lines. Line ratios can provide information about the temperatures or ionization structures of the emitter, and line profiles can be used to study dynamics of the emitting system. However, the lines may contain only 10% of the signal. Some lines contain more information than others; for instance, the existence of some lines or a ratio of certain lines, may determine the density of the emitting gas uniquely. Are there methods for weighting the data by the astrophysical information it carries, rather than simply by its signal-to-noise? Once again, χ^2 is not and adequate statistic.

3.4 Correlated Residuals

X-ray astronomers normally use χ^2 statistic to find a global best fit, and then examine the residuals of the fit (data–model) to find new structures or features (lines or edges) in the spectra. Often these residuals are obviously correlated (Figures 1-2). This usually indicates a localized feature (e.g. an emission line) that cannot drive χ^2. χ^2 though is blind to the clustering of the contributors to the statistic. Nonparametric 'run statistics' might be used, but these do not take into account the known measurement errors for each spectral channel. Are there other statistics available that include this information?

4 Source Detection

4.1 Analysis of events in N-space

Traditionally, X-ray astronomers find sources in their images, where all temporal and spectral information has been ignored. First the density of counts attributable to an "uninteresting" background level is evaluated, and then a window is passed across the field to locate regions where the local photon density is significantly above the background (Marshall 1992). Threshold levels for source existence are set using Poisson probabilities, the likelihood ratio test based on the Poisson distribution (Cash, W. 1979) or by Monte Carlo simulation. The window can be a simple square, or can be a filter matched to the known point spread function of the telescope (Vikhlinin, A. et al., 1995) These procedures are reasonably successful in locating constant point sources with continuous spectra, but suffer inefficiencies for spatially diffuse structure and unusual objects that are discontinuous in spectra or time.

Methods that search for clustering of events in the 4-dimensional position-time-energy space without resorting to binning may be more sensitive than standard techniques requiring binning, and may be sensitive to different types of X-ray sources (e.g. bursts, emission-line only sources). Such methods operate directly on the event files, so the data are not manipulated and all the original information remains there during the detection process. Percolation methods such as the "friends-of-friends" algorithm (known in statistics as single linkage hierarchical clustering) are commonly used to locate galaxy clusters in in studies of large-scale structure in the Universe (see Feigelson & Babu 1992). Recently percolation has been used in X-ray source detection algorithms (Ebeling et al. 1996). The main difficulty is to evaluate the statistical significance of detected structures and to reliably distinguish real physical structures from statistical noise in regions of diffuse low surface brightness.

4.2 Multi–scale Analysis of Complex Source Structures

Wavelets can be used in the source detection process as well, as described by Bijaoui; Damiani et al., Kashyap et al. in these proceedings. Binned images are correlated with wavelet functions at various scales and the resulting coefficients are compared across the scales in order to determine source parameters. Methods to extend the use of wavelets to entire fields-of-view and beyond the detection of point sources are under development (see the aforementioned contributions). Whereas methods to detect point sources are in good standing, much work is left to be done with source characterization: what is the significance of a multi-bin source (and what defines its extend and shape); and how do we combine information over multiple scales and statistically characterize the results?

Consider an ACIS image subject to a wavelet transform. Wavelet coefficients below some amplitude and/or spatial thresholds are deleted, and the image is reconstructed from the remaining coefficients. What is the statistical error of structures in the transformed image? If an extended source is present, how sensitive is its shape (size, eccentricity, etc.) to the manner of the reconstruction? And more general question how can we put the confidence limits on the source shapes?

In the Bayesian context, progress has been made recently on the development of "pixons" (Puetter 1996). The data are described by a model which is smoothed locally at the best scale for a given structure. Parts of the image with less structure (e.g. source-free regions) are sufficiently well represented at large scales, while parts of the image with more structure (e.g. many point sources in a small region) need smaller scales. As in the wavelet methods the main problems are related to characterization of detected sources.

