

Bayesian Statistics

Analysis of Hierarchical Models

Nan Lin

Department of Mathematics

Washington University in St. Louis

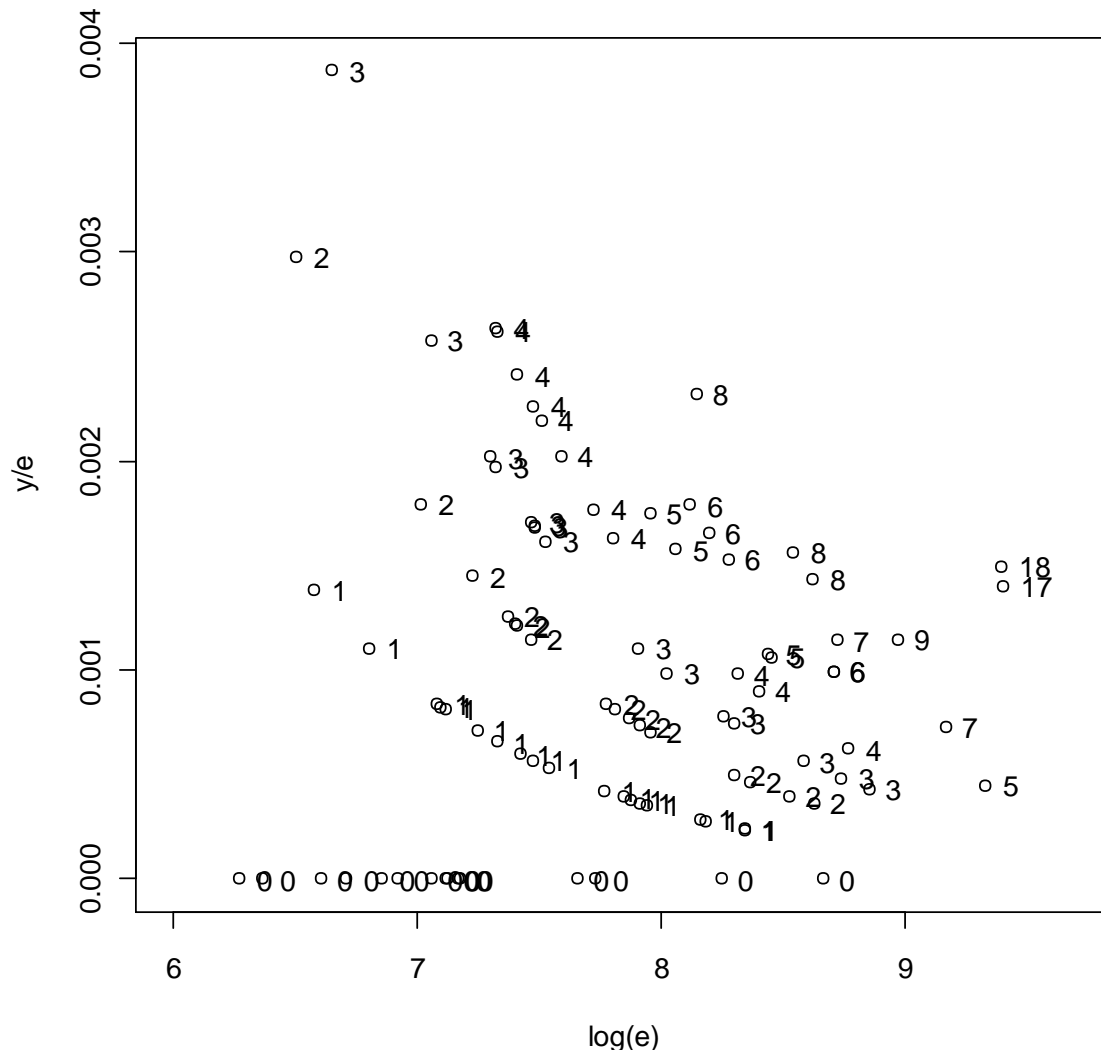
Hierarchical Modeling

- Heart transplant mortality rates (section 7.3)

- ▶ Goal: learn about the rate of success of heart transplant surgery of a particular hospital in the United States
- ▶ Prior: $\lambda \sim \text{gamma}(\alpha, \beta)$
 - ▶ Based on pilot data from 10 hospitals: $(\alpha, \beta) = (16, 15174)$
- ▶ Data: $y_i \sim \text{Poisson}(e_i \lambda_i), i = 1, \dots, 94$
 - ▶ e_i : exposure
 - ▶ λ_i : mortality rate per unit exposure
- ▶ Separate estimates, pooled estimate, or combined?

Death rates vs log(exposure) for all hospitals

Each point is labeled by the number of observed deaths.



- Death rate = y_i/e_i
- The estimated rates are highly variable, especially for programs with small exposures.
- The programs experiencing no deaths (a plotting label of 0) also are primarily associated with small exposures.

