Astrostatistics:
The Intersection of Statistics and Outer Space

Xiao-Li Meng (HU), Aneta Siemiginowska and Vinay Kashyap (CfA)

Joint work with Y. Chen (Michigan), X. Wang (Two Sigma Inc.), D. van Dyk (Imperial College London), H. Marshall (MIT)

October 29, 2019
Calibration Concordance Problem (Example: E0102)

- Supernova remnant E0102
- Four sources correspond to four spectral lines in E0102
Calibration Concordance Problem (Example: E0102)

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

1. Introduction

2. Scientific and Statistical Models

3. Concordance Model

4. Advantages of Our Approach
   - Multiplicative Shrinkages
   - Benefits of fitting the variances
   - Extentions to handle outliers
   - Results from Astronomy Data

5. Summary
Introduction

Scientific and Statistical Models

Concordance Model

Advantages of Our Approach
- Multiplicative Shrinkages
- Benefits of fitting the variances
- Extentions to handle outliers
- Results from Astronomy Data

Summary
Notation

- $N$ Instruments with true effective area $A_i$, $1 \leq i \leq N$.

- For each instrument $i$, we know estimated $a_i$ ($\approx A_i$) but not $A_i$. 
Notation

- \( N \) Instruments with true effective area \( A_i, 1 \leq i \leq N \).
  - For each instrument \( i \), we know estimated \( a_i \) (\( \approx A_i \)) but not \( A_i \).

- \( M \) Sources with fluxes \( F_j, 1 \leq j \leq M \).
  - For each source \( j \), \( F_j \) is unknown.
Notation

- \( N \) Instruments with true effective area \( A_i \), \( 1 \leq i \leq N \).
  - For each instrument \( i \), we know estimated \( a_i \) (\( \approx A_i \)) but not \( A_i \).

- \( M \) Sources with fluxes \( F_j \), \( 1 \leq j \leq M \).
  - For each source \( j \), \( F_j \) is unknown.

- Photon counts \( c_{ij} \): from measuring flux \( F_j \) with instrument \( i \).
Notation

- $N$ Instruments with true effective area $A_i$, $1 \leq i \leq N$.
  - For each instrument $i$, we know estimated $a_i$ ($\approx A_i$) but not $A_i$.

- $M$ Sources with fluxes $F_j$, $1 \leq j \leq M$.
  - 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:

\[
\frac{c_{ij}}{a_i} \neq \frac{c_{i'j}}{a_{i'}} \quad \text{for} \quad i \neq i'.
\]

Different instruments give different estimated flux of the same object!
Calibration Concordance Problem

1. Astronomers’ Dilemma:

\[
\frac{c_{ij}}{a_i} \neq \frac{c_{i'j}}{a_{i'}} \quad \text{for } i \neq i'.
\]

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?
Introduction

Scientific and Statistical Models

Concordance Model

Advantages of Our Approach
- Multiplicative Shrinkages
- Benefits of fitting the variances
- Extensions to handle outliers
- Results from Astronomy Data

Summary
Scientific and Statistical Models

Scientific Model

Multiplicative in original scale and additive on the log scale.

Counts = Exposure × Effective Area × Flux,

\[ C_{ij} = T_{ij}A_iF_j, \quad \Leftrightarrow \quad \log C_{ij} = B_i + G_j, \]

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

Statistical Model

\[ \log \text{counts} \ y_{ij} = \log c_{ij} - \alpha_{ij} = B_i + G_j + e_{ij}, \]
\[ e_{ij} \overset{\text{indep}}{\sim} \mathcal{N}(0, \sigma_{ij}^2); \]
where \( \alpha_{ij} = -0.5\sigma_{ij}^2 \) to ensure \( E(c_{ij}) = C_{ij} = A_i F_j. \)

- **Known Variances**: \( \sigma_{ij} \) known.
- **Unknown Variances**: \( \sigma_{ij} = \sigma_i \) unknown.
1. Introduction

2. Scientific and Statistical Models

3. Concordance Model

4. Advantages of Our Approach
   - Multiplicative Shrinkages
   - Benefits of fitting the variances
   - Extentions to handle outliers
   - Results from Astronomy Data

5. Summary
Bayesian Hierarchical Model

Log-Normal Hierarchical Model.

