Statistical Issues in Instrument Calibrations and Goodness-of-fit in Astrophysics

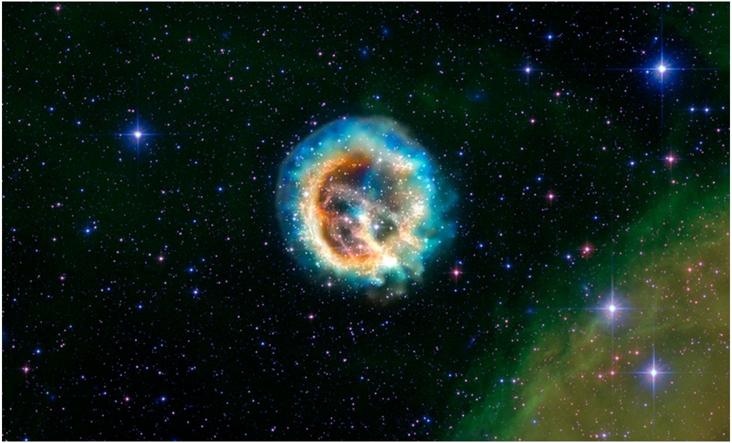
Yang Chen

Joint work with X.-L. Meng (Harvard University), X. Wang (Two Sigma Inc.), D. van Dyk (Imperial College London), V. Kashyap (Center for Astronomy), H. Marshall (MIT), X. Li (USTC), M. Bonamente (UAH)

July 19, 2023

Outline

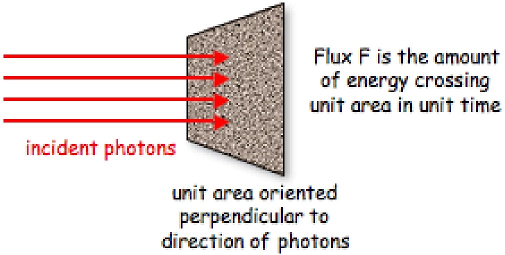
Calibration Concordance Problem (Example: E0102)

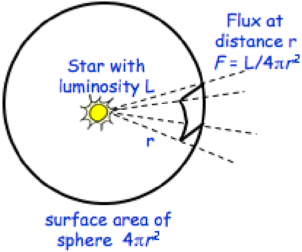


- Supernova remnant E0102
- Four sources correspond to four spectral lines in E0102

Measurements

Flux is the total amount of energy that crosses a unit area per unit time.

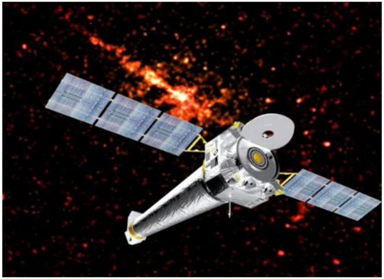




The flux of an astronomical source (F) depends on the luminosity of the object (L) and its distance from the Earth (r), $F = L/4\pi r^2$.

Observatory and Instruments

Current X-ray Observatory



USA: Chandra X-ray Observatory

High angular resolution (~0.5”)

Europe: XMM-Newton

High throughput (large effective area)

And

- Rossi X-ray Timing Explorer
- Swift
- INTEGRAL etc.

9 /72

Observatory and Instruments

CHANDRA
X-RAY OBSERVATORY

CXC Home | Proposer | Archive | Data Analysis | Instruments & Calibration | For the Public

Chandra Instruments and Calibration

The Chandra X-ray Observatory (CXO) is designed for high resolution (≤ 1/2 arcsec) X-ray imaging and spectroscopy. The High Resolution Mirror Assembly (HRMA) focuses X-rays onto one of two instruments, ACIS or HRC. Only one detector (HRC or ACIS) is in the focal plane at any given time. Two grating spectrometers (LETG or HETG) can be placed in the optical path behind the HRMA. The dispersed spectrum is read out by either ACIS or HRC. A high level overview of the instruments on-board the Chandra X-ray Observatory can be found on the About Chandra pages and a more detailed description can be found in the Proposers' Observatory Guide.

Current calibration data products for use in CIAO and other analysis systems can be found in the CALDB pages. A complete listing of all calibration products in the CALDB and a brief description of these products can be found in the Calibration Data Products.

- Calibration Status Summary
- ACIS
- HRC
- HETG
- LETG
- HRMA
- Calibration Database (CALDB)
- Cross-Calibration with other X-Ray Telescopes
- Aspect Information
- Calibration Workshops and Reviews
- SPIE Proceedings
- Science and Calibration Requirements

Advanced CCD Imaging Spectrometer (ACIS)

The ACIS has two arrays of CCDs, one (ACIS-I) optimized for imaging wide fields (16x16 arc minutes) the other (ACIS-S) optimized as a readout for the HETG transmission grating. One chip of the ACIS-S (S3) can also be used for on-axis (8x8 arc minutes) imaging and offers the best energy resolution of the ACIS system.

High Resolution Camera (HRC)

The HRC comprises two micro-channel plate imaging detectors, and offers the highest spatial (<0.5 arc second) and temporal (16 msec) resolutions. The HRC-I has the largest field-of-view (31x31 arc minutes) available on Chandra. The HRC-S is most commonly used to read out the dispersed spectrum from the LETG.