Pooled estimate (section 7.4)

- ▶ Assume equal mortality rates at all hospitals
- ▶ Use a noninformative prior

$$g(\lambda) \propto \frac{1}{\lambda}.$$

- ▶ Posterior: $gamma(\sum_j y_j, \sum_j e_j)$

$$\begin{aligned} g(\lambda|\text{data}) &\propto \frac{1}{\lambda} \prod_{j=1}^{94} \left[\lambda^{y_j} \exp(-e_j \lambda) \right] \\ &= \lambda^{\sum_{j=1}^{94} y_j - 1} \exp \left(- \sum_{j=1}^{94} e_j \lambda \right) \end{aligned}$$

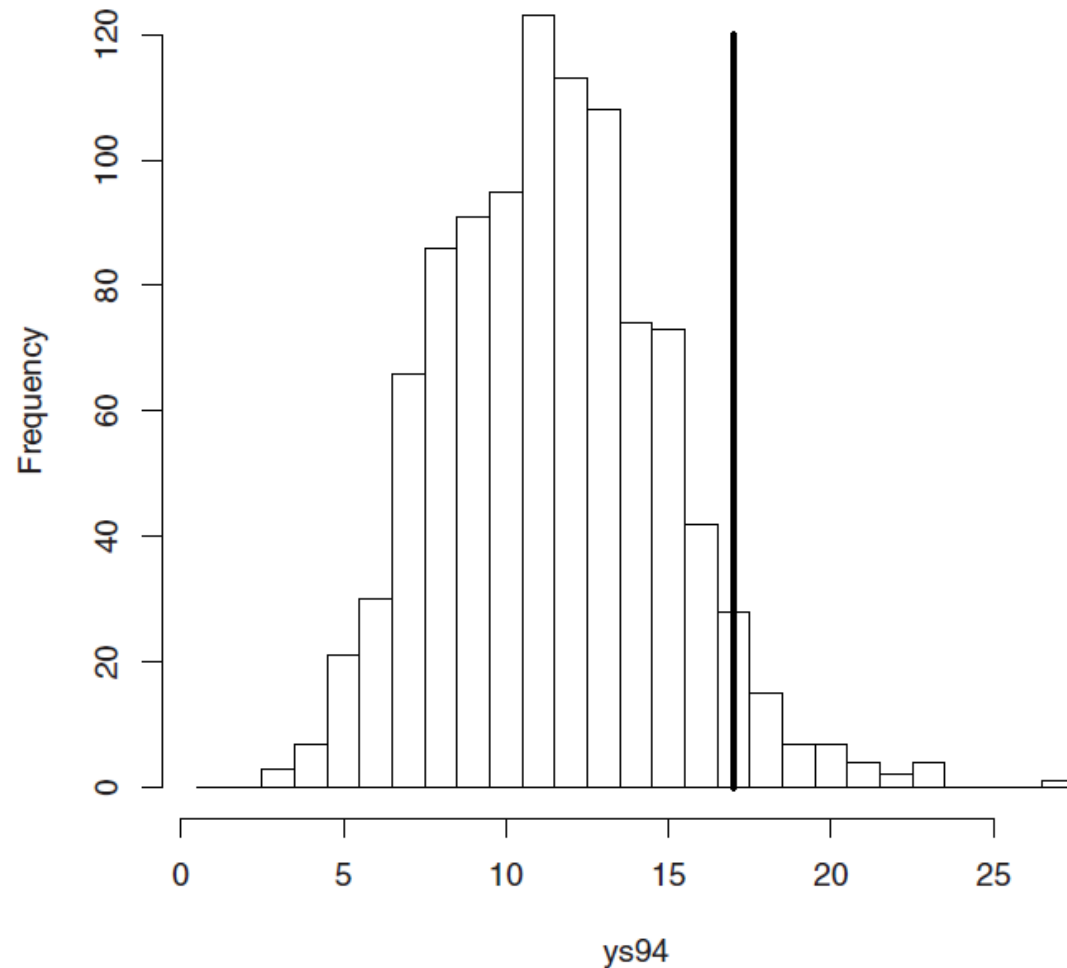
Pooled estimate

- ▶ Check model fit using predictive distribution

$$f(y_i^*|e_i, y) = \int f_P(y_i^*|e_i\lambda)g(\lambda|y)d\lambda,$$

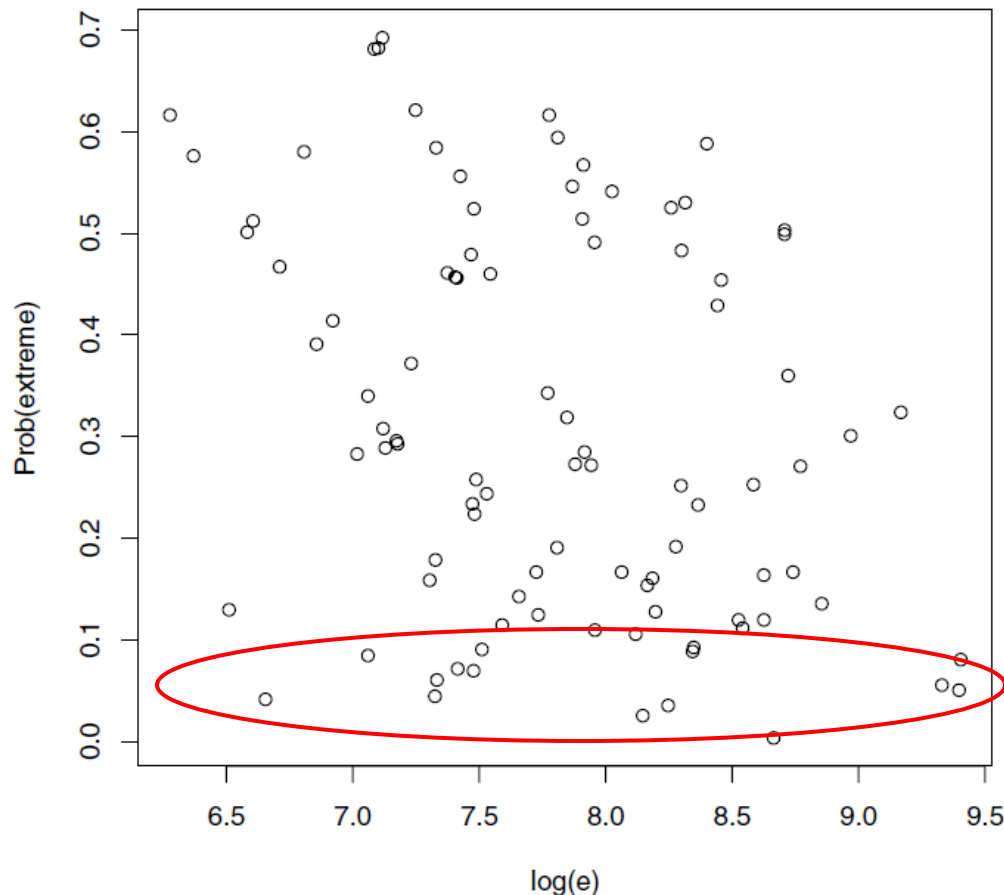
- ▶ If the actual number of observed deaths y_i is in the middle of this predictive distribution, then we can say that our observation is consistent with our model fit.
- ▶ On the other hand, if the observed y_i is in the extreme tails of the distribution $f(y_i^*|e_i, y)$, then this observation indicates that the model is inadequate in fitting this observation.

Hospital 94, $y_{94} = 17$



Probability that the future observation y_i^* is at least as extreme as y_i for all hospitals