\[
\begin{align*}
\text{log counts} \mid \text{area & flux & variance} & \sim \text{Gaussian distribution,} \\
y_{ij} \mid B_i, G_j, \sigma^2_i & \sim N (B_i + G_j, \sigma^2_i),
\end{align*}
\]
Log-Normal Hierarchical Model.

\[
\begin{align*}
\text{log counts} \mid \text{area & flux & variance} & \sim \text{Gaussian distribution}, \\
y_{ij} \mid B_i, \ G_j, \ \sigma_i^2 & \sim N (B_i + G_j, \ \sigma_i^2), \\
B_i & \sim N(b_i, \ \tau_i^2), \\
G_j & \sim \text{flat prior},
\end{align*}
\]
Bayesian Hierarchical Model

Log-Normal Hierarchical Model.

\[
\begin{align*}
\text{log counts} & \mid \text{area} & \& \text{flux} & \& \text{variance} \\
y_{ij} & \mid B_i, \ G_j, \ \sigma_i^2 \\
B_i & \sim \text{Gaussian distribution}, \\
G_j & \sim \text{flat prior}, \\
\text{If variance unknown: } \sigma_i^2 & \sim \text{Inv-Gamma}(df_g, \ \beta_g). 
\end{align*}
\]

Setting the prior parameters.

1. \( b_i = \log a_i, \ \tau_i \) are given by astronomers.
Bayesian Hierarchical Model

Log-Normal Hierarchical Model.

\[
\begin{align*}
\text{log counts} & \mid \text{area} \& \text{flux} \& \text{variance} \\
\log y_{ij} & \mid B_i, \quad G_j, \quad \sigma^2_i \quad \text{iindp} \sim \mathcal{N}(B_i + G_j, \quad \sigma^2_i), \\
B_i & \quad \text{iindp} \sim \mathcal{N}(b_i, \quad \tau_i^2), \\
G_j & \quad \text{iindp} \sim \text{flat prior}, \\
\sigma^2_i & \quad \text{iindp} \sim \text{Inv-Gamma}(df_g, \quad \beta_g).
\end{align*}
\]

Setting the prior parameters.

1. \( b_i = \log a_i, \quad \tau_i \) are given by astronomers.

2. \( df_g, \beta_g \) are given based on the variability in data.
Posterior Propriety and Identifiability

**Posterior Propriety.** The posterior is proper if each source is measured by at least one instrument, i.e., $|l_j| \geq 1$ for all $1 \leq j \leq M$. 

Identifiability
Concordance Model

Posterior Propriety and Identifiability

Posterior Propriety. The posterior is proper if each source is measured by at least one instrument, i.e., \(|l_j| \geq 1\) for all \(1 \leq j \leq M\).

Identifiability

- \(\tau_i^2 = \infty\): same posteriors with \(\{B_i, G_j\}\) and \(\{B_i + \delta, G_j - \delta\}\);

Identifiability of \(\dagger\)
Posterior Propriety and Identifiability

Posterior Propriety. The posterior is proper if each source is measured by at least one instrument, i.e., $|I_j| \geq 1$ for all $1 \leq j \leq M$.

Identifiability

• $\tau_i^2 = \infty$: same posteriors with $\{B_i, G_j\}$ and $\{B_i + \delta, G_j - \delta\}$;

• the condition number of $\Omega(\sigma^2)$ (conditional variance of $B, G$) is

$$\frac{\lambda_{\text{max}}(\Omega(\sigma^2))}{\lambda_{\text{min}}(\Omega(\sigma^2))} \geq \frac{u^\top \Omega(\sigma^2) u}{v^\top \Omega(\sigma^2) v} = 1 + \frac{4 \sum_{i=1}^{N} |J_i| \sigma_i^{-2}}{\sum_{i=1}^{N} \tau_i^{-2}},$$

(1)

where $u = (1_N, 1_M)^\top$ and $v = (1_N, -1_M)^\top$. 
Concordance Model

Posterior Propriety and Identifiability

**Posterior Propriety.** The posterior is proper if each source is measured by at least one instrument, i.e., $|I_j| \geq 1$ for all $1 \leq j \leq M$.