High Energy Transmission Grating (HETG)

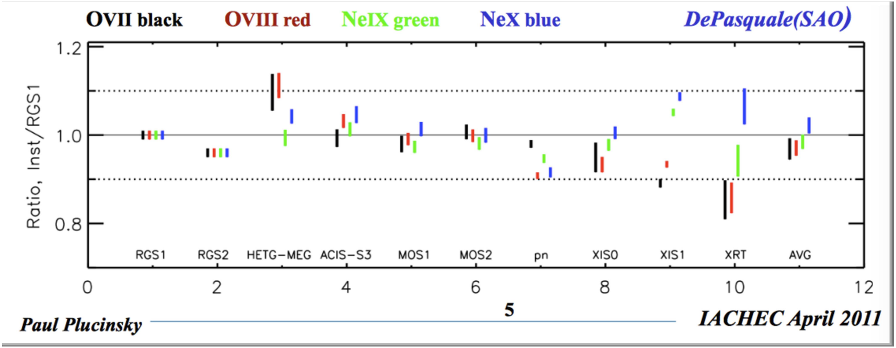
The HETG is optimized for high-resolution spectroscopy of bright sources over the energy band 0.4-10 keV. It is most commonly used with ACIS-S. The resolving power (E/ΔE) varies from ~800 at 1.5 keV to ~200 at 6 keV.

Low Energy Transmission Grating (LETG)

The LETG provides the highest spectral resolving power (E/ΔE > 1000) on Chandra at low energies (0.07 - 0.2 keV). The LETG/HRC-S combination is used extensively for high resolution spectroscopy of bright, soft sources such as stellar coronae, white dwarf atmospheres and cataclysmic variables.

Each of these instruments has a different photon collection efficiency – Effective Area. Reflectivity and vignetting, among other effects, cause the geometric area of a telescope to be reduced to a smaller "effective area".

Calibration Concordance Problem (Example: E0102)



- Four spectral lines observed with 11 X-ray detectors
- Main challenge – the data/instruments do not agree

Notation

- N Instruments with true effective area A_i, $1 \le i \le N$.
  - For each instrument i, we know estimated a_i ($\approx A_i$) but not A_i.
- M Sources with fluxes F_j, $1 \le j \le M$.
  - For each source j, F_j is unknown.
- Photon counts c_{ij}: from measuring flux F_j with instrument i.
- Lower cases: data / estimators.
- Upper cases: parameter / estimand.

Calibration Concordance Problem

1. Astronomers' Dilemma:

$$\frac{c_{ij}}{a_i} \neq \frac{c_{i'j}}{a_{i'}} \text{ for } i \neq i'.$$

Different instruments give different estimated flux of the same object!

2. Scientific Question:

    - Are there systematic errors in 'known' effective areas?
    - Can we derive properly adjusted effective areas?
    - Can we unify estimates of the same flux with different instruments?

Scientific and Statistical Models

Scientific Model

Multiplicative in original scale and additive on the log scale.

$$\text{Counts} = \text{Exposure} \times \text{Effective Area} \times \text{Flux},$$
$$C_{ij} = T_{ij} A_i F_j, \quad \Leftrightarrow \quad \log C_{ij} = B_i + G_j,$$

where log area $= B_i = \log A_i$, log flux $= G_j = \log F_j$; let $T_{ij} = 1$.

Statistical Model

$$\text{log counts } y_{ij} = \log c_{ij} - \alpha_{ij} = B_i + G_j + e_{ij}, \quad e_{ij} \overset{indep}{\sim} \mathcal{N}(0, \sigma_{ij}^2);$$

where $\alpha_{ij} = -0.5\sigma_{ij}^2$ to ensure $E(c_{ij}) = C_{ij} = A_i F_j$.

- **Known Variances**: σ_{ij} known.
- **Unknown Variances**: $\sigma_{ij} = \sigma_i$ unknown.

Bayesian Hierarchical Model

Log-Normal Hierarchical Model.

$$\begin{aligned}
\text{log counts } |\textit{area} \,\&\, \textit{flux} \,\&\, \textit{variance} &\overset{\text{indep}}{\sim} \text{Gaussian distribution}, \\
y_{ij} \mid B_i,\ G_j,\ \sigma_i^2 &\overset{\text{indep}}{\sim} \mathcal{N}\left(B_i + G_j,\ \sigma_i^2\right), \\
B_i &\overset{\text{indep}}{\sim} N(b_i,\ \tau_i^2), \\
G_j &\overset{\text{indep}}{\sim} \text{flat prior}, \\
\text{If variance unknown: } \sigma_i^2 &\overset{\text{indep}}{\sim} \text{Inv-Gamma}(df_g,\ \beta_g).
\end{aligned}$$

Setting the prior parameters.

1. $b_i = \log a_i$, τ_i are given by astronomers.
2. df_g, β_g are given based on the variability in data.

Extentions: Log-t Model

Question: Outliers? Less restrictions on the variances?

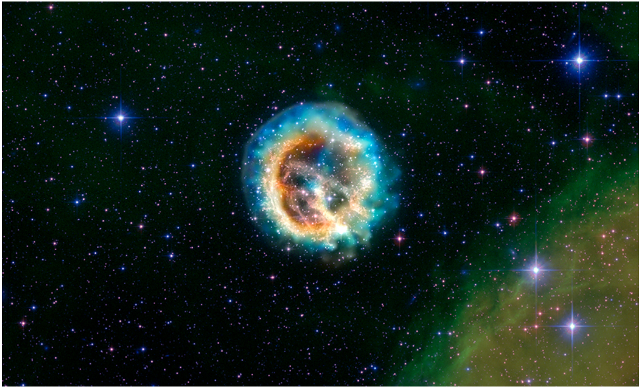
$$y_{ij} \mid B_i,\ G_j,\ \xi_{ij} = -\frac{\sigma^2}{2\xi_{ij}} + B_i + G_j + \frac{Z_{ij}}{\sqrt{\xi_{ij}}},$$
$$Z_{ij} \overset{\text{indep}}{\sim} N(0, \sigma^2),$$
$$B_i \overset{\text{indep}}{\sim} N(b_i, \tau_i^2).$$

If $\xi_{ij} \overset{\text{indep}}{\sim} \chi^2_\nu$, i.e. independent chi-squared distributions, the error term $Z_{ij}/\sqrt{\xi_{ij}}$ follows independent student-t distributions, i.e. $\frac{Z_{ij}}{\sqrt{\xi_{ij}}} \overset{\text{indep}}{\sim} \frac{\sigma}{\sqrt{\nu}} \mathrm{t}_\nu$.