$$\min\{P(y_i^* \leq y_i), P(y_i^* \geq y_i)\}.$$



Note that a number of these tail probabilities appear small (15 are smaller than 0.10), which means that the “equal-rates” model is inadequate for explaining the distribution of mortality rates for the group of 94 hospitals.

Hierarchical model (section 7.5)

- ▶ $y_i \sim \text{Poisson}(e_i \lambda_i), i = 1, \dots, 94$
- ▶ Exchangeability $\rightarrow \lambda_1, \dots, \lambda_{94} \sim \text{gamma}\left(\alpha, \frac{\alpha}{\mu}\right)$

$$f(\lambda|\alpha, \mu) = \frac{\left(\frac{\alpha}{\mu}\right)^\alpha \lambda^{\alpha-1} \exp(-\alpha\lambda/\mu)}{\Gamma(\alpha)}$$

- ▶ Prior mean: μ
- ▶ Prior variance: $\frac{\mu^2}{\alpha}$
- ▶ Hyper-prior:
 - ▶ $\mu \sim \text{Inv-gamma}(a, b)$
 - ▶ $\alpha \sim g(\alpha)$
 - ▶ μ and α are independent
- ▶ This model induces positive correlation between λ_i
 - ▶ See Albert (2009)

Practical setup

▶ Noninformative prior

$$g(\mu) \propto \frac{1}{\mu}, \mu > 0.$$

$$g(\alpha) = \frac{z_0}{(\alpha + z_0)^2}, \alpha > 0.$$

This is a relatively flat distribution.



- ▶ z_0 is the prior median of α . We set $z_0 = 0.53$.

