Luminosity Functions

Lazhi Wang

Department of Statistics, Harvard University

May 15th, 2012

Introduction: What is a Luminosity Function?

Figure: A galaxy cluster.

Introduction: Project Goal

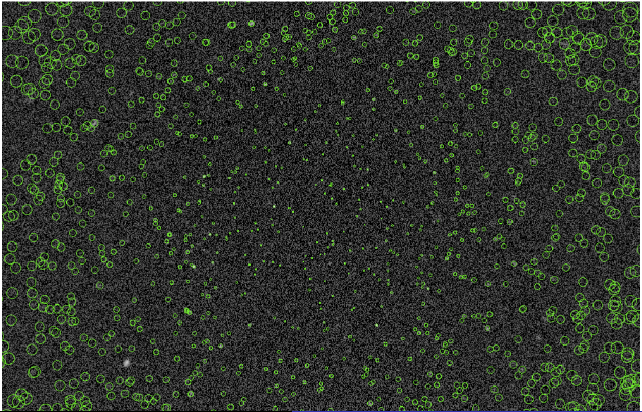
- The Luminosity Function specifies the relative number of sources at each luminosity.

Introduction: Project Goal

- The Luminosity Function specifies the relative number of sources at each luminosity.
- Goal of the Project: To develop a fully Bayesian model to infer the distribution of the luminosities of all the sources in a population.

Introduction: Data

- Y_i, observed photon counts, contaminated with background in a source exposure.
- X, observed photon counts in the exposure of pure background .



Bayesian Model

- Level I model:

$$
\begin{aligned}
X|\xi &\sim Pois(\xi), \\
Y_i &= Y_{iB} + Y_{iS}, \text{ where } Y_{iB}|\xi \sim Pois(a_i\xi), \\
Y_{iS}|\lambda_i &\sim Pois(b_i\lambda_i) \sim \begin{cases} \delta_0, & \text{if } \lambda_i = 0; \\ Pois(b_i\lambda_i), & \text{if } \lambda_i \neq 0. \end{cases}
\end{aligned}
$$

- ξ is the background intensity,
- λ_i is the intensity of source i,
- a_i is ratio of source area to background area (known constant),
- b_i is the telescope effective area (known constant).

Bayesian Model

- Level II model:

$$\xi \sim Gamma(\alpha_0, \beta_0),$$
$$\lambda_i|\alpha,\beta,\pi \begin{cases} = 0, & \text{with probability } 1-\pi; \\ \sim Gamma(\alpha,\beta), & \text{with probability } \pi. \end{cases}$$

- Level III model:

$$P(\alpha,\beta,\pi) \propto \frac{1}{\beta^3}\pi^{c_1-1}(1-\pi)^{c_2-1}.$$

Bayesian Model: Summary

- Model

$$
\begin{aligned}
X|\xi \sim Pois(\xi),\, Y_i = Y_{iB} + Y_{iS}, \\
Y_{iS}|\xi \sim Pois(a_i\xi),\, Y_{iB}|\lambda_i \sim Pois(b_i\lambda_i), \\
\xi \sim Gamma(\alpha_0, \beta_0), \\
\lambda_i|\alpha, \beta, \pi \begin{cases} = 0, & \text{with probability } 1-\pi, \\ \sim Gamma(\alpha, \beta), & \text{with probability } \pi, \end{cases} \\
P(\alpha, \beta, \pi) \propto \frac{1}{\beta^3}.
\end{aligned}
$$

- Research interest:
  - The posterior distribution of intensities λ,
  - The posterior distribution of $1-\pi$, the proportion of dark sources.

Simulation Study 1: Setup

- Number of sources, N=500.
- Distribution of λ's:

$$\lambda_i \overset{\text{iid}}{\sim} \begin{cases} \delta_0, \text{ with prob } 1-\pi = 0.15, \\ Gamma[10,35] \text{ with prob } 0.85. \end{cases}$$

