

Chapter 10.3 Hierarchical Beta-Binomial Modeling

Jim Albert and Monika Hu

Chapter 10 Bayesian Hierarchical Modeling

Example: Deaths after heart attack

- ▶ The New York State (NYS) Department of Health
- ▶ Data on mortality after Acute Myocardial Infarction (AMI), commonly known as a heart attack
- ▶ The 2015 report: 13 hospitals in Manhattan, New York City

Example: Deaths after heart attack cont'd

Hospital	Cases	Deaths	Death %
Bellevue Hospital Ctr.	129	4	3.101
Harlem Hospital Ctr.	35	1	2.857
Lenox Hill Hospital	228	18	7.894
Metropolitan Hospital Ctr.	84	7	8.333
Mount Sinai Beth Israel	291	24	8.247
Mount Sinai Hospital	270	16	5.926
Mount Sinai Roosevelt	46	6	13.043
Mount Sinai St. Luke's	293	19	6.485
NYU Hospitals Center	241	15	6.224
NYP Hospital - Allen Hospital	105	13	12.381
NYP Hospital - Columbia Presbyterian Ctr.	353	25	7.082
NYP Hospital - New York Weill Cornell Ctr.	250	11	4.400
NYP/Lower Manhattan Hospital	41	4	9.756

A hierarchical Beta-Binomial model

- ▶ Treat “cases” as trials and “deaths” as successes: the Binomial sampling model
- ▶ The objective is to learn about the death probability p of the hospitals
- ▶ Data shows that some hospitals have much higher death rates than other hospitals
- ▶ Separate estimates? Combined estimates?

A hierarchical Beta-Binomial model

- ▶ Treat “cases” as trials and “deaths” as successes: the Binomial sampling model
- ▶ The objective is to learn about the death probability p of the hospitals
- ▶ Data shows that some hospitals have much higher death rates than other hospitals
- ▶ Separate estimates? Combined estimates?
- ▶ A hierarchical model!

Model specification

- ▶ Y_i : the number of resulted deaths from heart attack
- ▶ n_i : the number of heart attack cases
- ▶ p_i : the death rate for hospital i

Model specification

- ▶ Y_i : the number of resulted deaths from heart attack
- ▶ n_i : the number of heart attack cases
- ▶ p_i : the death rate for hospital i
- ▶ Sampling, for $i, \dots, 13$:

$$Y_i \sim \text{Binomial}(n_i, p_i) \quad (1)$$

- ▶ Prior for p_i , $i = 1, \dots, 13$:

$$p_i \sim \text{Beta}(a, b) \quad (2)$$

- ▶ Review:

$$p_i \mid y_i \sim \text{Beta}(a + y_i, b + n_i - y_i) \quad (3)$$

Model specification cont'd

- ▶ Sampling, for $i, \dots, 13$:

$$Y_i \sim \text{Binomial}(n_i, p_i) \quad (4)$$

- ▶ Prior for p_i , Stage1: for $i = 1, \dots, 13$:

$$p_i \sim \text{Beta}(a, b) \quad (5)$$

- ▶ Prior for p_i , Stage 2: the hyperprior:

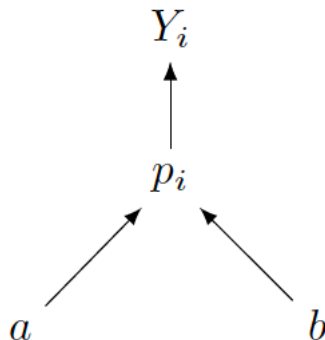
$$a, b \sim \pi(a, b) \quad (6)$$

Graphical representations of the hierarchical model

$$Y_i \sim \text{Binomial}(n_i, p_i)$$

$$p_i \sim \text{Beta}(a, b)$$

$$a, b \sim \pi(a, b)$$

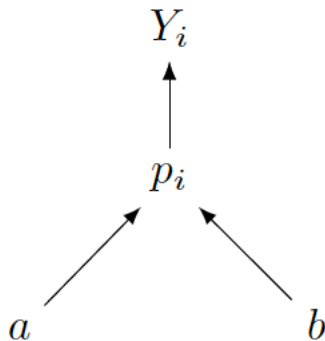


Another hierarchical model

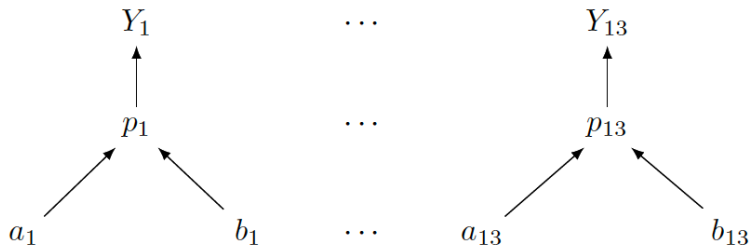
- If the beta parameters a and b are known constants