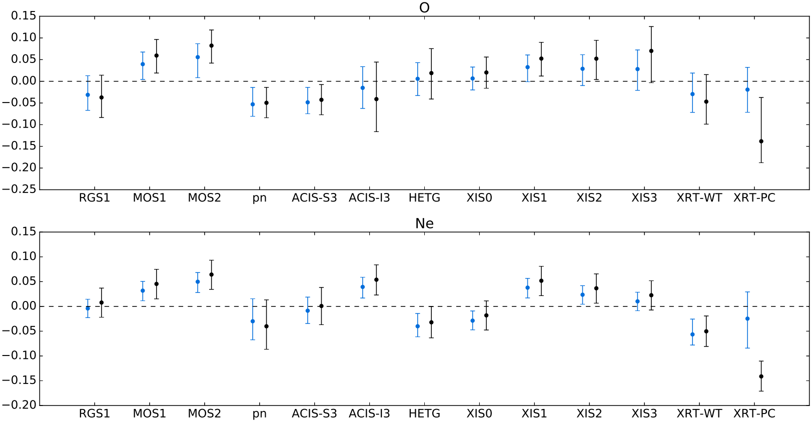
Numerical Results (E0102)

Recap: Supernova remnant E0102.

Four sources are four spectral lines in E0102.



Estimates of $B_i = \log A_i$ ($M = 2$ each panel)



- Adjusted so that default effective area, $b_i = \log a_i = 0$.
- 95% posterior intervals (black:$\tau = 0.05$; blue: $\tau = 0.025$).
- Some instruments systematically high, others low.

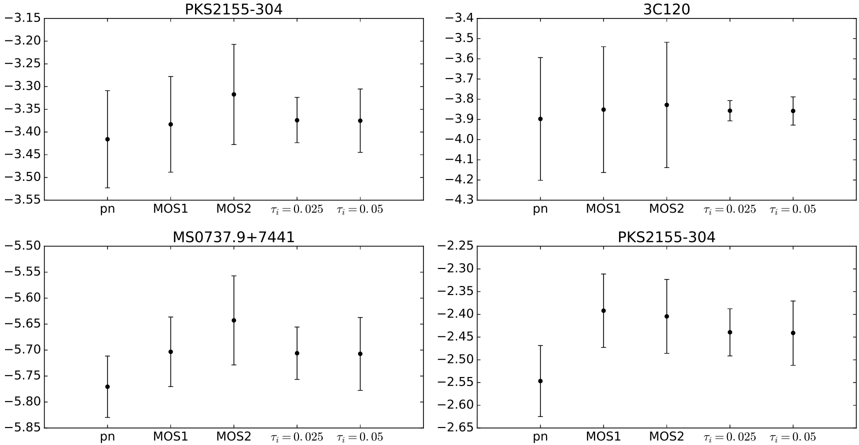
Numerical Results (XCAL)

- **XCAL data**: Bright active galactic nuclei from the XMM-Newton cross-calibration sample.
  - Observed in hard ($n = 94$), medium ($n = 103$), soft ($n = 108$) bands.
- **Pileup**: Image data are clipped to eliminate the regions affected by pileup, determined using `epatplot`.
- **Three detectors**: MOS1, MOS2 and pn.
- We fit our model and show results on

Sources: M=103 (in medium band).

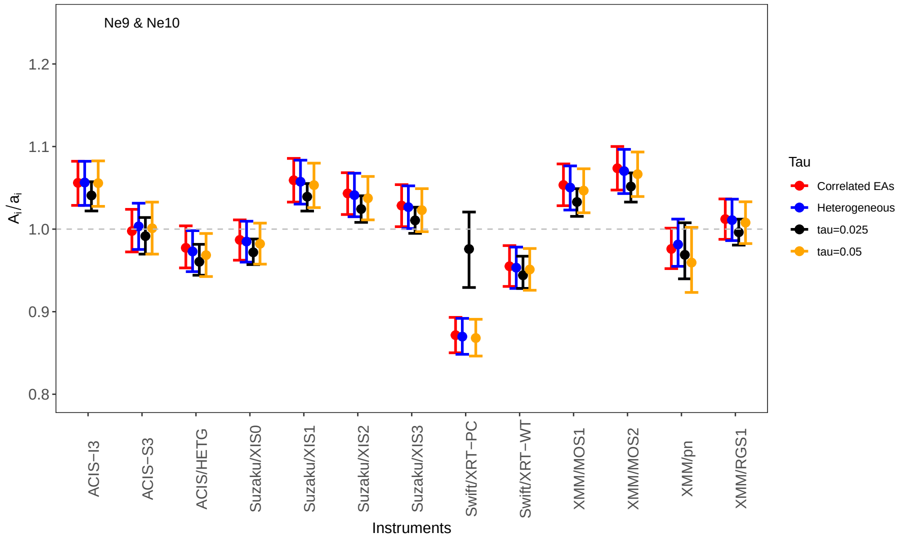
The hard and soft bands data are fitted similarly – treating hard/medium/soft band as three different data sets.