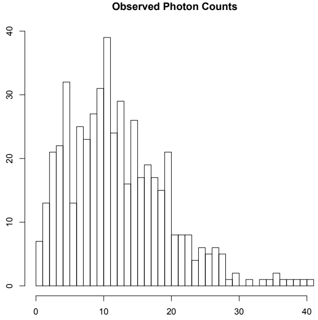
- Distribution of Y_{iS}:

$$Y_{iS}|\lambda_i \sim Pois(b_i\lambda_i).$$

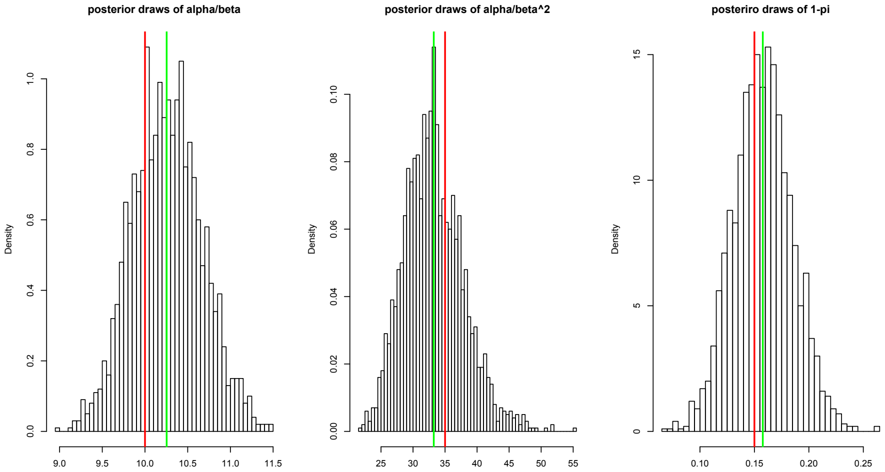
- Distribution of background noise Y_{iB}:

$$Y_{iB}|\xi \sim Pois(4), \text{ approximately}.$$

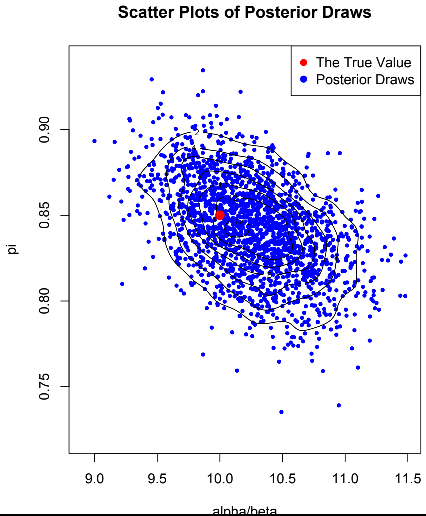
$$Y_i = Y_{iS} + Y_{iB}.$$

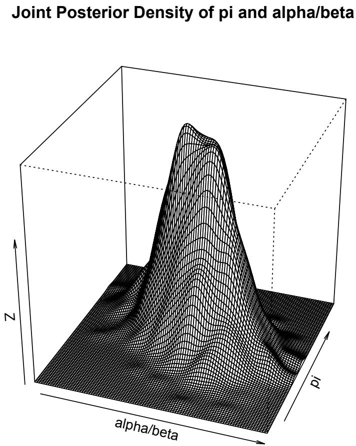


Posterior Distributions of the Hyper-parameters



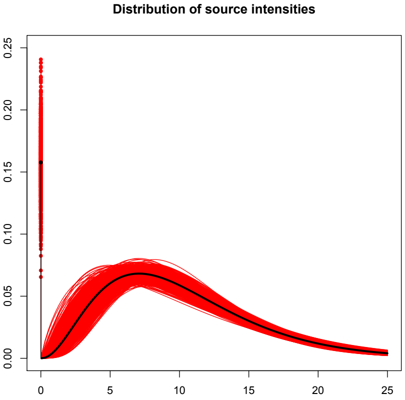
Posterior Distributions of the Hyper-parameters





Distribution of source intensities

$$\lambda_i \overset{\text{iid}}{\sim} \begin{cases} \delta_0, & \text{with prob } 1-\pi, \\ Gamma(\alpha,\beta), & \text{with prob } \pi. \end{cases}$$



Simulation Study 2: Setup

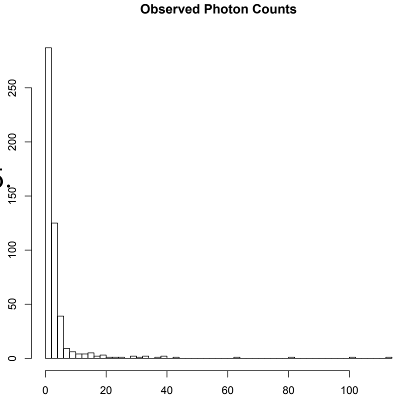
- Number of sources, N=500.
- Distribution of λ's:

$$\lambda_i \overset{iid}{\sim} \begin{cases} \delta_0, & \text{with prob } 1-\pi = 0.15, \\ Gamma[2,60] & \text{with prob } 0.85. \end{cases}$$