$$Y_i \sim \text{Binomial}(n_i, p_i)$$

$$p_i \sim \text{Beta}(a, b)$$



Separate estimates model



Beta prior revisit

- ▶ Review from Chapter 7
 - ▶ `beta.select()` function for choosing a and b
 - ▶ The parameter a as the prior count of “successes”
 - ▶ The parameter b as the prior count of “failures”
 - ▶ The sum $a + b$ represents the prior sample size
 - ▶ The expectation of $\text{Beta}(a, b)$ is $\frac{a}{a+b}$

Beta prior revisit

- ▶ Review from Chapter 7
 - ▶ `beta.select()` function for choosing a and b
 - ▶ The parameter a as the prior count of “successes”
 - ▶ The parameter b as the prior count of “failures”
 - ▶ The sum $a + b$ represents the prior sample size
 - ▶ The expectation of $\text{Beta}(a, b)$ is $\frac{a}{a+b}$
- ▶ A natural parameterization
 - ▶ $\pi(\mu, \eta)$
 - ▶ $\mu = \frac{a}{a+b}$: mean
 - ▶ $\eta = a + b$: (prior) sample size

$$\mu, \eta \sim \pi(\mu, \eta) \tag{7}$$

where $a = \mu\eta$ and $b = (1 - \mu)\eta$

Second-stage prior

- ▶ Assume μ and η are independent
- ▶ The hyperprior expectation μ is the mean measure for p_i , the average death rate across 13 hospitals
 - ▶ if we have little prior knowledge about the expectation μ , we can assign a uniform prior (Beta(1, 1))

Second-stage prior

- ▶ Assume μ and η are independent
- ▶ The hyperprior expectation μ is the mean measure for p_i , the average death rate across 13 hospitals
 - ▶ if we have little prior knowledge about the expectation μ , we can assign a uniform prior ($\text{Beta}(1, 1)$)
- ▶ Consider known hyperparameter values

$$E(p^* \mid y^*) = \frac{y^* + \mu\eta}{n^* + \eta} \quad (8)$$

Second-stage prior cont'd

- ▶ The posterior mean is rewritten as

$$E(p^* | y^*) = (1 - \lambda) \frac{y^*}{n^*} + \lambda \mu \quad (9)$$

- ▶ λ is the shrinkage fraction $\lambda = \frac{\eta}{n^* + \eta}$
- ▶ λ falls in the interval $(0, 1)$ and represents the degree of shrinkage of the posterior mean away from the sample proportion y^*/n^* towards the prior mean μ

Second-stage prior cont'd

- ▶ Suppose we believe *a priori* that, for a representative sample size n^* , the shrinkage λ is Uniform(0, 1)
- ▶ The prior density for the prior sample size η has the form

$$\pi(\eta) = \frac{n^*}{(n^* + \eta)^2}, \quad \eta > 0 \quad (10)$$

- ▶ The logarithm of η , $\theta = \log \eta$, has a Logistic distribution with location $\log n^*$ and scale 1

$$\pi(\theta) = \frac{e^{-(\theta - \log n^*)}}{(1 + e^{-(\theta - \log n^*)})^2} \quad (11)$$

Model specification

- ▶ Sampling, for $i, \dots, 13$:

$$Y_i \sim \text{Binomial}(n_i, p_i) \quad (12)$$

- ▶ Prior for p_i , Stage 1: for $i = 1, \dots, 13$:

$$p_i \sim \text{Beta}(a, b) \quad (13)$$

- ▶ Prior for p_i , Stage 2:

$$\mu \sim \text{Beta}(1, 1), \quad (14)$$

$$\log \eta \sim \text{Logistic}(\log n^*, 1) \quad (15)$$

where $a = \mu\eta$ and $b = (1 - \mu)\eta$

JAGS step 1: describe the model by a script

```
modelString <-"
model {
  ## likelihood
  for (i in 1:N){
    y[i] ~ dbin(p[i], n[i])
  }
  ## priors and hyperpriors
  for (i in 1:N){
    p[i] ~ dbeta(a, b)
  }
  a <- mu*eta
  b <- (1-mu)*eta
  mu ~ dbeta(mua, mub)
  eta <- exp(logeta)
  logeta ~ dlogis(logn, 1)
}
"
```

