How to handle calibration uncertainties in high-energy astrophysics

Vinay Kashyap
How to handle calibration uncertainties in high-energy astrophysics

Vinay Kashyap

CHASC Astrostatistical Collaboration
Chandra X-ray Center
Smithsonian Astrophysical Observatory

Hyunsook Lee, Aneta Siemiginowska, Jonathan McDowell, Arnold Rots, Jeremy Drake, Pete Ratzlaff, Andreas Zezas, Doug Burke, Rima Izem, Alanna Connors, David van Dyk, Taeyoung Park
bottom line

there is now a way
to include calibration uncertainty
in astrophysical data analysis
in a flexible way
for any instrument, mission, or detector.
Part I
 calibration uncertainty
 in data analysis

Part II
 practical issues of
 storage, retrieval, flexibility
The three most important effects that affect data analysis

- astrophysical model uncertainty
- statistical uncertainty (measurement error)
- calibration uncertainty (systematic error)
examples of calibration uncertainty
power-law residuals with current calibration
power-law residuals with 20Å contamination overlayer
general model of HEA data

\[ M(E', p', t) = \int dE dp \, S(E, p, t; \theta) \, A(E, p'; p, t) \, R(E, E', p'; t) \, P(p, p', E; t) \]

\((E, p, t)\) : photon energy, location, arrival time
\((E', p')\) : detector channel, chip location
\(S\) : astrophysical source model
\(R\) : energy redistribution function (RMF)
\(P\) : position redistribution function (PSF)
\(A\) : effective area (ARF)
\(M\) : predicted model counts
effect on model parameter uncertainty
but how exactly?
DATA  \[\rightarrow\]  CALIBRATION

- Draw parameters
- Compute likelihood
- Update parameters
DATA

Draw effective areas

Draw parameters

Compute likelihood

Update parameters

CALIBRATION
error bars on power-law indices of radio loud quasars
but where do the effective areas come from?
\[ A'(E_j) = A_0(E_j) + \delta A(E_j) + \sum_{k=1}^{N_{max}} r_k e_k \nu_k(E_j) + r(N_{max}+1) \xi(E_j) \]

\[ A = A_0 + \text{bias} + \text{components} + \text{residual} \]
Part II
practical issues of
storage, retrieval, flexibility

\[ A = A_0 + \text{bias} + \text{components} + \text{residual} \]
A = A_0 + bias + components + residual

store in same format as A_0

e.g., SPECRESP

case specific secondary FITS extension

e.g., PCA1D
SIMS
POLY1D
PCPC
MULTISCALE
### TABLE SPECRESP

<table>
<thead>
<tr>
<th>ENERG_LO</th>
<th>ENERG_HI</th>
<th>SPECRESP</th>
<th>BIAS</th>
<th>BIN_LO</th>
<th>BIN_HI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Units</td>
<td>Types</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>keV</td>
<td>float</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>keV</td>
<td>float</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.22</td>
<td>0.23</td>
<td>98.221</td>
<td>0.337985</td>
<td>53.9066</td>
<td>56.3569</td>
</tr>
<tr>
<td>0.23</td>
<td>0.24</td>
<td>115.87</td>
<td>0.473241</td>
<td>51.6605</td>
<td>53.9066</td>
</tr>
<tr>
<td>0.24</td>
<td>0.25</td>
<td>134.755</td>
<td>0.591328</td>
<td>49.5941</td>
<td>51.6605</td>
</tr>
<tr>
<td>0.25</td>
<td>0.26</td>
<td>156.063</td>
<td>0.689149</td>
<td>47.6866</td>
<td>49.5941</td>
</tr>
<tr>
<td>0.26</td>
<td>0.27</td>
<td>175.412</td>
<td>0.726106</td>
<td>45.9204</td>
<td>47.6866</td>
</tr>
<tr>
<td>0.27</td>
<td>0.28</td>
<td>196.854</td>
<td>0.706191</td>
<td>44.2804</td>
<td>45.9204</td>
</tr>
<tr>
<td>0.28</td>
<td>0.29</td>
<td>57.5639</td>
<td>0.0294997</td>
<td>42.7535</td>
<td>44.2804</td>
</tr>
</tbody>
</table>

View Mode: Read/Write  Displaying rows 1 - 20  (1078 total rows)
### COMPONENT, FVARIANCE, EIGENVAL, EIGENVEC

<table>
<thead>
<tr>
<th>Units</th>
<th>Types</th>
<th>FVARIANCE</th>
<th>EIGENVAL</th>
<th>EIGENVEC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>short</td>
<td>double</td>
<td>float</td>
<td>float</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0.43916979432106</td>
<td>261.972</td>
<td>float[1078]</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.236995741724968</td>
<td>192.446</td>
<td>float[1078]</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>0.107716666471898</td>
<td>129.742</td>
<td>float[1078]</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0.0846366658806801</td>
<td>115.005</td>
<td>float[1078]</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>0.0422603040933609</td>
<td>81.2653</td>
<td>float[1078]</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>0.0225162785500288</td>
<td>59.3181</td>
<td>float[1078]</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>0.0158293452113867</td>
<td>49.736</td>
<td>float[1078]</td>
</tr>
</tbody>
</table>

View Mode: Read/Write
Displaying rows 1 - 20  (20 total rows)
what have we got so far?

- realistic error bars
- implemented in BLoCXS
- 500x speed up in analysis
- Sherpa on the way

- unified file format for use in XSPEC/Sherpa
- 100x drop in storage
- generalize to any instrument
- extendable to model lacunae
- roll your own
there are a number of steps between a calibration scientist saying “the error on the effective area is X% at energy Y” to then have it fold into a spectral model fit and inflate the error bars on the parameters, and we believe that we have connected the dots.
uncertainty in energy response
one more thing..
what’s next?

unified file format implemented in XSPEC/Sherpa
other schemes of dimensionality reduction
RMFs: 2D PCA, within and between PCA
PSFs: multiscale residuals
The Three Poisson Model

\[ n \sim \text{Pois}(c \cdot s + b) \]
\[ y \sim \text{Pois}(t \cdot b) \]
\[ z \sim \text{Pois}(u \cdot \epsilon) \]