Numerical Results (XCAL): Calibration Concordance

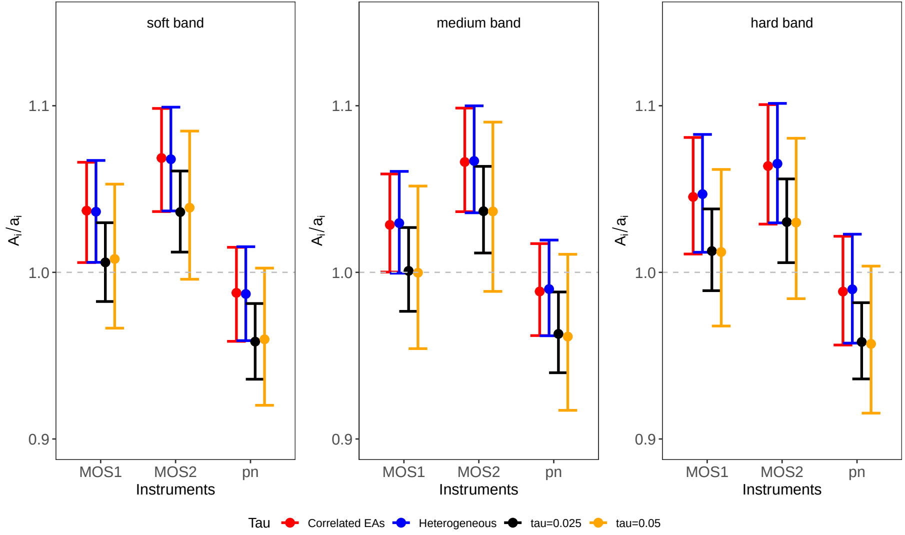


4 out of 103 Sources in medium band. y-axis: G (log flux); vertical bars (left 3 in each panel): mean $\pm$ 2 s.d. based on observed fluxes, vertical bars (right 2 in each panel): 95% posterior intervals based on our model.

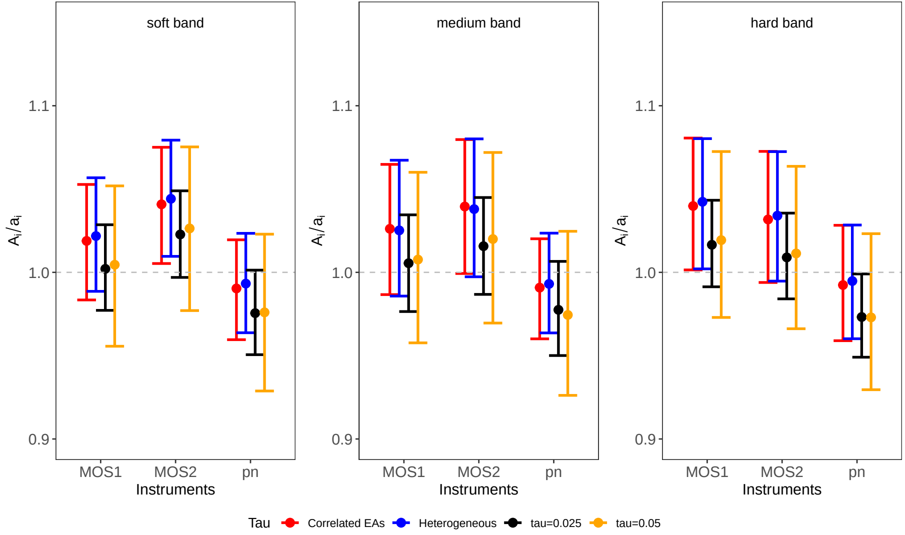
Extensions to Account for Correlated Energy Bands



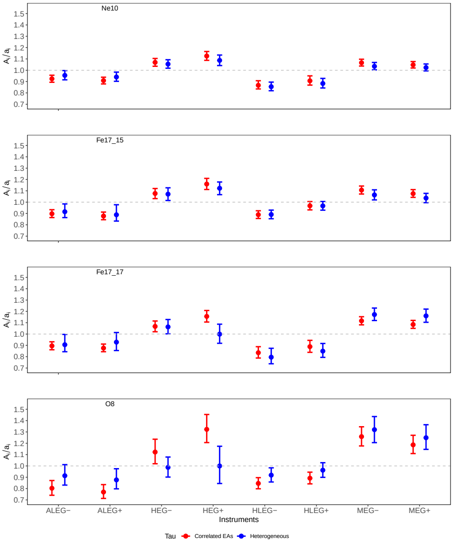
Extensions to Account for Correlated Energy Bands