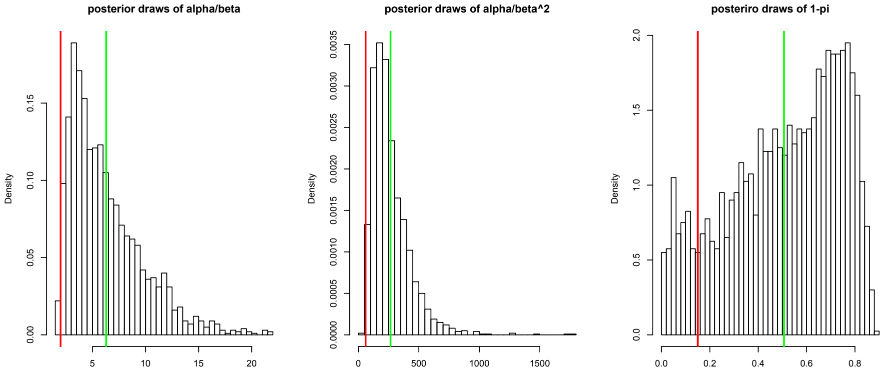
- Distribution of Y_{iS}:

$$Y_{iS}|\lambda_i \sim Pois(b_i\lambda_i).$$

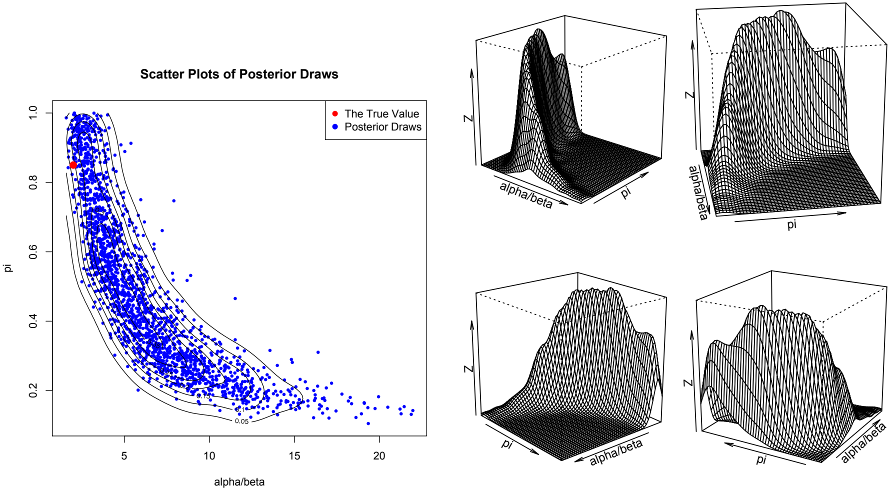
- $Y_{iB}|\xi \sim Pois(2)$, approximately.



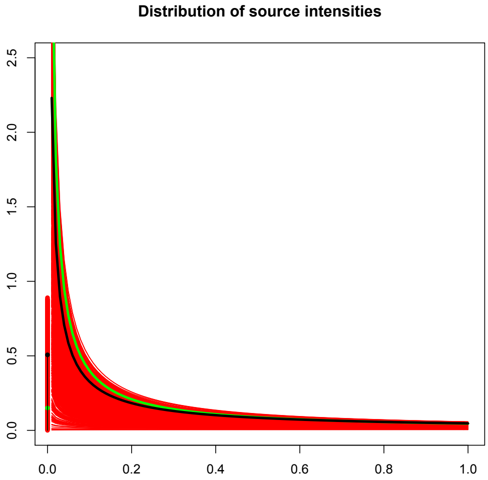
Simulation Study 2



Simulation Study 2



Simulation Study 2



Choices of Priors for the Hyper-parameters

- Recall

$$\lambda_i|\alpha,\beta,\pi \overset{\text{iid}}{\sim} \begin{cases} \delta_0, & \text{with prob } 1-\pi, \\ Gamma[\frac{\alpha}{\beta},\frac{\alpha}{\beta^2}] = Gamma[\mu,\frac{\mu}{\beta}] & \text{with prob } \pi. \end{cases}$$

Choices of Priors for the Hyper-parameters

- Recall

$$\lambda_i|\alpha,\beta,\pi \overset{\text{iid}}{\sim} \begin{cases} \delta_0, & \text{with prob } 1-\pi, \\ Gamma[\frac{\alpha}{\beta}, \frac{\alpha}{\beta^2}] = Gamma[\mu, \frac{\mu}{\beta}] & \text{with prob } \pi. \end{cases}$$