5 Instrument Related Issues

5.1 Photon Pile-up in ACIS

The time between readouts of the AXAF ACIS is 2.7 seconds. If in that time two photons land in the same pixel, the electric charge created by the two will be summed. The existence of two photons and their individual energies will be unknown. This problem is called photon 'pile-up'. Since 50% of the photons from a point source in AXAF fall into a single pixel this will be a common occurrence for bright X-ray sources.

As Poisson statistics apply to photon arrival times (assuming a constant source), it is easy to calculate when pile-up will set in. Pile-up becomes a 10% effect at $\sim$0.1 ct/s (Figure 4). It will not be clear which photons are doubly counted. The long tail of the Poisson distribution also means that triple and quadruple countings will be significant too. If the total charge exceeds that from a $\sim$15 keV photon then the total charge collected will exceed the capacity of the telemetry, and the single "overscale" event is considered a background event (due to a particle not a photon), and is lost. Worse still, real events can be 'split' over two or more pixels, so if two counts arrive next to each other they will be considered as a single event in normal processing. This lowers the pile-up count rate limit by almost a factor 10, so it will be very common. Figure 5 shows a simulated high count rate spectrum with pile-up and the corresponding clean spectrum with no pile-up. It is clear that many soft energy photons were redistributed into the high energy band changing the slope of the spectrum significantly and smearing out the structure at high energies. Are there ways to recover the initial spectrum given an estimate of the fraction of pile-up events, and possibly given a clean spectrum involving about 10% of the events?

5.2 Overlapping orders in Low Energy Grating data

When a grating spectrum is projected onto the AXAF detectors, the different orders of diffraction overlap in space. With the ACIS CCD detector this is a minor problem since the orders separate quite cleanly in pulse height

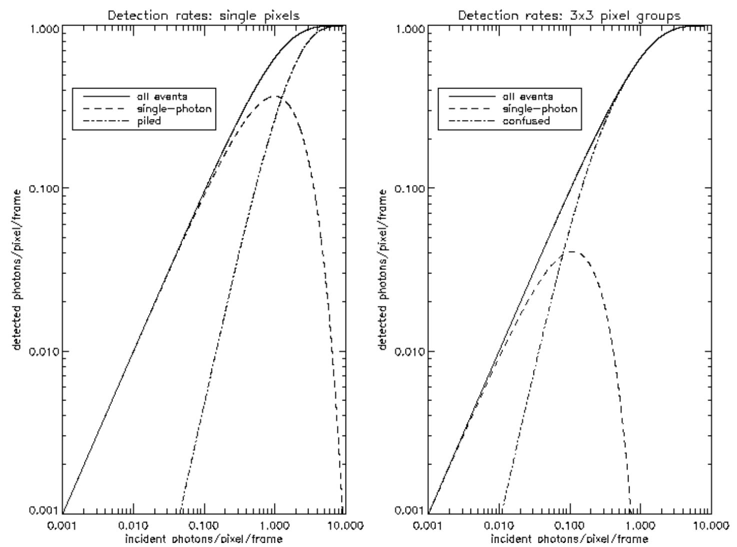

FIGURE 4. Detected event rate vs. incident photon rate for single-pixel case, left, and (more realistic) 3×3 pixel case, right. For the single-pixel case, the likelihood of multiple-photon ("piled") events becomes significant ($\sim$ 10%) at incident rates of only ~ 0.2 counts per frame. In the 3×3 pixel case, photon confusion (photons arriving in neighboring pixels during a single frame) lowers this count rate threshold by an order of magnitude. Courtesy Joel Kastner.

space. However, when the other major instrument, the High Resolution Camera ("HRC") is used, as it must be for low energy spectra, there is no direct way to discriminate the different orders, since the HRC has almost no inherent energy resolution (Figures 6a,6b).

