Astronomical Data Analysis Software and Systems IV
ASP Conference Series, Vol. 77, 1995
R. A. Shaw, H. E. Payne, and J. J. E. Hayes, eds.

Bias-Free Parameter Estimation with Few Counts, by Iterative Chi-Squared Minimization

K. Kearns, F. Primini, and D. Alexander

Smithsonian Astrophysical Observatory, 60 Garden St., Cambridge, MA 02138

Abstract. We present a modified χ^2 fitting technique, useful for fitting models to binned data with few counts per bin. We demonstrate through numerical simulations that model parameters estimated with our technique are essentially bias-free, even when the average number of counts per bin is $\sim$1. This is in contrast to the results from traditional χ^2 techniques, which exhibit significant biases in such cases (see, for example, Nousek & Shue 1989; Cash 1979). Moreover, our technique can explicitly handle bins with 0 counts, obviating the need to ignore such bins or rebin the data. We conclude with a discussion of the problem of estimating goodness-of-fit in the limit of few counts using our modified χ^2 statistic.

1. Introduction

When fitting models to data with few counts, two of the most common methods used are the standard χ^2 method and the C statistic. Use of the χ^2 method requires that one avoid bins with 0 counts by either ignoring them or rebinning, and produces significantly biased results for data with few counts. The C-statistic gives unbiased results but is difficult to interpret in terms of goodness-of-fit. Neither approach is ideal, though each is useful in some cases. The Iterative Weighting Technique which we investigate here both addresses the deficiencies inherent in using the standard method for data with few counts, and provides a goodness-of-fit parameter which is indistinguishable from the standard χ^2 parameter for many datasets.

2. Iterative Weighting

Iterative Weighting (IW) is an example of the class of weighted least-squares estimators described by Wheaton et al. (1994), in which χ^2 is expressed as a weighted sum of squared deviations,

$$\chi^2 \;=\; \sum_i W_i[O_i \;-\; M_i(p_1, p_2 ...)]^2,$$

where O_i are the observed counts in bin i, $M_i(p_1, p_2...)$ are the counts predicted by the model M with parameters $(p_1, p_2...)$, and the weights W_i are the inverses of the true variances σ_i^2. As Wheaton et al. (1994) point out, the approximation

$W_i \simeq O_i^{-1}$ leads to significant biases in the best-fit parameters, due to the strong anti-correlation between W_i and $[O_i - M_i]^2$. Similar biases are encountered if the approximation $W_i \simeq M_i^{-1}$ is used (Nousek & Shue 1989). The IW technique avoids such biases by estimating W_i through successive iterations, where for each iteration, j, $W_i^j \simeq [M_i(p_1^{*,\, j-1}, p_2^{*,\, j-1}...)]^{-1}$, and the best-fit parameters $p_1^{*,\, j}, p_2^{*,\, j}...$ are determined by minimization of

$$\chi^{2,\, j} \;=\; \sum_i \frac{[O_i \;-\; M_i(p_1^j, p_2^j...)]^2}{M_i(p_1^{*,\, j-1}, p_2^{*,\, j-1}...)}.$$

For the first iteration, all weights are set to 1. In our sample, we find that the minimum χ^2 values and best-fit parameters converge after about 6 iterations.

3. Data Simulation

To demonstrate the IW technique, we repeat the simple numerical experiment of Nousek & Shue (1989). For a range of total counts, N, from 25 to 1000, we generate an ideal power-law spectrum such that:

$$\overline{n_i} = N_o \int_{E_1+(i-1)\Delta E}^{E_1+i\Delta E} E^{-\gamma} dE, \quad N = N_o \int_{0.095}^{0.845} E^{-\gamma} dE$$

for $i = 1 \rightarrow 15$, $\Delta E = (0.845 - 0.095)/15$, and $\gamma = 2.0$ For each ideal spectrum, we simulate 1000 sample spectra $\{\mathrm{n}_i\}$, where $\{\mathrm{n}_i\}$ are random deviates drawn from Poisson distributions with means = $\overline{n_i}$. We then determine best-fit model parameters N_o^{calc} and γ^{calc} for each simulated spectrum, using IW and Powell's method for function minimization (Press 1988). For each N, we then compute the average N_o^{calc}/N_o and γ^{calc}/γ; compile the distributions of minimum χ^2 for comparison with the theoretical distribution; and compute the percentage of simulations for which the χ^2_{min} and $\Delta\chi^2$ contours include N_o and γ, for comparison with the expected percentages.