-

$$P(\mu,\beta,\pi)d\mu d\beta d\pi \propto P(\beta)P(\pi)d\mu d\beta d\pi \propto \frac{1}{\beta}P(\beta)P(\pi)d\alpha d\beta d\pi,$$

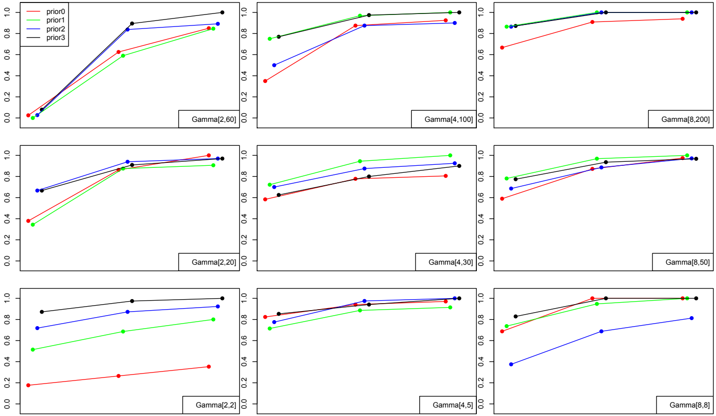
$$P(\alpha,\beta,\pi)d\alpha d\beta d\pi \propto \frac{1}{\beta^{c_3+1}}\pi^{c_1-1}(1-\pi)^{c_2-1}d\alpha d\beta d\pi$$

Choices of Priors: Simulation Study

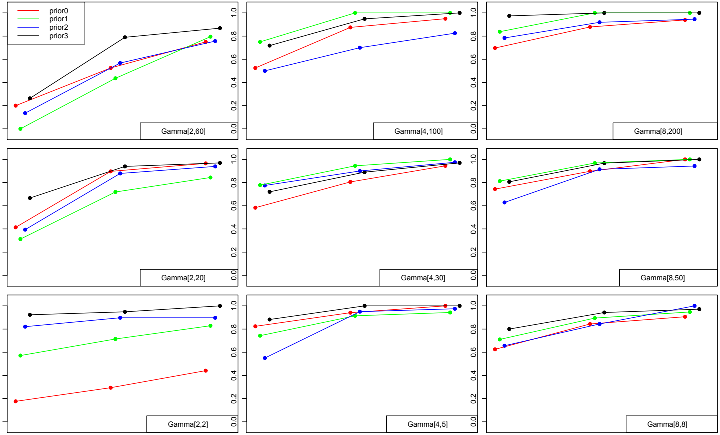
- $\pi \sim Beta(c_1, c_2)$, we can use informative prior if we have some prior information about the distribution of π. Otherwise, we can let $c_1 = c_2 = 1$, so $\pi \sim Unif(0, 1)$.
- Priors for (α, β):

$$\begin{aligned}\text{Prior 0: } & P(\alpha,\beta) \propto 1, \\ \text{Prior 1: } & P(\alpha,\beta) \propto \frac{1}{\beta}, \\ \text{Prior 2: } & P(\alpha,\beta) \propto \frac{1}{\beta^2}, \\ \text{Prior 3: } & P(\alpha,\beta) \propto \frac{1}{\beta^3}.\end{aligned}$$

Choices of Priors: Coverage for π



Choices of Priors: Coverage for $\frac{\alpha}{\beta}$



Choices of Priors

- Conclusion: the prior

$$P(\alpha, \beta) \propto \frac{1}{\beta^3}$$

  gives the highest frequency coverage in most simulation studies.
- This prior is called Stein's Harmonic Prior. The SHP prior is shown to provide estimators that have adequate frequency coverage.

Speeding up MCMC

- It takes about 3 hours to get 100,000 draws.
- Metropolis-Hastings algorithm within Gibbs Sampler:

$$P(\alpha|\beta, \pi, \xi, \underset{\sim}{\lambda}, \underset{\sim}{Y_B}, X, \underset{\sim}{Y}) \propto \left(\frac{(\beta\lambda^*)^\alpha}{\Gamma(\alpha)}\right)^K,$$

where $K = \sum 1_{\lambda_i \neq 0}$, and $\lambda^* = (\prod_{i,\lambda_i \neq 0} \lambda_i)^{1/K}$.

- Is there a good way to sample from the conditional posterior distribution?

Acknowledgement

- Vinay Kashyap,
- David van Dyk,
- Andreas Zezas.