An iterative deconvolution method may work for continuum points, while a pattern matching technique may be effective for lines, at least if they are not so numerous that they are heavily blended with one another. Higher orders have higher wavelength resolution, which would separate blended lines, so changing the pattern. The higher orders can dominate the counts, especially in spectral lines, so the problem is not a perturbative one, and error propagation and 'blow-up' is a concern. Can alternative techniques be considered?

Acknowledgments: We would like to thank Prof. Ernst Linder, Herald Ebeling, Frank Primini, Joel Kastner, David Huenemoerder, Fabrizio Nicastro for discussions. AS and ME are supported by NASA grant NAS8-39073. AC is supported by the COMPTEL project on board CGRO, which is supported in part through NASA grant NAS 5-26646, DARA grant 50 QV

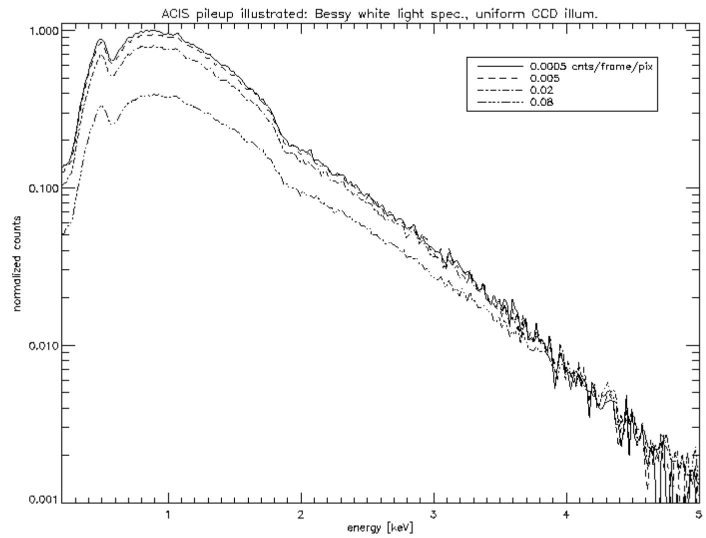

FIGURE 5. Modifications to an input continuum spectrum caused by photon pileup in the detector. Pileup results in false high-energy events, effectively flattening the spectrum. The effect grows more prominent with source strength. Based on simulated ACIS data produced by Andy Rasmussen.

90968, and the Netherlands Organization for Scientific Research (NWO). EDF is supported by NASA grants: NAGW-2120 and NAS8-38252. PF and VK acknowledge support from AXAF Sience Center.