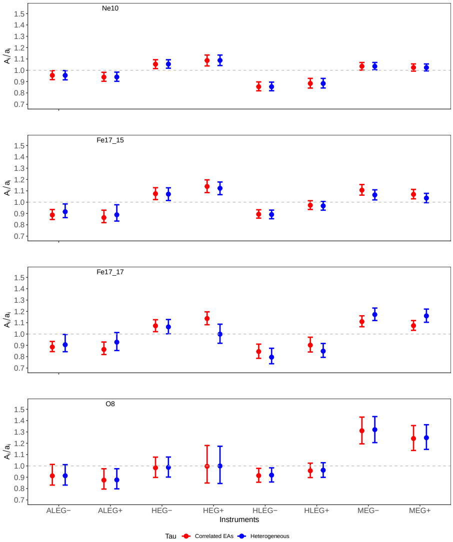


Extensions to Account for Correlated Energy Bands



Extensions to Account for Correlated Energy Bands





"History" of C statistics in Astrophysics

- Cash, W., *Parameter estimation in astronomy through application of the likelihood ratio*, Astrophysical Journal, Part 1, vol. 228, Mar. 15, 1979, p. 939-947.
  - Inference: MLE & confidence intervals
  - Goodness-of-fit Test: χ^2 for difference of likelihood ratios – if there exists a hypothesized fixed subset of parameters.
- Kaastra, J. S. *On the use of C-stat in testing models for X-ray spectra*, Astronomy & Astrophysics 605 (2017): A51.
  - Goodness-of-fit Test: Approximate Gaussian
- Bonamente, Massimiliano. *Distribution of the C statistic with applications to the sample mean of Poisson data.* Journal of Applied Statistics 47.11 (2020): 2044-2065.
  - Homogeneous Poisson rates & Approximate Gaussian interval.

Mathematical Notations for C-stat

Let null set be $\mathcal{S} = \{s_i(\theta), 1 \le i \le I\} \subset \mathbb{R}^I$. Under the null model, the maximum likelihood estimate for θ is

$$\hat{\theta}_I = \mathrm{argmax}_{\theta \in \mathbb{R}^d} \left\{ L(s_1(\theta), \ldots, s_I(\theta) | N_1, \ldots, N_I) = p(N_1, \ldots, N_I | \theta) \right\}.$$

The saturated model is $N_i \overset{\text{indep.}}{\sim} \mathrm{Poisson}(s_i)$, $1 \le i \le I$. The maximum likelihood estimate for s_i is $\hat{s}_i = N_i$. The log likelihood ratio statistics is

$$\begin{aligned} \mathrm{LR}_I = -2 \log \Lambda_I &= -2 \log \frac{\sup_{\mathcal{S}} L(s_1, \ldots, s_I | N_1, \ldots, N_I)}{\sup_{\mathbb{R}^n} L(s_1, \ldots, s_I | N_1, \ldots, N_I)} \\ &= 2 \sum_{i=1}^{I} \left[s_i(\hat{\theta}_n) - N_i \log s_i(\hat{\theta}_I) - N_i + N_i \log N_i \right]. \end{aligned}$$

This is the `C-stat` after plugging in the MLE $\hat{\theta}_I$, i.e. $\mathrm{LR}_I = C_I(\hat{\theta}_I)$, where the `C-stat`, denoted by $C_I(\theta)$, is defined as

$$C_I(\theta) = 2 \sum^{I} \left[s_i(\theta) - N_i \log s_i(\theta) - N_i + N_i \log N_i \right].$$

Likelihood Ratio Test and C-stat

The plug-in C statistic is not equal to the "true" C statistic:

Lemma (Wilk's Theorem)

For any n, $-C_n(\hat{\theta}_n) + C_n(\theta_0) = \mathrm{LR}_n^*$*, where* LR_n^* *is given by*

$$\begin{aligned}\mathrm{LR}_n^* &= -2\log\frac{L(s_1(\theta_0),\ldots,s_n(\theta_0)|N_1,\ldots,N_n)}{L(s_1(\hat{\theta}_n),\ldots,s_n(\hat{\theta}_n)|N_1,\ldots,N_n)} \\ &= 2\sum_{i=1}^{n}\left[N_i\log s_i(\hat{\theta}_n) - N_i\log s_i(\theta_0) + s_i(\theta_0) - s_i(\hat{\theta}_n)\right],\end{aligned}$$

which is the likelihood ratio statistics for testing the null hypothesis $H_0:\theta=\theta_0$ *versus the alternative* $H_1:\{s_i(\theta), 1\le i\le n\}\in\mathcal{S}$*. As* $n\to\infty$*,* $\mathrm{LR}_n^* \xrightarrow{\mathcal{D}} \chi_d^2$*.*

Asymptotic Normality

Without loss of generality, we can assume that all $s_i(\boldsymbol{\theta}^\star)$ are bounded from below and n is large.

Lemma (Problem Reduction due to Infinite Divisibility)

If $\sum_{i=1}^n s_i(\boldsymbol{\theta}^\star) \to \infty$*, then there exists* $\{m_1, \ldots, m_I\}$ *such that (1)* $\sum_{i=1}^n m_i \to \infty$*, (2)* $m_i = 1$ *when* $s_i(\boldsymbol{\theta}^\star) \leq 1$*, (3)* $0.5 < s_i(\boldsymbol{\theta}^\star)/m_i < 1$ *when* $s_i(\boldsymbol{\theta}^\star) > 1$*, (4) the likelihood is equivalent to the likelihood of the following model*

$$\tilde{N}_{ij} \overset{\text{indep.}}{\sim} \text{Poisson}\left(\frac{s_i(\boldsymbol{\theta})}{m_i}\right), \quad \sum_{j=1}^{m_i} \tilde{N}_{ij} = N_i. \tag{1}$$

Under mild regularity conditions, we have

$$\frac{C_n(\hat{\boldsymbol{\theta}}) - E[C_n(\hat{\boldsymbol{\theta}})]}{\sqrt{\text{Var}(C_n(\hat{\boldsymbol{\theta}}))}} \to N(0,1), \quad \text{as} \quad n \to \infty.$$