**Identifiability**

- $\tau_i^2 = \infty$: same posteriors with $\{B_i, G_j\}$ and $\{B_i + \delta, G_j - \delta\}$;
- the *condition number* of $\Omega(\sigma^2)$ (conditional variance of $B, G$) is

$$
\frac{\lambda_{\text{max}}(\Omega(\sigma^2))}{\lambda_{\text{min}}(\Omega(\sigma^2))} \geq \frac{u^\top \Omega(\sigma^2) u}{v^\top \Omega(\sigma^2) v} = 1 + \frac{4 \sum_{i=1}^{N} |J_i| \sigma_i^{-2}}{\sum_{i=1}^{N} \tau_i^{-2}},
$$

where $u = (1_N, 1_M)^\top$ and $v = (1_N, -1_M)^\top$.
- $\{\tau_i^2\} >> \{\sigma_i^2\}$: elongated posterior contours.
Posterior Propriety and Identifiability

**Posterior Propriety.** The posterior is proper if each source is measured by at least one instrument, i.e., $|l_j| \geq 1$ for all $1 \leq j \leq M$.

**Identifiability**
- $\tau_i^2 = \infty$: same posteriors with $\{B_i, G_j\}$ and $\{B_i + \delta, G_j - \delta\}$;
- the *condition number* of $\Omega(\sigma^2)$ (conditional variance of $B, G$) is
  \[
  \frac{\lambda_{\text{max}}(\Omega(\sigma^2))}{\lambda_{\text{min}}(\Omega(\sigma^2))} \geq \frac{u^\top \Omega(\sigma^2) u}{v^\top \Omega(\sigma^2) v} = 1 + \frac{4 \sum_{i=1}^N |J_i| \sigma_i^{-2}}{\sum_{i=1}^N \tau_i^{-2}},
  \]
  \[
  (1)
  \]
  where $u = (1_N, 1_M)^\top$ and $v = (1_N, -1_M)^\top$.
- $\{\tau_i^2\} >> \{\sigma_i^2\}$: elongated posterior contours.

Alternative: setting $B_1 = 0$ or $\tau_1 = 0$. 
Bayesian Computation (Unknown Variances)

Markov Chain Monte Carlo (MCMC) algorithms.
Bayesian Computation (Unknown Variances)

Markov Chain Monte Carlo (MCMC) algorithms.

- Gibbs Sampling: update parameters one-at-a-time.

The joint distribution of the $B_i$ and $G_j$ is Gaussian.

Hamiltonian Monte Carlo (HMC) – Stan package.

Highly correlated parameters, high-dim parameter space.
Bayesian Computation (Unknown Variances)

Markov Chain Monte Carlo (MCMC) algorithms.

- Gibbs Sampling: update parameters one-at-a-time.
- Block Gibbs Sampling: update vectors of parameters.
Bayesian Computation (Unknown Variances)

Markov Chain Monte Carlo (MCMC) algorithms.

- Gibbs Sampling: update parameters one-at-a-time.
- Block Gibbs Sampling: update vectors of parameters.
  - The joint distribution of the $B_i$ and $G_j$ is Gaussian.
Bayesian Computation (Unknown Variances)

Markov Chain Monte Carlo (MCMC) algorithms.

- Gibbs Sampling: update parameters one-at-a-time.
- Block Gibbs Sampling: update vectors of parameters.
- The joint distribution of the $B_i$ and $G_j$ is Gaussian.
- Hamiltonian Monte Carlo (HMC) – Stan package.
Bayesian Computation (Unknown Variances)

Markov Chain Monte Carlo (MCMC) algorithms.

- Gibbs Sampling: update parameters one-at-a-time.
- Block Gibbs Sampling: update vectors of parameters.
  - The joint distribution of the $B_i$ and $G_j$ is Gaussian.
- Hamiltonian Monte Carlo (HMC) – Stan package.
  - Highly correlated parameters, high-dim parameter space.
1 Introduction

2 Scientific and Statistical Models

3 Concordance Model

4 Advantages of Our Approach
   - Multiplicative Shrinkages
   - Benefits of fitting the variances
   - Extentions to handle outliers
   - Results from Astronomy Data

5 Summary
Introduction

Scientific and Statistical Models

Concordance Model

Advantages of Our Approach
- Multiplicative Shrinkages
  - Benefits of fitting the variances
  - Extentions to handle outliers
  - Results from Astronomy Data

Summary
Shrinkage Estimators: Known Fluxes and Errors

Hierarchical model $\Rightarrow$ Shrinkage estimators (weighted averages of evidence from 'Prior' and evidence from 'Data').
Shrinkage Estimators: Known Fluxes and Errors

Hierarchical model \(\Rightarrow\) Shrinkage estimators (weighted averages of evidence from 'Prior' and evidence from 'Data').