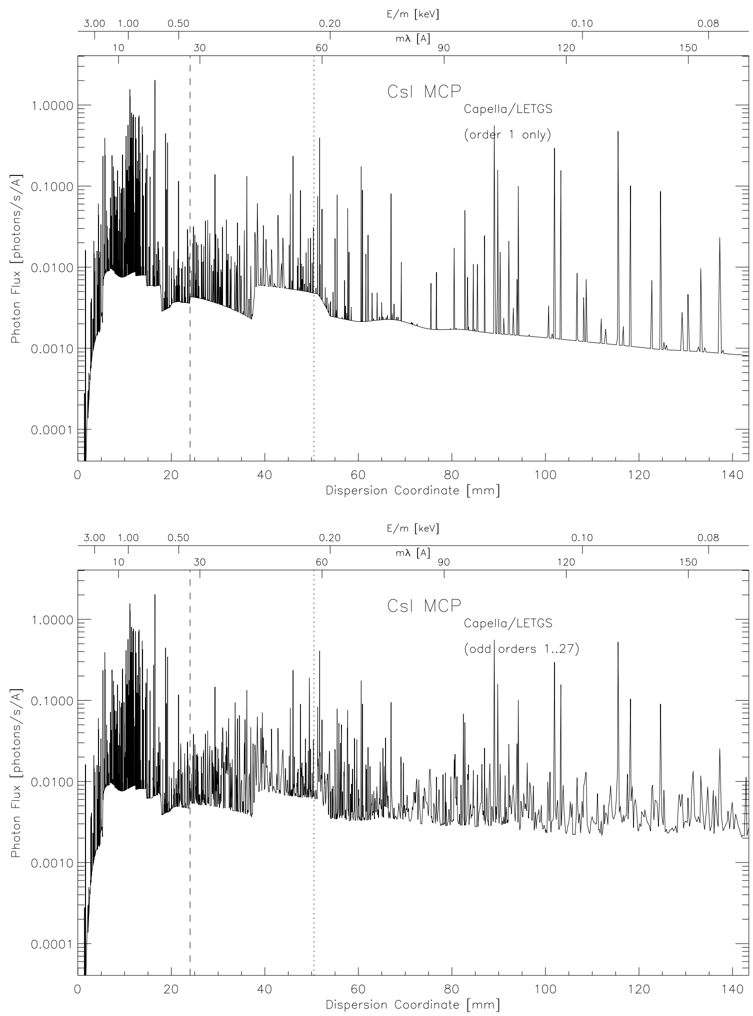

FIGURE 6. Simulated high resolution spectra of a star (Capella) observed with AXAF HRC-S/LETG. It is a grating spectrum with many diffraction orders. a) Only the first order spectrum is plotted. b) The contributions from the odd orders (1–27) are added and the resulted spectrum is plotted (from Internal ASC Memorandum by David Huenemoerder, November 1994). Courtesy David Huenemoerder.

6 References

[1] Bijaoui, A. 1996, these proceedings

[2] Bradt et al., 1992, ARAA 30, 391

[3] Cash, W. 1979, Ap.J. 228, 939

[4] Damiani, F. et al. 1996, these proceedings

[5] Ebeling et al. 1996, MNRAS, in press

[6] Feigelson, E.D. & Babu, G.J., 1992, "Statistical Challenges in Modern Astronomy", Springer-Verlag, New York, 1992.

[7] Fasano et al. 1993, ApJ, 416, 546

[8] Giommi et al. 1995, ADASS IV, Ed. Shaw, R.A., Payne, H.E. & Hayes, J.J.E. p.117

[9] Huenemoerder, D., 1994, Internal ASC Memorandum

[10] Kashyap, V. 1996, ADASS V Ed. Jacoby, G.H. & Barnes, J., p.25

[11] Kashyap, V. et al. 1996, these proceedings

[12] Marshall, H., 1992, in "Statistical Challenges in Modern Astronomy", Ed. Feigelson, E.D., & Babu, G.J., p. 247.

[13] Puetter 1996, in XV Intl. Workshop on Maximum Entropy and Bayesian Methods, ed. Hanson and Silver, in press

[14] Swank, J., 1996, these proceedings

[15] van der Klis, M., 1996, these proceedings

[16] Vikhlinin, A., Forman, W. Jones, C. & Murray, S, 1995, Ap.J. 451, 542

[17] Zombeck, M. 1996, "Advanced X-ray Astrophysics Facility" to be published in the Proceedings of the International School of Space Science Course on "X-ray Astronomy" L'Aquila, Italy August 29 - September 10, 1994

[18] Zombeck, M.V, 1982, COSPAR and International Astronomical Union, Symposium on Advanced Space Instrumentation in Astronomy, 4th, Ottawa, Canada, May 20-22, 1982. Advances in Space Research, vol. 2, no. 4, p. 259-270.

[19] Zombeck, M.V., 1979, in: Space optics: Imaging X-ray optics workshop; Proceedings of the Seminar, Huntsville, Ala., May 22-24, 1979. Bellingham, Wash., Society of Photo-Optical Instrumentation Engineers p. 50; Discussion, p. 62.

Discussion by Joseph Horowitz[5]

There are many statistical questions touched on in the paper of Siemiginowska et al. (referred to briefly as [S+]), most of which are either implicit or very general. Thus, some of my comments will also be implicit or very general, consisting, in some cases, only of pointing out some relevant references.