4. Results

In Table 1 we compare the biases (as measured by the ratios of average best-fit parameter values to true values) in 1000 IW runs with those found for traditional χ^2 and the C statistic by Nousek & Shue in 250 runs. We find that the IW biases are comparable to those encountered using the C statistic for all N. These results are displayed in Figure 1. In Figure 2 we compare both differential and cumulative theoretical χ^2 distributions with our observed distributions. We apply a KS-test to the cumulative distributions and find that at N=25 the match is poor, but by N=100 the two distributions are in good agreement.

The percentage of simulations for which the $\chi^2_{min} + \Delta\chi^2$ contours include N_o and γ, for $\Delta\chi^2$ values appropriate to various joint two-parameter confidence levels, is shown in table 2. For most N, the measured and expected confidence levels are in good agreement.

	Traditional χ^2			Iterative Weighting			C-Statistic		
N	N_{calc}/N_o	γ_{calc}/γ	%cnvg	N_{calc}/N_o	γ_{calc}/γ	%cnvg	N_{calc}/N_o	γ_{calc}/γ	%cnvg
25	0.709	1.152	96	1.145	1.003	98	1.269	0.958	86
50	0.647	1.134	100	1.055	1.008	99.6	1.079	0.998	100
75	0.636	1.130	100	1.025	1.009	100	1.078	0.995	100
100	0.673	1.109	100	1.008	1.008	100	1.053	0.996	100
150	0.707	1.094	100	1.025	1.001	100	1.015	1.005	100
250	0.767	1.072	100	1.019	1.000	100	1.019	1.000	100
500	0.863	1.040	100	1.007	1.000	100	0.997	1.004	100
1000	0.937	1.017	100	1.005	1.000	100	1.001	0.999	100

Table 1. Comparison of three fitting techniques.

	Minimum χ^2 plus ... (for 2 parameters of interest)					
N	2.30 (68.3%)	4.61 (90%)	6.17 (95.4%)	9.21 (99%)	11.80 (99.73%)	18.40 (99.99%)
25	69.8	87.0	92.1	96.8	98.4	99.9
50	68.9	88.1	93.7	97.2	98.2	98.3
75	67.7	87.8	93.5	98.4	99.3	99.8
100	68.1	89.1	94.1	98.2	99.0	99.8
150	67.0	87.1	93.6	97.7	99.3	100
250	68.0	90.3	95.6	99.2	99.8	100
500	69.1	90.1	95.5	98.9	99.6	99.9
1000	69.6	88.9	95.0	99.0	99.7	100

Table 2. Estimating Confidence Limits: Percentage of Best Fits Within Various χ^2 Boundaries, from a total of 1000 spectral fits.

5. Conclusions

We find that unbiased parameter estimates by χ^2 minimization are possible for binned data with few or no counts in some bins, provided the χ^2 calculation is modified slightly. Except for very small N, this modified χ^2 statistic is distributed according to the theoretical χ^2 distribution. Goodness-of-fit can therefore be assessed using traditional techniques. Further, this χ^2 statistic can be used to estimate confidence levels from standard $\chi^2_{min} + \Delta\chi^2$ boundaries.

Acknowledgments. This work is partially supported by NASA contract NAS5-30934.

References

Cash, W. 1979 ApJ, 228, 939

Nousek, J. A., & Shue, D. R. 1989, ApJ, 342, 1207

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1986, Numerical Recipes (New York, Cambridge University Press)