(1) When fluxes and variances are known,

**Original Scale**

\[
\hat{A}_i = a_i^{W_i} \left[ (\bar{c}_i \cdot \bar{f}^{-1}) e^{\sigma^2_i / 2} \right]^{1-W_i},
\]

where

\[
\bar{c}_i = \prod_j c_{ij}^{1/M}, \quad \bar{f} = \prod_j f_j^{1/M}
\]

are geometric means.

**Log-Scale**

\[
\hat{B}_i = W_i b_i + (1 - W_i)(\bar{y}_i - \bar{G}),
\]

where

\[
\bar{G} = \frac{\sum_j g_j}{M}, \quad \bar{y}_i = \frac{\sum_j y_{ij}}{M}
\]

are arithmatic means.

The ‘weights’, \(W_i = \frac{\tau_i^{-2}}{\tau_i^{-2} + M\sigma_i^{-2}}\), represents the direct information in \(b_i\) relative to indirect information in fluxes.
Shrinkage Estimators: Known Errors

(2) When fluxes are unknown and variances are known,

\[
\hat{B}_i = W_i b_i + (1 - W_i)(\bar{y}_i - \bar{G}_i), \quad \hat{G}_j = \bar{y}_j - \bar{B},
\]

where \( \bar{G}_i = \frac{\sum_j \hat{G}_j}{M} \), \( \bar{B} = \frac{\sum_i \hat{B}_i \sigma_i^{-2}}{\sum_i \sigma_i^{-2}} \), \( \bar{y}_i = \frac{\sum_j y_{ij}}{M} \), \( \bar{y}_j = \frac{\sum_i y_{ij} \sigma_i^{-2}}{\sum_i \sigma_i^{-2}} \).
Advantages of Our Approach  Multiplicative Shrinkages

Shrinkage Estimators: Known Errors

(2) When fluxes are unknown and variances are known,

\[ \hat{B}_i = W_i b_i + (1 - W_i)(\bar{y}_i - \bar{G}_i), \quad \hat{G}_j = \bar{y}_j - \bar{B}, \]

where \( \bar{G}_i = \frac{\sum_j \hat{G}_j}{M} \), \( \bar{B} = \frac{\sum_i \hat{B}_i \sigma_i^{-2}}{\sum_i \sigma_i^{-2}} \), \( \bar{y}_i = \frac{\sum_j y_{ij}}{M} \), \( \bar{y}_j = \frac{\sum_i y_{ij} \sigma_i^{-2}}{\sum_i \sigma_i^{-2}} \).

(3) When variances are unknown, shrinkage estimator of variance,

\[ \hat{\sigma}_i^2 = \frac{2}{1 + \sqrt{1 + S_{y,i}^2}} \sigma_{y,i}^2, \quad S_{y,i}^2 = \frac{1}{|J_i| + \alpha} \left[ \sum_{j \in J_i} (y_{ij} - \hat{B}_i - \hat{G}_j)^2 + \beta \right] \]
1 Introduction

2 Scientific and Statistical Models

3 Concordance Model

4 Advantages of Our Approach
   - Multiplicative Shrinkages
   - Benefits of fitting the variances
   - Extensions to handle outliers
   - Results from Astronomy Data

5 Summary
Benefits of Fitting $\sigma_i^2$

- Tolerance to model/error model misspecification.
Benefits of Fitting $\sigma_i^2$

- Tolerance to model/error model misspecification.

- Pitfalls of assuming ‘known’ variances:
  - Overly optimistic ‘known variances’
    - $\Rightarrow$ overly narrow confidence intervals
  - $\Rightarrow$ possible false discoveries
Benefits of Fitting $\sigma_i^2$

- Tolerance to model/error model misspecification.