JAGS step 2: define the data and prior parameters

```
deathdata <- read.csv("DeathDataset.csv", header=T)
deathdata$DeathRatio <- deathdata$Deaths /
  deathdata$Cases*100

y <- deathdata$Deaths
n <- deathdata$Cases
N <- length(y)
the_data <- list("y" = y, "n" = n, "N" = N,
                 "mua" = 1, "mub" = 1,
                 "logn" = log(100))
```

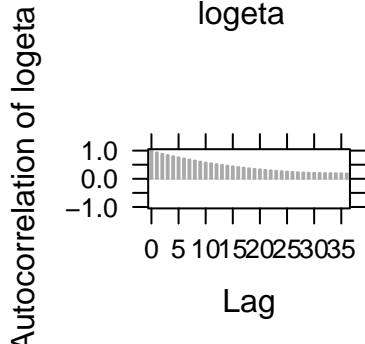
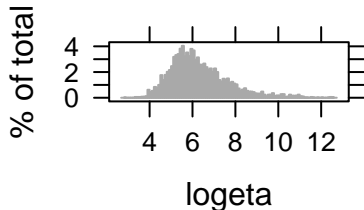
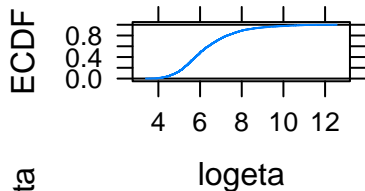
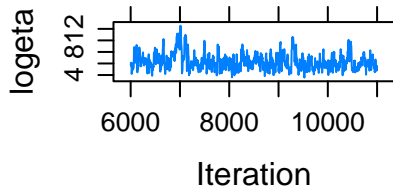
JAGS step 2: define the data and prior parameters cont'd

```
posterior <- run.jags(modelString,  
                      n.chains = 1,  
                      data = the_data,  
                      monitor = c("p", "mu", "logeta"),  
                      adapt = 1000,  
                      burnin = 5000,  
                      sample = 5000)
```

MCMC diagnostics and summarization

```
plot(posterior, vars = "logeta")
```

Generating plots...



MCMC diagnostics and summarization cont'd

##

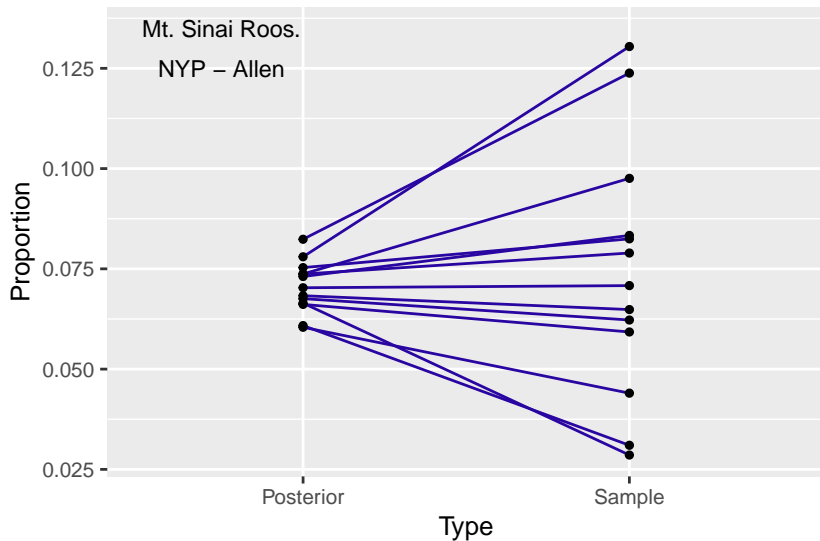
JAGS model summary statistics from 5000 samples (adaptive)