▶ Posterior

- ▶ $\lambda_i | \mu, \alpha, y \sim \text{gamma}(y_i + \alpha, e_i + \frac{\alpha}{\mu})$
- ▶ By integrating the joint posterior $f(\alpha, \mu, \lambda_1, \dots, \lambda_{94} | y)$ w.r.t. λ_i

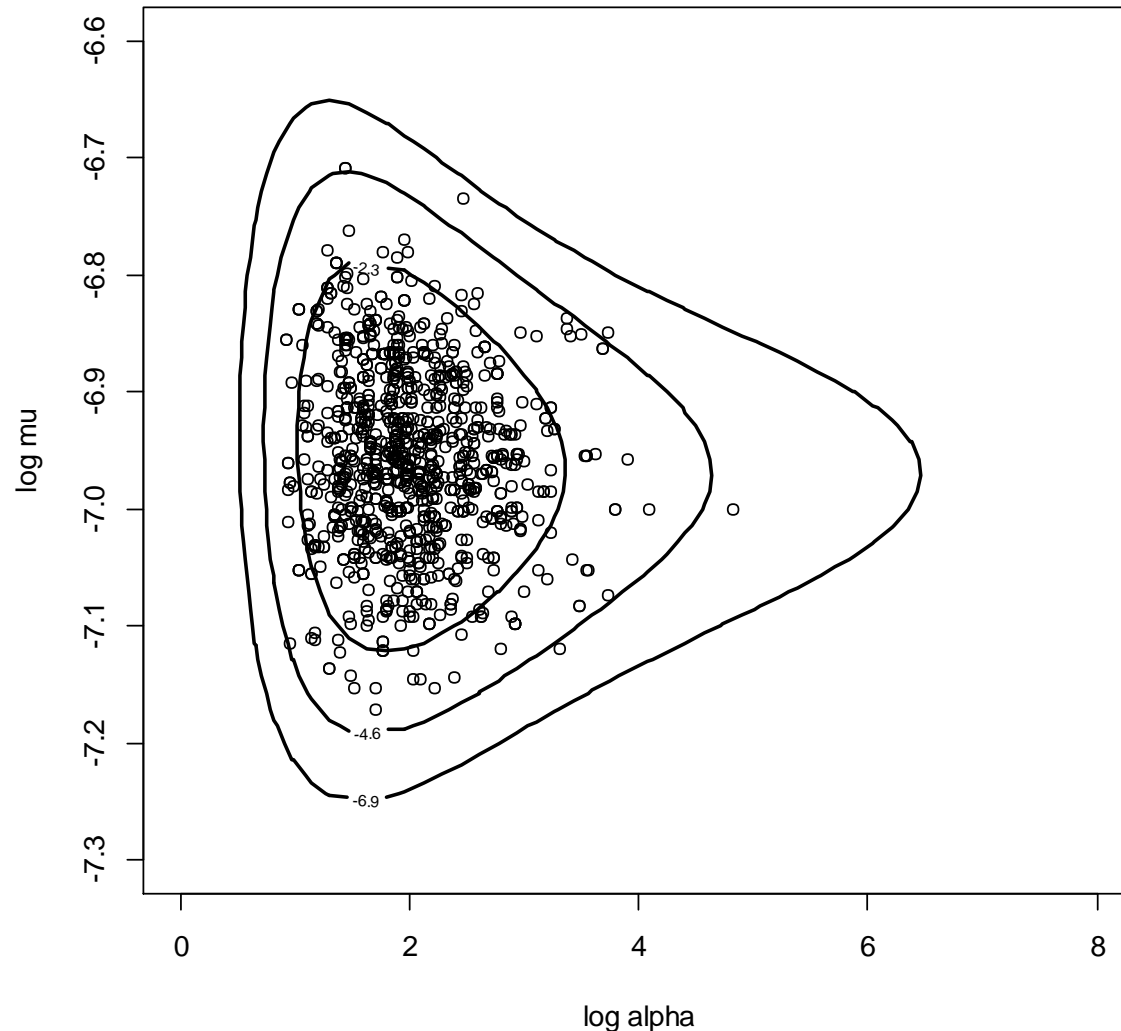
$$p(\alpha, \mu | \text{data}) = K \frac{1}{\Gamma^{94}(\alpha)} \prod_{j=1}^{94} \left[\frac{(\alpha/\mu)^\alpha \Gamma(\alpha + y_i)}{(\alpha/\mu + e_i)^{(\alpha + y_i)}} \right] \frac{z_0}{(\alpha + z_0)^2} \frac{1}{\mu},$$

Simulating from the posterior distribution (section 7.7)