Wheaton, W. A. et al. 1995, ApJ, submitted

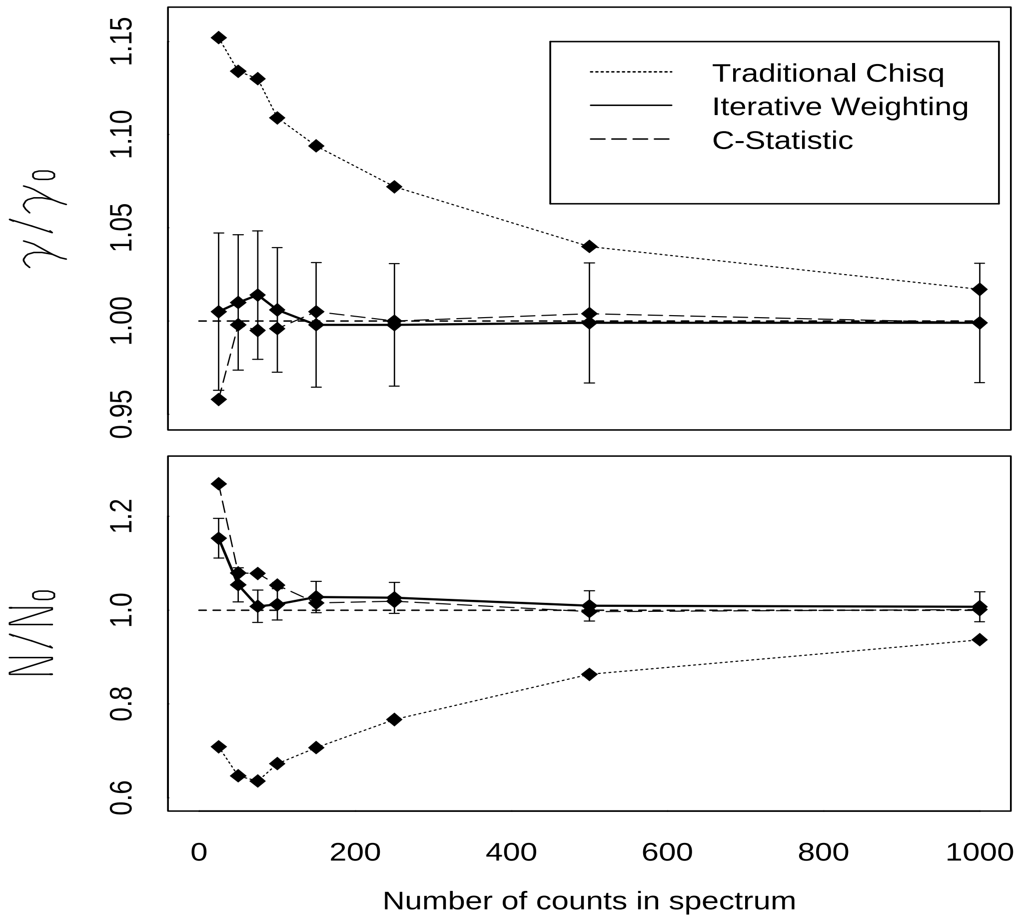

Figure 1. Bias in best-fit parameters for three fitting techniques.

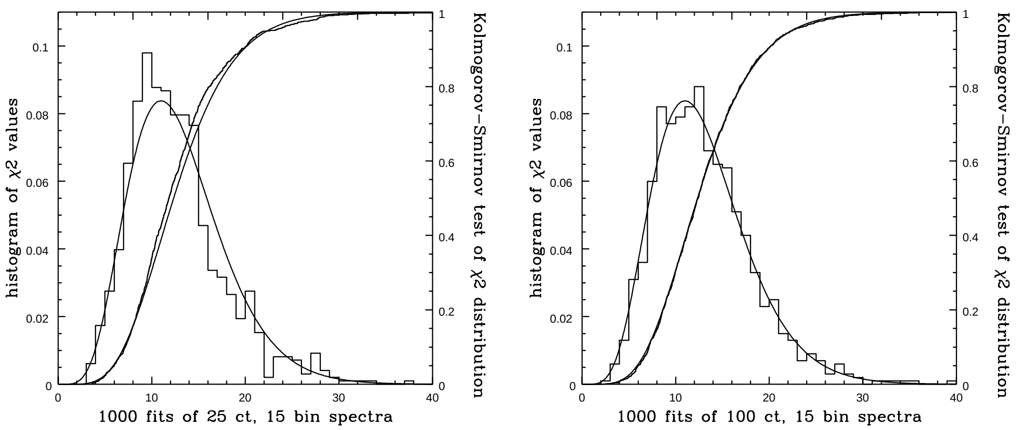

Figure 2. Comparison of theoretical χ^2 distribution with observed distribution for IW by KS-test, with overlaid histograms.