##

##	Lower95	Median	Upper95	Mean	SD	Mode	
## p[1]	0.0348	0.0619	0.0842	0.0609	0.0128	--	0.
## p[2]	0.0351	0.0667	0.0925	0.0665	0.0142	--	0.
## p[3]	0.0521	0.0725	0.0979	0.0737	0.0115	--	0.
## p[4]	0.0485	0.0716	0.104	0.0731	0.0137	--	0.
## p[5]	0.0551	0.074	0.0998	0.0753	0.0114	--	0.
## p[6]	0.0459	0.0659	0.0865	0.0661	0.0103	--	0.
## p[7]	0.0485	0.0747	0.114	0.078	0.0169	--	0.
## p[8]	0.0488	0.068	0.0886	0.0683	0.01	--	0.
## p[9]	0.0466	0.0675	0.0883	0.0676	0.0107	--	0.
## p[10]	0.0545	0.0786	0.117	0.0824	0.0168	--	0.
## p[11]	0.0506	0.0701	0.0888	0.0703	0.00963	--	0.
## p[12]	0.038	0.0611	0.0809	0.0604	0.0111	--	0.
## p[13]	0.0443	0.0715	0.105	0.0737	0.0153	--	0.

Inferences

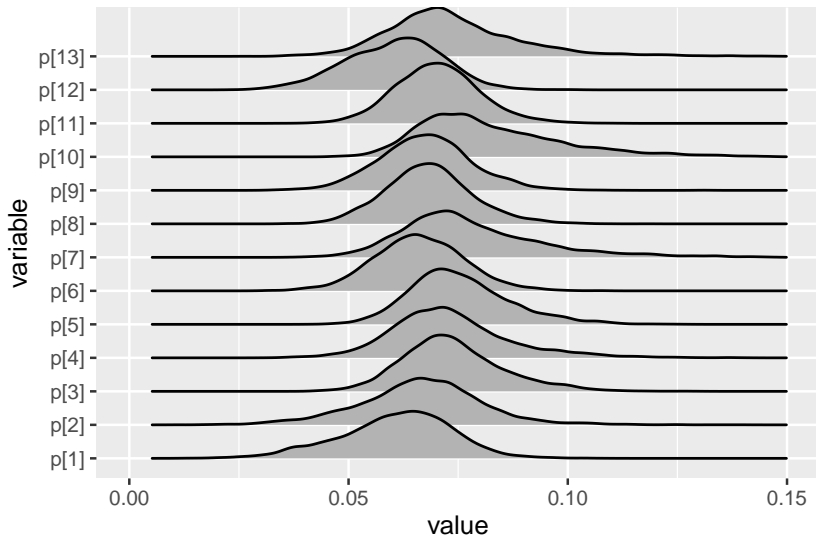
► Pooling / shrinkage effects



Inferences

► Posterior densities

Picking joint bandwidth of 0.00189



Inferences

- ▶ Comparison of hospitals
 - ▶ e.g. compare the death rates of two hospitals directly

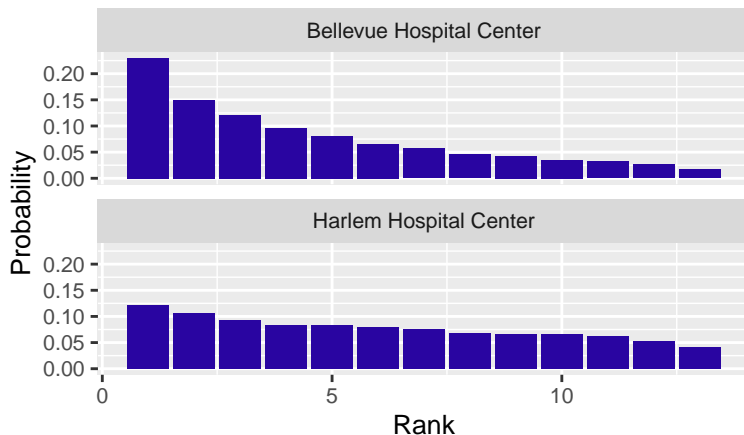
```
p11draws <- as.mcmc(posterior, vars = "p[11]")  
p12draws <- as.mcmc(posterior, vars = "p[12]")  
diff = p11draws - p12draws  
sum(diff > 0)/5000
```

```
## [1] 0.7572
```

Inferences

- ▶ Comparison of hospitals
 - ▶ e.g. ranking

$$P(p_1 < p_2, \dots, p_1 < p_{13} \mid y) \quad (16)$$



Inferences

- ▶ Comparison of hospitals
 - ▶ e.g. probability of rank 1