- Pitfalls of assuming ‘known’ variances:
  - Overly optimistic ‘known variances’
    - $\Rightarrow$ overly narrow confidence intervals
  - $\Rightarrow$ possible false discoveries
  - ‘known variances’ $\geq$ true variability
    - $\Rightarrow$ noninformative results
1 Introduction

2 Scientific and Statistical Models

3 Concordance Model

4 Advantages of Our Approach
   - Multiplicative Shrinkages
   - Benefits of fitting the variances
   - Extentions to handle outliers
   - Results from Astronomy Data

5 Summary
Question: Outliers? Less restrictions on the variances?
Auxiliary Model

**Question:** Outliers? Less restrictions on the variances?

\[ y_{ij} \mid B_i, G_j, \xi_{ij} = -\frac{\sigma^2}{2\xi_{ij}} + B_i + G_j + \frac{Z_{ij}}{\sqrt{\xi_{ij}}}, \]

\[ Z_{ij} \overset{\text{indep}}{\sim} N(0, \sigma^2), \]

\[ B_i \overset{\text{indep}}{\sim} N(b_i, \tau_i^2). \]
**Question:** Outliers? Less restrictions on the variances?

\[
y_{ij} \mid B_i, G_j, \xi_{ij} = -\frac{\sigma^2}{2\xi_{ij}} + B_i + G_j + \frac{Z_{ij}}{\sqrt{\xi_{ij}}},
\]

\[
Z_{ij} \overset{\text{iid}}{\sim} N(0, \sigma^2),
\]

\[
B_i \overset{\text{iid}}{\sim} N(b_i, \tau_i^2).
\]

If \( \xi_{ij} \overset{\text{iid}}{\sim} \chi^2_\nu \), i.e. independent chi-squared distributions, the error term

\[
Z_{ij}/\sqrt{\xi_{ij}}
\]

follows independent student-t distributions, i.e.

\[
\frac{Z_{ij}}{\sqrt{\xi_{ij}}} \overset{\text{iid}}{\sim} \frac{\sigma}{\sqrt{\nu}} t_\nu.
\]
A Numerical Example with Outliers

Simulation: $N = 10$, $M = 40$, $G_1 = -1$ and $G_j = 3, j > 1$.
Asymptotic variance of log-counts: $e^{-B_i - G_j} \Rightarrow$ outliers.
A Numerical Example with Outliers

Simulation: $N = 10$, $M = 40$, $G_1 = -1$ and $G_j = 3, j > 1$.
Asymptotic variance of log-counts: $e^{-B_i - G_j} \Rightarrow$ outliers.

\[ \hat{R}_{ij} = \frac{y_{ij} - \hat{B}_i - \hat{G}_j + 0.5 \times \hat{\sigma}_i^2}{\hat{\sigma}_i}, \quad \hat{R}_{ij} = \frac{y_{ij} - \hat{B}_i - \hat{G}_j + 0.5 \times \kappa^2 / \hat{\xi}_{ij}}{\kappa / \hat{\xi}_{ij}^{1/2}} \]
A Numerical Example with Outliers

Simulation: \( N = 10, M = 40, G_1 = -1 \) and \( G_j = 3, j > 1 \).

Asymptotic variance of log-counts: \( e^{-B_i - G_j} \Rightarrow \) outliers.

\[
\hat{R}_{ij} = \frac{y_{ij} - \hat{B}_i - \hat{G}_j + 0.5 \times \hat{\sigma}_i^2}{\hat{\sigma}_i}, \quad \hat{R}_{ij} = \frac{y_{ij} - \hat{B}_i - \hat{G}_j + 0.5 \times \kappa^2 / \hat{\xi}_{ij}}{\kappa / \hat{\xi}_{ij}^{1/2}}
\]
Coverage Properties With Outliers, Misspecification

<table>
<thead>
<tr>
<th>Poisson Model</th>
<th>Para.</th>
<th>Coverage Probability</th>
<th>Length of Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>log-Normal</td>
<td>log-t</td>
</tr>
<tr>
<td>N = 10</td>
<td>$B$</td>
<td>[0.941, 0.959]</td>
<td>0.067±0.005</td>
</tr>
<tr>
<td></td>
<td>$G_1$</td>
<td>0.399</td>
<td>0.090±0.015</td>
</tr>
<tr>
<td></td>
<td>$G_{2:M}$</td>
<td>[0.967, 0.977]</td>
<td>0.077±0.003</td>
</tr>
<tr>
<td>N = 40</td>
<td>$B$</td>
<td>[0.953, 0.969]</td>
<td>0.041±0.007</td>
</tr>
<tr>
<td></td>
<td>$G_1$</td>
<td>0.398</td>
<td>0.045±0.003</td>
</tr>
<tr>
<td></td>
<td>$G_{2:M}$</td>
<td>[0.965,0.977]</td>
<td>0.038±0.001</td>
</tr>
</tbody>
</table>