► Simulation

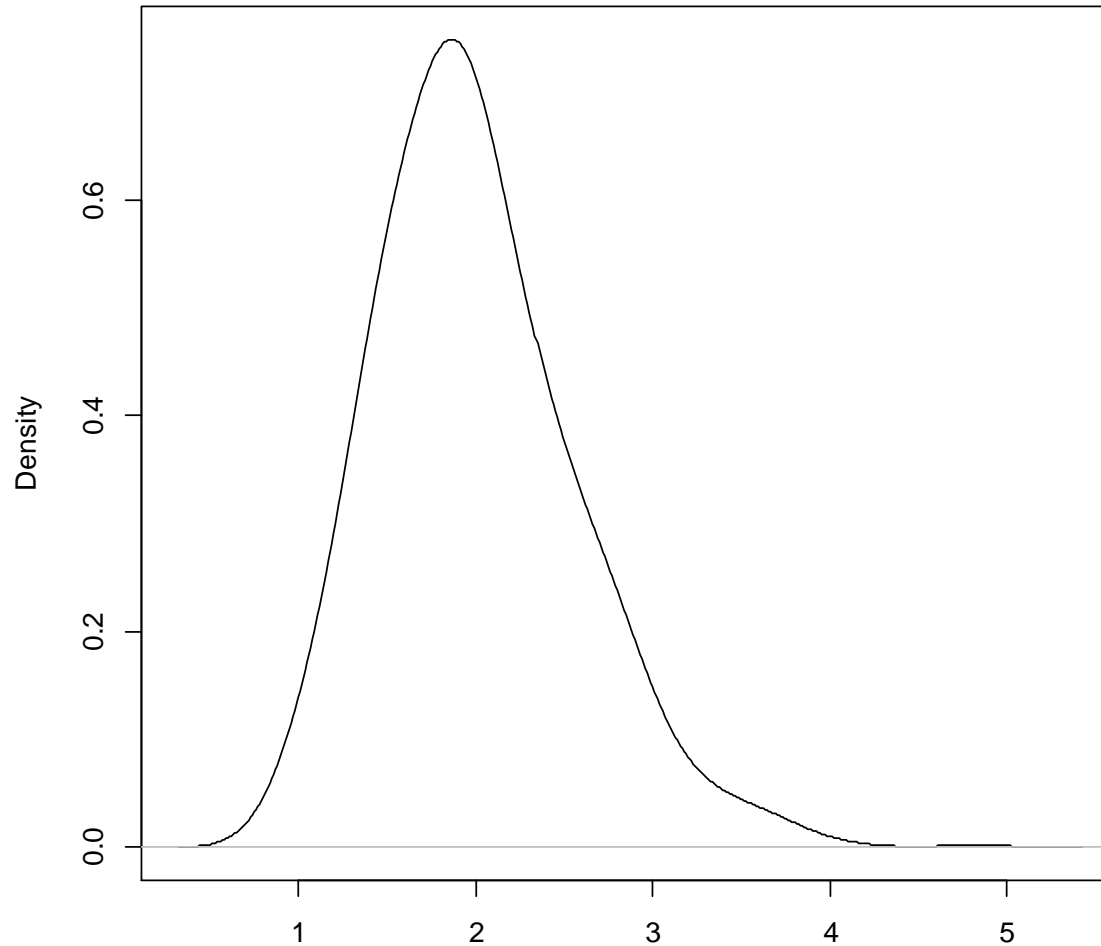
- Simulate from $p(\mu, \alpha | \text{data})$
- Simulate from $\lambda_i | \mu, \alpha, y \sim \text{gamma}(y_i + \alpha, e_i + \frac{\alpha}{\mu})$
- Since both μ and α are positive, consider parameter transformation: Let $\theta_1 = \log(\alpha)$, $\theta_2 = \log(\mu)$
- $$p(\theta_1, \theta_2 | \text{data}) = K \frac{1}{\Gamma^{94}(\alpha)} \prod_{i=1}^{94} \left[\frac{\left(\frac{\alpha}{\mu}\right)^\alpha \Gamma(\alpha + y_i)}{\left(\frac{\alpha}{\mu} + e_i\right)^{\alpha + y_i}} \right] \frac{z_0 \alpha}{(\alpha + z_0)^2}$$
- $\log p(\theta_1, \theta_2 | \text{data})$ is defined by R function `poissgamexch` in the package `LearnBayes`
- A Metropolis-within-Gibbs algorithm is used to sample from this distribution using the function `gibbs`

Contour plot of the posterior density of $(\log \alpha, \log \mu)$ with simulate values imposed