High-Order Asymptotics

Assume s_i follows log-linear model $\boldsymbol{\eta} = \boldsymbol{X\theta}$ where $\eta_i = \log s_i$. Let $V = \mathrm{diag}(s_i)$, $Q = (Q_{ij}) = X(X^\top VX)^{-1}X$, $\kappa_1^{(i)} = E(C_i)$, $\kappa_2^{(i)} = E(C_i - \kappa_1^{(i)})^2$, $\kappa_3^{(i)} = E(C_i - \kappa_1^{(i)})^3$, $\kappa_{11}^{(i)} = E\{(C_i - \kappa_1^{(i)})(N_i - s_i)\}$, $\kappa_{12}^{(i)} = E\{(C_i - \kappa_1^{(i)})(N_i - s_i)^2\}$, $\kappa_{21}^{(i)} = E\{(C_i - \kappa_1^{(i)})^2(N_i - s_i)\}$ and $\kappa_{03}^{(i)} = E(N_i - s_i)^3$. Then under regularity conditions,

$$E(C_{\min}|\hat{\boldsymbol{\theta}}) = \hat{\kappa}_1^{(\cdot)} - \frac{1}{2}\mathbf{1}^\top X^\top \hat{\Sigma} X(X^\top \hat{V} X)^{-1}\mathbf{1} + O(n^{-1/2}),$$

$$\mathrm{Var}(C_{\min}|\hat{\boldsymbol{\theta}}) = \hat{\kappa}_2^{(\cdot)} - \hat{\kappa}_{11}^\top X(X^\top \hat{V} X)^{-1} X^\top \hat{\kappa}_{11} + O(n^{-1/2}),$$

where $\Sigma = \mathrm{diag}\{\kappa_{12}^{(i)} - (\sum_j \kappa_{11}^{(j)} Q_{ji})\kappa_{03}^{(i)}\}$, $\kappa_{11} = (\kappa_{11}^{(1)}, \cdots, \kappa_{11}^{(n)})^\top$, $\kappa_1^{(\cdot)} = \sum_{i=1}^n \kappa_1^{(i)}$ and $\kappa_2^{(\cdot)} = \sum_{i=1}^n \kappa_2^{(i)}$.

Algorithms for Goodness-of-fit Assessment

Algorithm 1 Likelihood ratio with χ^2-statistics

Require: Data points: the N_i's, the number of bins n, and the number of unknown parameters to be estimated d.

1: Obtain $\hat{\boldsymbol{\theta}}$ via the following maximum likelihood estimation

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \log L_n(N_1, \cdots, N_n|\boldsymbol{\theta})$$

2: Calculate $s_i(\hat{\boldsymbol{\theta}}) = f_i(\hat{\boldsymbol{\theta}})$ and

$$C_{\min} = 2\sum_{i=1}^{n}[s_i(\hat{\boldsymbol{\theta}}) - N_i \log s_i(\hat{\boldsymbol{\theta}}) - N_i + N_i \log N_i]$$

3: Determine the p-value by

$$p = \max_p\{\chi^2_{n-d}(\frac{p}{2}) \le C_{\min} \le \chi^2_{n-d}(1-\frac{p}{2})\}.$$

4: **return** p

Algorithm 2 Asymptotic Normality – Bootstrap Mean/Variance

Require: Data points N_i's, the number of bins n, the number of parameters to be estimated d, and the number of bootstrap repetitions B.

1: Obtain $\hat{\boldsymbol{\theta}}$ via the maximum likelihood estimation $\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \log L_n(N_1, \cdots, N_n|\boldsymbol{\theta})$.

2: Calculate $s_i(\hat{\boldsymbol{\theta}}) = f_i(\hat{\boldsymbol{\theta}})$ and

$$C_{\min} = 2\sum_{i=1}^{n}[s_i(\hat{\boldsymbol{\theta}}) - N_i \log s_i(\hat{\boldsymbol{\theta}}) - N_i + N_i \log N_i]$$

3: **for** $m \in \{1, 2, \cdots, B\}$ **do**

4: Generate n Poisson samples denoted by $N_i^{(m)}$, $i = 1, \cdots, n$.

5: Obtain $\hat{\boldsymbol{\theta}}^{(m)}$ via the following maximum likelihood estimation

$$\hat{\boldsymbol{\theta}}^{(m)} = \arg\min_{\boldsymbol{\theta}^{(m)}\in\boldsymbol{\Theta}} \log L_n(N_1^{(m)}, \cdots, N_n^{(m)}|\boldsymbol{\theta})$$

6: Calculate $s_i^{(m)}(\hat{\boldsymbol{\theta}}^{(m)}) = f_i(\hat{\boldsymbol{\theta}}^{(m)})$ and

$$C_{\min}^{(m)} = 2\sum_{i=1}^{n}[s_i^{(m)}(\hat{\boldsymbol{\theta}}^{(m)}) - N_i^{(m)} \log s_i^{(m)}(\hat{\boldsymbol{\theta}}^{(m)}) - N_i^{(m)} + N_i^{(m)} \log N_i^{(m)}]$$

7: **end for**

8: Determine the bootstrap mean and variance

$$\mathbb{E}_b(C_{\min}) \approx \frac{\sum_{m=1}^{B} C_{\min}^{(m)}}{B}, \mathrm{Var}_b(C_{\min}) \approx \frac{\sum_{m=1}^{B}(C_{\min}^{(m)} - \mathbb{E}_b(C_{\min}))^2}{B-1}.$$

9: Determine the p-value by

$$p = \max_p\left\{Z(\frac{p}{2}) \le \frac{C_{\min} - \mathbb{E}_b(C_{\min})}{\sqrt{\mathrm{Var}_b(C_{\min})}} \le Z(1-\frac{p}{2})\right\},$$

where Z is the cumulative distribution function of the standard normal distribution.