Table 1: $M = 40$. Coverage of nominal 95% posterior intervals calculated from 2000 datasets simulated under a Poisson model. The intervals in columns 3 and 4 give the smallest and largest coverage observed for the corresponding parameter. The last two columns give the lengths of nominal 95% intervals in the format: mean ± standard deviation.
1. Introduction

2. Scientific and Statistical Models

3. Concordance Model

4. Advantages of Our Approach
   - Multiplicative Shrinkages
   - Benefits of fitting the variances
   - Extentions to handle outliers
   - Results from Astronomy Data

5. Summary
**Recap**: Supernova remnant E0102.

Four sources are four spectral lines in E0102.
Advantages of Our Approach
Results from Astronomy Data

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.
Prior Influence

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Oxygen $\tau = 0.025$</th>
<th>Oxygen $\tau = 0.05$</th>
<th>Neon $\tau = 0.025$</th>
<th>Neon $\tau = 0.05$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGS1</td>
<td>0.570</td>
<td>0.205</td>
<td>0.063</td>
<td>0.016</td>
</tr>
<tr>
<td>MOS1</td>
<td>0.279</td>
<td>0.077</td>
<td>0.075</td>
<td>0.019</td>
</tr>
<tr>
<td>MOS2</td>
<td>0.355</td>
<td>0.065</td>
<td>0.077</td>
<td>0.017</td>
</tr>
<tr>
<td>pn</td>
<td>0.250</td>
<td>0.041</td>
<td>0.620</td>
<td>0.218</td>
</tr>
<tr>
<td>ACIS-S3</td>
<td>0.218</td>
<td>0.040</td>
<td>0.270</td>
<td>0.088</td>
</tr>
<tr>
<td>ACIS-I3</td>
<td>0.906</td>
<td>0.640</td>
<td>0.099</td>
<td>0.026</td>
</tr>
<tr>
<td>HETG</td>
<td>0.648</td>
<td>0.341</td>
<td>0.129</td>
<td>0.034</td>
</tr>
<tr>
<td>XIS0</td>
<td>0.180</td>
<td>0.051</td>
<td>0.069</td>
<td>0.018</td>
</tr>
<tr>
<td>XIS1</td>
<td>0.298</td>
<td>0.078</td>
<td>0.071</td>
<td>0.019</td>
</tr>
<tr>
<td>XIS2</td>
<td>0.463</td>
<td>0.140</td>
<td>0.063</td>
<td>0.016</td>
</tr>
<tr>
<td>XIS3</td>
<td>0.772</td>
<td>0.364</td>
<td>0.062</td>
<td>0.018</td>
</tr>
<tr>
<td>XRT-WT</td>
<td>0.726</td>
<td>0.278</td>
<td>0.154</td>
<td>0.026</td>
</tr>
<tr>
<td>XRT-PC</td>
<td>0.934</td>
<td>0.235</td>
<td>0.906</td>
<td>0.017</td>
</tr>
</tbody>
</table>

Table 2: Proportion of prior influence, as defined by $1 - W_i$, for E0102 data.
Numerical Results (2XMM)

- 2XMM catalog: used to generate large, well-defined samples of various types of astrophysical objects; collected with the XMM-Newton European Photon Imaging Cameras (EPIC).
Numerical Results (2XMM)

- 2XMM catalog: used to generate large, well-defined samples of various types of astrophysical objects; collected with the XMM-Newton European Photon Imaging Cameras (EPIC).

- Three EPIC instruments: the EPIC-pn, and the two EPIC-MOS detectors (pn, MOS1, and MOS2).
Numerical Results (2XMM)

- 2XMM catalog: used to generate large, well-defined samples of various types of astrophysical objects; collected with the XMM-Newton European Photon Imaging Cameras (EPIC).