Chi-square

The notion of χ^2 comes up often in the astronomy literature, and in [S+] as well. Now χ^2 has several possible meanings to statisticians, and Eric Feigelson was kind enough to clue me in to what astronomers mean by "χ^2". (I am not the only one who has been confused on this point; see [B], p.322.)

According to him, the "bible" of statistics for astronomers is ... *Numerical Recipes* [NR]! In ch.15 of [NR], χ^2 refers to the (scaled) error sum of squares in nonlinear regression, which, under certain conditions, has a χ^2-distribution with the appropriate degrees of freedom.

In more familiar terms (for statisticians), the *Pearson* χ^2*-statistic*, designed to test hypotheses about Poisson (or multinomial) observations,

$$\chi_P^2 = \sum_{i=1}^{k} (N_i - \lambda_i(\theta))^2 / \lambda_i(\theta),$$

is often used in astronomy.

Here $N_1, \ldots, N_k$ are independent Poisson counts, with means $\lambda_i(\theta)$, typically representing photon counts in k distinct energy or spatial bins, and the parameter θ is some physical characteristic of, for example, an x-ray source.

When the means are large, the N_i have approximately normal distributions, but then the variances are necessarily equal to the means (also pointed out in [B]). Then the nonlinear regression material in [NR] becomes relevant, but the constraint on the variances is often ignored.

The usual strategy is to fit a model by finding the value $\hat{\theta}$ that minimizes χ_P^2, and to search the residuals for further structure. The χ_P^2-statistic has a good intuitive motivation, and its asymptotic distribution is known. The minimizer $\hat{\theta}$ is asymptotically equivalent, and in some cases identical to, the maximum likelihood (ML) estimator of θ. For details, see [C], chs. 30, 33.

[5] Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003

Another version of χ^2 is the *likelihood ratio* - χ^2,

$$\chi^2_{LR} = 2\sum_{i=1}^{k}(N_i \log(N_i/\lambda_i(\theta)) - (N_i - \lambda_i(\theta))),$$

which is asymptotically equivalent to χ^2_P and has the same limiting distribution. But χ^2_{LR} is the correct likelihood ratio statistic for testing the "null" hypothesis, H_0, that the Poisson means are given by the model $\lambda_1(\theta), \ldots, \lambda_k(\theta)$, for some θ, against the alternative, that the means are not of that form. Tests of H_0 against specific alternative models allow rigorous assessment of the reality of features that do not conform to the H_0-model.

In this connection, χ^2_{LR} has good decomposition properties, not shared by χ^2_P, for certain sequences of nested models of successively greater complexity (see [MN]).

The "Bible"

Although it is a great read, [NR] is no more suitable as a statistical bible than Ptolemy is for astronomy. A glance at the main statistical references in [NR], ch.15, confirms this: Bevington 1969, von Mises 1964, Brownlee 1965, Martin 1971, plus two 1976 papers in *Ap. J.*. It is as though nothing had happened in statistics over the last 25 years or so. For more recent, though not necessarily astronomer- (or statistician-) friendly, expositions of nonlinear regression, see [BW], [SW], [G1]. Some recent *linear* regression books, for instance, [M], [R], also contain some nonlinear material.

Searching the Parameter Space

The optimization problems hinted at in [S+], §3.1, are standard, difficult ones in numerical analysis. Some statistical theory that could be applied to grid searches for ML and similar estimators is available (e.g., [WS] and references therein) to the intrepid. The added feature is that accuracy of the grid estimator *with high probability* can sometimes be asserted.

Recently, effective Markov Chain Monte Carlo (MCMC) calculations have been developed for ML and other function optimization problems; see [G2] and [B+].

The question of "unphysicality" ([S+], §3.1) is of course not a statistical one. If it can be expressed in a reasonable mathematical form, physicality can, in principle, be added as a constraint to the model.