10: **return** p

Algorithms for Goodness-of-fit Assessment

Algorithm 3 Asymptotic Normality – High Order

Require: Data points N_i's, the number of bins n and the number of parameters to be estimated d.

1: Obtain $\hat{\boldsymbol{\theta}}$ via the following maximum likelihood estimation

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \log L_n(N_1, \cdots, N_n|\boldsymbol{\theta})$$

2: Calculate $s_i(\hat{\boldsymbol{\theta}}) = f_i(\hat{\boldsymbol{\theta}})$ and

$$C_{\min} = 2\sum_{i=1}^{n}[s_i(\hat{\boldsymbol{\theta}}) - N_i \log s_i(\hat{\boldsymbol{\theta}}) - N_i + N_i \log N_i]$$

3: Determine the cumulants $\hat{\kappa}_1^{(i)}$, $\hat{\kappa}_{11}^{(i)}$, $\hat{\kappa}_{12}^{(i)}$, $\hat{\kappa}_{03}^{(i)}$, $\hat{V}$, $\hat{Q}$ and $\hat{\Sigma}$ via direct summation over each Poisson data N_i.

4: Determine the theoretical asymptotic mean and variance

$$E(C_{\min}|\hat{\boldsymbol{\theta}}) = \hat{\kappa}_1^{(\cdot)} - \frac{1}{2}\mathbf{1}^\top X^\top \hat{\Sigma} X (X^\top \hat{V} X)^{-1}\mathbf{1} + O(n^{-1/2}),$$

$$\mathrm{Var}(C_{\min}|\hat{\boldsymbol{\theta}}) = \hat{\kappa}_2^{(\cdot)} - \hat{\kappa}_{11}^\top X (X^\top \hat{V} X)^{-1} X^\top \hat{\kappa}_{11} + O(n^{-1/2}).$$

5: Determine the p-value by

$$p = \max_p \left\{ Z\left(\frac{p}{2}\right) \le \frac{C_{\min} - \mathbb{E}(C_{\min}|\hat{\boldsymbol{\theta}})}{\sqrt{\mathrm{Var}(C_{\min}|\hat{\boldsymbol{\theta}})}} \le Z\left(1 - \frac{p}{2}\right) \right\},$$

where Z is the cumulative distribution function of the standard normal distribution.

6: **return** p

Algorithm 4 Parametric Bootstrap

Require: Data points N_i's, the number of bins n, the number of parameters to be estimated d, and the number of bootstrap repetitions B.

1: Obtain $\hat{\boldsymbol{\theta}}$ via the following maximum likelihood estimation

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \log L_n(N_1, \cdots, N_n|\boldsymbol{\theta})$$

2: Calculate $s_i(\hat{\boldsymbol{\theta}}) = f_i(\hat{\boldsymbol{\theta}})$ and

$$C_{\min} = 2\sum_{i=1}^{n}[s_i(\hat{\boldsymbol{\theta}}) - N_i \log s_i(\hat{\boldsymbol{\theta}}) - N_i + N_i \log N_i]$$

3: **for** $m \in \{1, 2, \cdots, B\}$ **do**

4: Generate n Poisson samples denoted by $N_i^{(m)}$, $i = 1, \cdots, n$.

5: Obtain $\hat{\boldsymbol{\theta}}^{(m)}$ via the following maximum likelihood estimation

$$\hat{\boldsymbol{\theta}}^{(m)} = \arg\min_{\boldsymbol{\theta}^{(m)}\in\boldsymbol{\Theta}} \log L_n(N_1^{(m)}, \cdots, N_n^{(m)}|\boldsymbol{\theta})$$

6: Calculate $s_i^{(m)}(\hat{\boldsymbol{\theta}}^{(m)}) = f_i(\hat{\boldsymbol{\theta}}^{(m)})$ and

$$C_{\min}^{(m)} = 2\sum_{i=1}^{n}[s_i^{(m)}(\hat{\boldsymbol{\theta}}^{(m)}) - N_i^{(m)} \log s_i^{(m)}(\hat{\boldsymbol{\theta}}^{(m)}) - N_i^{(m)} + N_i^{(m)} \log N_i^{(m)}]$$

7: **end for**

8: Rearrange $C_{\min}^{(m)}$, $m = 1, 2, \cdots, B$ such that $C_{\min}^{(1)} \le C_{\min}^{(2)} \le \cdots \le C_{\min}^{(B)}$. And determine k such that $k = \min_k\{k|C_{\min}^{(k-1)} \le C_{\min} < C_{\min}^{(k)}\}$.

9: Determine the p-value by

$$p = \frac{2}{B}\min\{k, B - k\}.$$

10: **return** p

Numerical Studies: A simple example

We consider this example: $n = 100$, $\theta_1 = 2$, $\theta_2 = 1$, and

$$s_i = \theta_1 \exp\left(\theta_2 \times i/n\right), \quad i = 1, \ldots, n.$$

Table 1: The p-values of five numerical studies, $\theta_1 = 2.0$.

Test	1	2	3	4	5
Bootstrap test	0.112	0.732	0.316	0.124	0.610
C_{min} test	0.109	0.730	0.302	0.113	0.649
χ^2 test	0.028^{**}	0.184	0.063	0.025^{**}	0.153

Numerical Studies: Systematic Comparisons

Model A: Constant Rate Poisson Model, $s_i = \mu$, $\mu = \{0.5, 2, 5, 10\}$.

Model B: Varying Rate Poisson Model

- Pareto/Powerlaw Rates: $s_i(\boldsymbol{\theta}) = \mu(1 + i \times c_0)^{-k}$, where $c_0 = \frac{1}{n}$, $\mu = \{0.5, 2, 5, 10\}$ and $k = 1$.
- Exponential Rates: $s_i(\boldsymbol{\theta}) = \mu \exp(-i\eta)$, where $\mu = 5, 10, 100$ and $\eta = n^{-1}$.