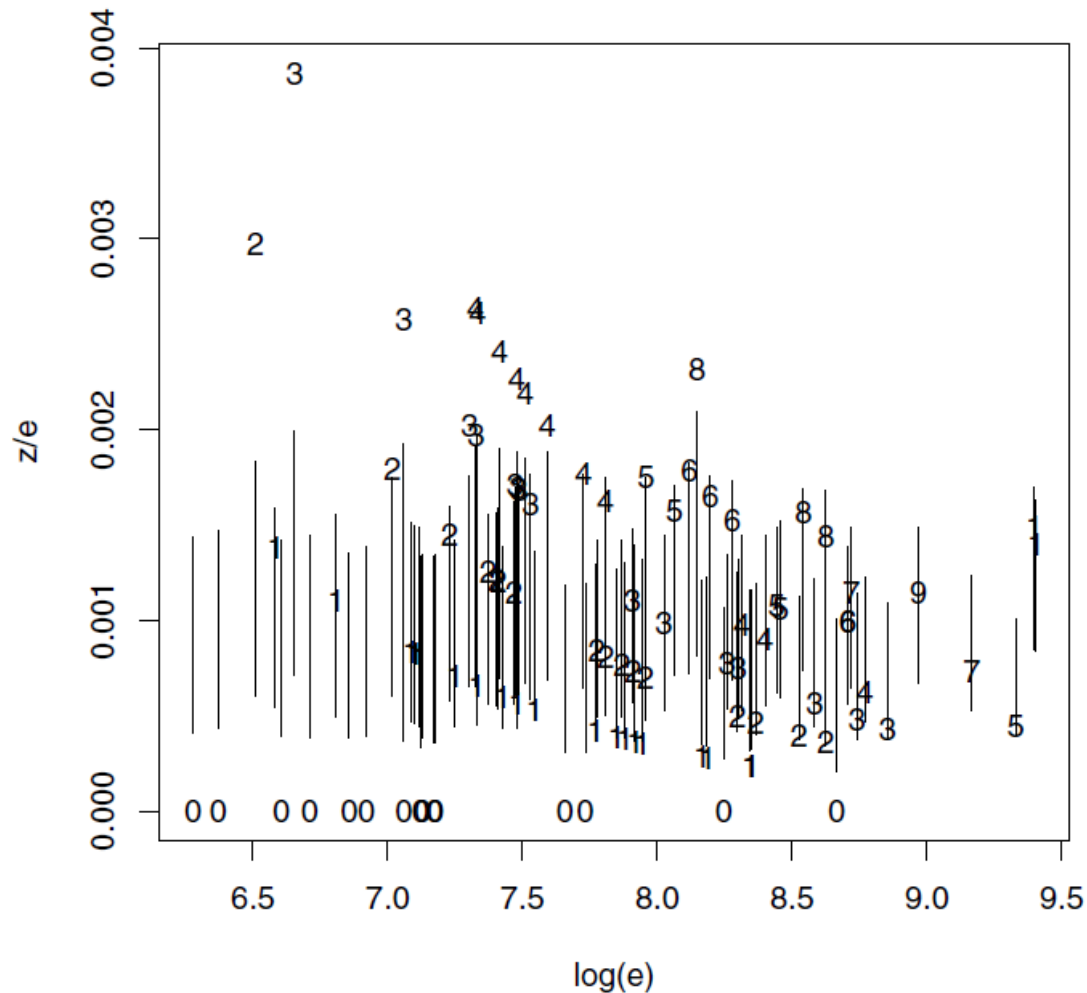
Marginal posterior density of $\log\alpha$

`density.default(x = fitgibbs$par[, 1], bw = 0.2)`



N = 1000 Bandwidth = 0.2

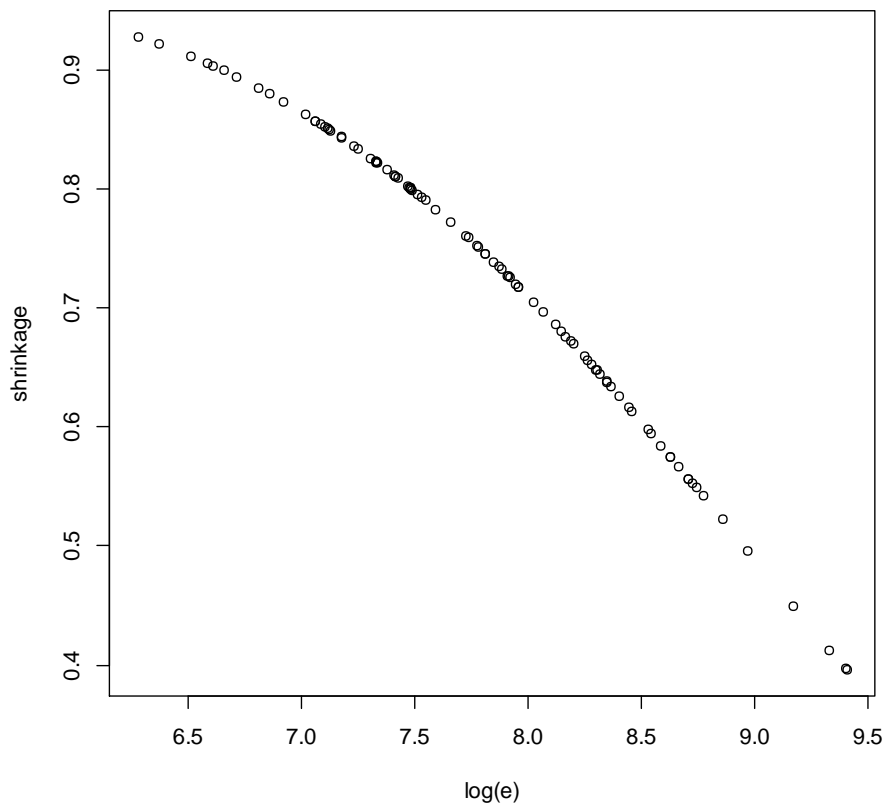
Plot of observed death rates against $\log(\text{exposure})$ together with intervals representing 90% posterior probability bands for the true rates $\{\lambda_i\}$.