- Three EPIC instruments: the EPIC-pn, and the two EPIC-MOS detectors (pn, MOS1, and MOS2).

- Three datasets: hard (2.5 - 10.0 keV), medium (1.5 - 2.5 keV) and soft (0.5 - 1.5 keV) energy bands. The three instruments (pn, MOS1 and MOS2) measured 41, 41, and 42 sources respectively in hard, medium, and soft bands. Faint sources.
Figure 3: Adjustments of the log-scale Effective Areas for hard band (left), medium band (middle) and soft band (right) of the 2XMM datasets.
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 ± 2 s.d. based on observed fluxes, vertical bars (right 2 in each panel): 95% posterior intervals based on our model.
## Prior Influence

<table>
<thead>
<tr>
<th>Data Name</th>
<th>$\tau_i = 0.025$</th>
<th></th>
<th>$\tau_i = 0.05$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pn</td>
<td>mos1</td>
<td>mos2</td>
<td>pn</td>
</tr>
<tr>
<td>hard band 2XMM</td>
<td>0.093</td>
<td>0.075</td>
<td>0.082</td>
<td>0.025</td>
</tr>
<tr>
<td>medium band 2XMM</td>
<td>0.250</td>
<td>0.216</td>
<td>0.222</td>
<td>0.076</td>
</tr>
<tr>
<td>soft band 2XMM</td>
<td>0.093</td>
<td>0.075</td>
<td>0.069</td>
<td>0.025</td>
</tr>
<tr>
<td>hard band XCAL</td>
<td>0.010</td>
<td>0.019</td>
<td>0.031</td>
<td>0.003</td>
</tr>
<tr>
<td>medium band XCAL</td>
<td>0.023</td>
<td>0.016</td>
<td>0.028</td>
<td>0.006</td>
</tr>
<tr>
<td>soft band XCAL</td>
<td>0.021</td>
<td>0.011</td>
<td>0.007</td>
<td>0.005</td>
</tr>
</tbody>
</table>

**Table 3:** Proportion of prior influence.
1 Introduction

2 Scientific and Statistical Models

3 Concordance Model

4 Advantages of Our Approach
   - Multiplicative Shrinkages
   - Benefits of fitting the variances
   - Extensions to handle outliers
   - Results from Astronomy Data

5 Summary
Summary

Statistics

1. *Multiplicative* mean modeling:

   log-Normal hierarchical model.
Statistics

1. Multiplicative mean modeling:

   log-Normal hierarchical model.

2. Shrinkage estimators.
Summary

Statistics

1. *Multiplicative* mean modeling:

   log-Normal hierarchical model.

2. Shrinkage estimators.

Summary

Statistics

1. *Multiplicative* mean modeling:

   log- Normal hierarchical model.

2. Shrinkage estimators.


4. The potential pitfalls of assuming 'known' variances.
Summary

Statistics

1. *Multiplicative* mean modeling:

   log-Normal hierarchical model.

2. Shrinkage estimators.


4. The potential pitfalls of assuming 'known' variances.

Astronomy

1. Adjustments of effective areas of each instrument.
Summary

Statistics
1. *Multiplicative* mean modeling:
   
   log-Normal hierarchical model.

2. Shrinkage estimators.


4. The potential pitfalls of assuming 'known' variances.

Astronomy
1. Adjustments of effective areas of each instrument.
2. Calibration concordance.
Discussions: Ongoing and Future Work

- Correlations among instruments.
  - Estimated correlations based on theoretical simulations.
  - Prior? Extra data? Uncertainty?
Discussions: Ongoing and Future Work

- Correlations among instruments.
  - Estimated correlations based on theoretical simulations.
  - Prior? Extra data? Uncertainty?

- Robustness $\Rightarrow$ Misspecified models.
Discussions: Ongoing and Future Work

- Correlations among instruments.
  - Estimated correlations based on theoretical simulations.
  - Prior? Extra data? Uncertainty?

- Robustness $\Rightarrow$ Misspecified models.

- Better quantification of prior influence.
Discussions: Ongoing and Future Work

- Correlations among instruments.
  - Estimated correlations based on theoretical simulations.
  - Prior? Extra data? Uncertainty?

- Robustness $\Rightarrow$ Misspecified models.

- Better quantification of prior influence.

- Coverage properties when outliers exist.
Acknowledgement

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