Model C: Unstructured Rate Poisson Model: $s_i \sim \Gamma(\alpha, \beta)$, where $\beta = \sqrt{\alpha}$ and $\alpha = 25, 4, 0.25$, representing large, mixed and small count settings.

Numerical Studies: Systematic Comparisons

	Alg.1			Alg.2			Alg.3			Alg.4		
Model	n=10,50,100			n=10,50,100			n=10,50,100			n=10,50,100		
A-L-B	0.07	0.06	0.03	0.05	0.05	0.05	0.05	0.05	0.03	0.05	0.04	0.04
A-M-B	0.05	0.11	0.11	0.03	0.03	0.03	0.03	0.02	0.03	0.05	0.02	0.03
A-S-B	**0**	**0**	**0**	0.04	0.03	0.02	0.06	0.03	0.10	0.02	0.02	0.02
B-P-L	0.07	0.16	0.08	0.06	0.11	0.06	0.03	0.11	0.04	0.04	0.11	0.04
B-P-M	0.01	0.16	0.19	0.04	0.08	0.09	0.03	0.07	0.07	0.04	0.09	0.09
B-P-S	0.07	0.01	0.06	**0**	**0.02**	**0.01**	0.09	0.04	0.04	0.07	0.02	0.01
B-E-L	0.08	0.08	0.09	0.05	0.05	0.07	0.05	0.05	0.07	0.05	0.06	0.06
B-E-M	0.02	0.04	0.14	0.02	0.04	0.06	0.02	0.06	0.06	0.01	0.05	0.07
B-E-S	0.13	0.06	0.11	**0.03**	**0.01**	**0.01**	0.15	0.04	0.06	0.12	0	0.01

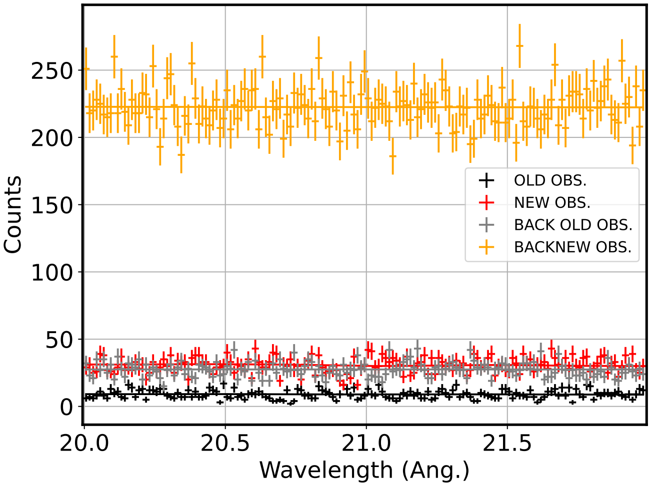
Table 1: Type I Error from 100 repeated simulation experiments with three different count settings under Models A and B: the null hypothesis is true.

Numerical Studies: Systematic Comparisons

	Alg.1			Alg.2			Alg.3			Alg.4		
Model	n=10,50,100			n=10,50,100			n=10,50,100			n=10,50,100		
C-U-L	0.90	0.43	0.21	0.92	0.41	0.22	0.92	0.42	0.21	0.93	0.47	0.25
C-U-M	0.83	0.38	0.05	0.80	0.44	0.19	0.84	0.45	0.17	0.85	0.49	0.23
C-U-S	0.79	0.38	0.19	0.61	0.16	0.01	**0.55**	**0.07**	**0**	0.70	0.18	0.02

Table 2: Type II Error from 100 repeated simulation experiments with three different count settings under Model C: the null hypothesis is not true..

Real Data Application



Real Data Application

Spectrum	$\hat{\mu}$	$C_{\min}$	Algorithm	$\mathbb{E}[C_{\min}]$	$\mathrm{Var}(C_{\min})$	p-value
Spec.I	8.962	190.72	Algo.1	158	316	0.078*
			Algo.2	162.48	338.74	0.125
			Algo.3	161.40	334.37	0.109
			Algo.4	N/A	N/A	0.128
Spec.II	30.704	167.67	Algo.1	158	316	0.568
			Algo.2	161.21	329.64	0.722
			Algo.3	158.89	321.70	0.624
			Algo.4	N/A	N/A	0.690
Spec.III	27.478	171.39	Algo.1	158	316	0.441
			Algo.2	160.81	328.85	0.560
			Algo.3	159.00	322.17	0.490
			Algo.4	N/A	N/A	0.548
Spec.IV	222.54	153.46	Algo.1	158	316	0.826
			Algo.2	159.20	324.53	0.750
			Algo.3	158.12	318.43	0.794
			Algo.4	N/A	N/A	0.760

Table 3: Performance of four test methods in each spectrum.

Summary

Instrument Calibration

1. *Multiplicative* mean modeling:

   log-Normal hierarchical model.

2. Shrinkage estimators.
3. Bayesian computation: MCMC & Stan.
4. The potential pitfalls of assuming 'known' variances.
5. Adjustments of effective areas of each instrument.

Goodness-of-fit

1. Systematic study of options for Goodness-of-fit and Python Package.

Acknowledgement

Yang Chen (UMich), Xufei Wang (Two Sigma), Xiao-Li Meng (Harvard), David van Dyk (ICL), Herman Marshall (MIT) & Vinay Kashyap (cfA)