Shrinkage effect

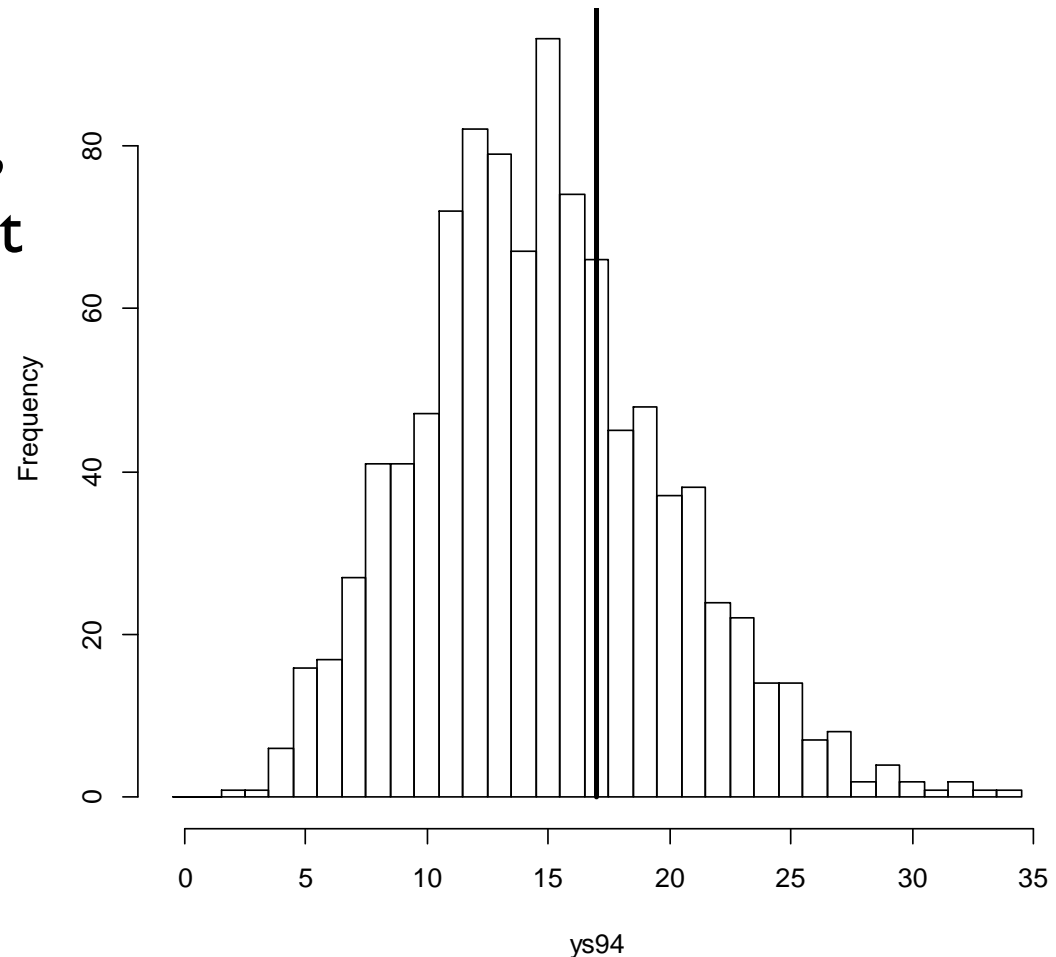
plot of posterior shrinkage vs log(exposure)

- ▶ $E(\lambda_i|data) \approx \left(1 - E(B_i|data)\right) \frac{y_i}{e_i} + E(B_i|data) \frac{\sum_{j=1}^{94} y_j}{\sum_{j=1}^{94} e_j}$
- ▶ Size of shrinkage: $B_i = \alpha/(\alpha + e_i\mu)$



Posterior Predictive Model Checking (section 7.10)

- ▶ $p(y_{94}^*|y)$
- ▶ The observed value of $y_{94} = 17$ is in the middle, which indicates agreement of this observation with the fitted model.