Which Residuals?

Residuals are mentioned several times in [S+] (§§2.1, 3.4). In dealing with low counts, where the Poisson distribution must be respected, there are various types of residuals specifically tailored for Poisson data, viz., Anscombe/Cox-Snell and deviance residuals, all discussed in [MN].

To Bin or Not to Bin

Many of the phenomena discussed in [S+] can be modeled directly, with no binning.

Let B be the energy $\times$ time "box", $E_1 \leq E \leq E_2$, $t_1 \leq t \leq t_2$. Astrophysical models for the photon count $N(B)$ in B often specify that, as a random variable, $N(B)$ have a Poisson distribution with mean $(t_2 - t_1) \int_{E_1}^{E_2} g(E)\, dE$, and that, for disjoint energy $\times$ time regions, the counts be statistically independent. The physics is contained in the function $g(E)$. This type of model is a *space-time Poisson point process*, although "energy-time" would be a better term for this example.

If a photon of energy E is detected in the energy interval $E' \pm dE'$ with probability $k(E, E')dE'$, where $k(E, E')$ models the detector, then (theorem) $N'(B)$, the number of detections in the box B, also follows a Poisson point process model with g-function $g'(E') = \int k(E, E')\, g(E)\, dE$. The observations are the counts $N'(B)$, for all boxes B, which is equivalent to the full, unbinned data set. The inference problem is to find out information about the *original* $g(E)$.

There is an elaborate statistical literature on modeling and inference for Poisson point processes, some of which is cited elsewhere in these proceedings. For astronomers, a good source for this point is [SM].

Conclusion

Reading between the lines of [S+], many of the statistical questions sound really fascinating, but it is usually not possible to say whether there are statistical techniques for this or that purpose without knowing the details of the problem. Choosing a statistical technique is not like choosing a pair of shoes off the shelf, especially for such complex phenomena as those discussed in [S+]. Rather than "statistician as shoe clerk", a more appropriate metaphor might be "statistician as psychotherapist". Serious collaboration between astronomers and statisticians requires lots of conversation, and should start early in the project. Recent technological advances are generating fundamental new statistical and scientific challenges that would be best met by such collaborative efforts.

7 References

[BW] D.M. Bates and D.G. Watts, *Nonlinear Regression and its Applications.* Wiley 1988.

[B+] J. Besag, P. Green, D. Higdon, and K. Mengersen, Bayesian computation and stochastic systems. *Stat. Sci.*, 10, 1995, 3-66.

[B] P. Bickel, in *Statistical Challenges in Modern Astronomy*, E. Feigelson and G. Babu (eds.). Springer 1992.

[C] H. Cramèr, *Mathematical Methods of Statistics.* Princeton Univ. Press 1946.

[G1] A.R. Gallant, *Nonlinear Statistical Models.* Wiley 1987.

[G2] C. Geyer, On the convergence of Monte Carlo maximum likelihood calculations. *J. Roy. Statist. Soc. Ser. B* 56, 1994, 261-274.

[MN] P. McCullagh and J.A. Nelder, *Generalized Linear Models* (2nd ed.). Chapman and Hall 1989.

[M] R.H. Meyers, *Classical and Modern Regression with Applications* (2nd ed.). PWS-Kent 1990.

[NR] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes* (2nd ed.). Cambr. Univ. Press 1992.

[R] J.O. Rawlings, *Applied Regression Analysis: A Research Tool.* Wadsworth 1988.

[SW] G.A.F. Seber and C.J. Wild, *Nonlinear Regression.* Wiley 1989.

[SM] D.L. Snyder and M.I. Miller, *Random Point Processes in Space and Time* (2nd ed.). Springer 1991.

[WS] X. Shen and W.H. Wong, Convergence rate of sieve estimates. *Ann. Stat.* 22, 1994, 580-615.