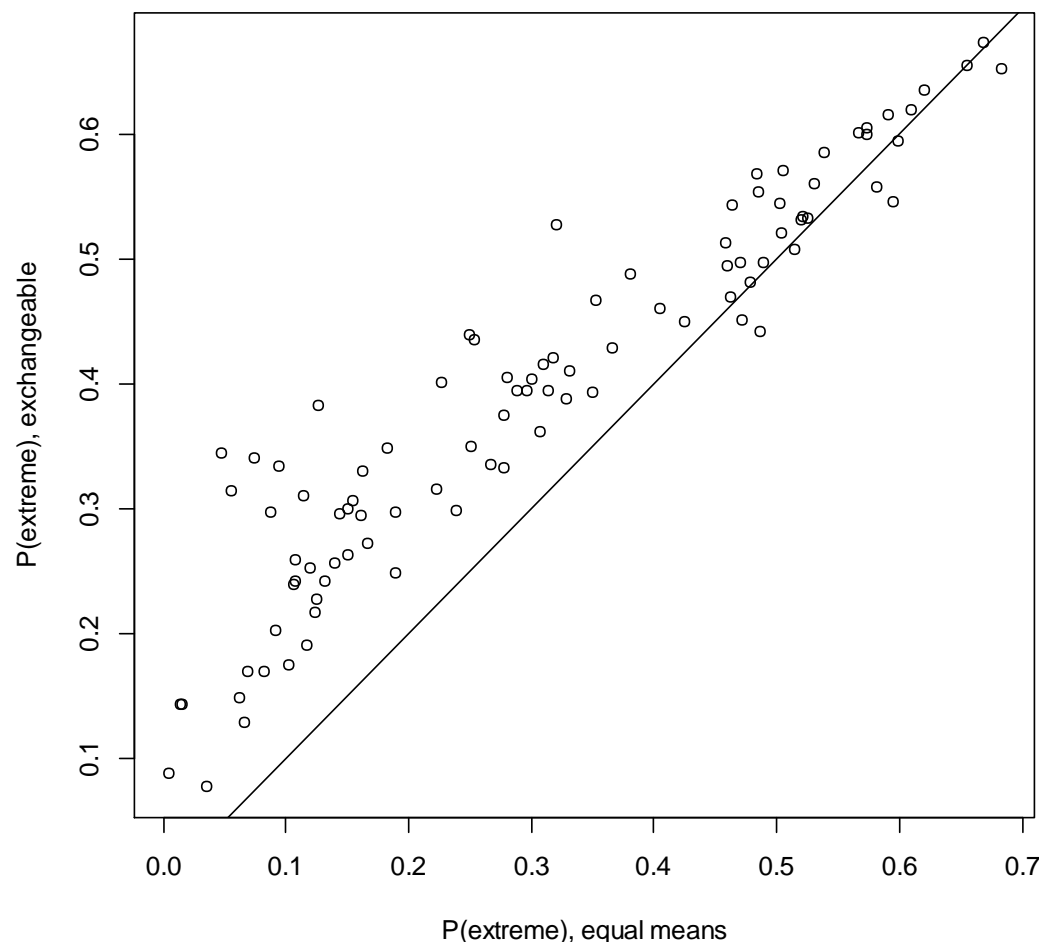


Posterior Predictive Model Checking

- ▶ Perform this for all $y_i, i = 1, \dots, 94$ and in each case calculate the posterior predictive pvalue (ppp)

$$\min\{P(y_i^* \leq y_i), P(y_i^* \geq y_i)\}.$$

- ▶ The plot on the right shows a comparison between the ppp from the hierarchical model and the simple equal-rate model



Basic analysis of MCMC output (section 6.8)

- ▶ R package `coda`
- ▶ To use various diagnostic and summary tools in `coda`, one first needs to convert the MCMC output using the function `mcmc`
 - ▶ You may perform thinning when using `mcmc()`
 - ▶ It is usually necessary to discard a burn-in period
- ▶ Graphics
 - ▶ Trace plot (can use any plotting function in R)
 - ▶ Autocorrelation (`autocorr.plot`)
- ▶ Basic numeric summary
 - ▶ `summary.mcmc()`

Exercise

- ▶ Recall that we used a Metropolis-within-Gibbs algorithm to simulate from $p(\theta_1, \theta_2 | \text{data})$ in the hospital example

```
library(LearnBayes)
data(hearttransplants)
attach(hearttransplants)

datapar = list(data = hearttransplants, z0 = 0.53)
start = c(4, -7)
fitgibbs = gibbs(poissgamexch, start, 1000, c(1,.15), datapar)
```

- ▶ Perform a basic analysis to the MCMC output

```
xyplot(mcmc(fitgibbs$par))  
autocorr.plot(mcmc(fitgibbs$par))  
summary(mcmc(fitgibbs$par))  
batchSE(mcmc(fitgibbs$par))
```