Denis Ostroushko - PUBH 7440 - HW4 - Part 2

Problem 1

Full hierarchical model:

$$
\begin{aligned}
p(\theta_{ir}, \beta_{0r}, \sigma_r^2 | y_{ir}, n_{ir}) &\propto \prod_{i=1}^{N=67} \prod_{r=1}^{R=2} Bin(n_{ir}, \pi_{ir}) \times Norm(\beta_{0r}, \sigma_r^2) \times Norm(0, \tau^2) \times IG(0.001, 0.001) \\
&= \prod_{i=1}^{N=67} \prod_{r=1}^{R=2} Bin(n_{ir}, \frac{exp(\theta_{ir})}{1 + exp(\theta_{ir})}_{ir}) \times Norm(\beta_{0r}, \sigma_r^2) \times Norm(0, \tau^2) \times IG(0.001, 0.001)
\end{aligned}
$$

where $\tau^2 = 10,000$.

Since π_{ir} and θ_{ir} have a deterministic relationship, and θ_{ir} is the random variable, we placed θ_{ir} into the full conditional model.

Problem 2

Full conditional for β_{0r}

We will have a total of 2 full conditional models for each $\beta_{0r}, r = 1, 2$. For each β_{0r} there are 67 parameters θ_{ir}, and observed values of data y_{ir} and n_{ir}.

Full conditional distribution for β_{0r}:

$$p(\beta_{0r}|.) \propto [\prod_{i=1}^{67} Norm(\theta_{ir}|\beta_{0r}, \sigma_r^2)] \times Norm(\beta_{0r}|0, \tau^2)$$
$$\propto \frac{\beta_{0r}}{\sigma_r^2} \sum_{i=1}^{67} \theta_{ir} - \frac{1}{2}\beta_{0r}^2 (\frac{67}{\sigma^2} + \frac{1}{\tau^2})$$

After more careful consideration, I worked out that this is actually a kernel of a normal distribution. Parameters for this normal distribution are:

$Mean = \frac{\sum_{i=1}^{67} \theta_{ir}}{67+\frac{\sigma_r^2}{\tau^2}}$ and

$Variance = \frac{1}{\frac{67}{\sigma^2}+\frac{1}{\tau^2}}$

We can use Gibbs sampling from a normal distribution to obtain posterior samples for β_{0r}

Full conditional for θ_{ir} **:**

π_{ir} is a function of θ_{ir}, which is a random variable. Therefore, we will obtain a full conditional for θ_{ir} using distributions given in the

For each of \$2 $\times$67 =\$134 parameters θ_{ir} we have a full conditional function. First, let's rewrite distribution of Y_{ir} by using $\frac{exp(\theta_{ir})}{1+exp(\theta_{ir})}$. Then we have:

$$Y_{ir} \propto [\frac{exp(\theta_{ir})}{1 + exp(\theta_{ir})}]^{y_{ir}} \times [\frac{1}{1 + exp(\theta_{ir})}]^{n_{ir}-y_{ir}}$$

Full conditional is given by:

$$p(\theta_{ir}|.) \propto [\frac{exp(\theta_{ir})}{1 + exp(\theta_{ir})}]^{y_{ir}} \times [\frac{1}{1 + exp(\theta_{ir})}]^{n_{ir}-y_{ir}} \times exp(-\frac{1}{2}(\theta_{ir} - \beta_{0r})^2)$$
$$= exp(\theta_{ir})^{y_{ir}} \times [1 + exp(\theta_{ir})]^{-n_{ir}} \times exp(-\frac{1}{2\sigma_r^2}(\theta_{ir} - \beta_{0r})^2)$$

This is not a recognizable kernel, use Metropolis updates with a symmetric normal candidate density to obtain posterior samples.

Full conditional for σ_r^2 **:**

As always, when prior for variance of a normal distribution has inverse gamma, we expect to have an inverse gamma distribution as a posterior.

$$p(\sigma_r^2|.) \propto [\prod_{i=1}^{67} Norm(\theta_{ir}|\beta_{0r}, \sigma_r^2)] \times IG(0.001, 0.001)$$
$$\propto [\sigma_r^2]^{-n/2-0.001-1} \times exp(1/\sigma^2 \sum_i (\theta_{ir} - \beta_{0r})^2 - 0.001/\sigma_r^2)$$

Therefore, full conditional for σ_r^2 is given by $IG(0.001 + n/2, 0.001 + \frac{1}{2}\sum_i(\theta_{ir} - \beta_{0r})^2$

Details with derivations of all full conditionals are attached in the Appendix as hand written notes

Problem 3

I used provided starting values, and 50,000 iterations of Metropolis and Gibbs sampling.

History plots for β_{0r}

Figure 1 and Figure 2 show history values for $\beta_{0r}, r = 1, 2$. While there are no issues with β_{01}, history plot after 1,000 burnin period for β_{02} can be a cause for concern. While the values fluctuate somewhat randomly around the final average value for this parameter, I can see how some Bayesian statisticians might argue that there is a periodic trend in the sampled values. To me, this looks okay, and I would attribute this behavior to the small values of y_{i2} and n_{i2} in the data, with some values being 0 and 0 respectively for a given county.

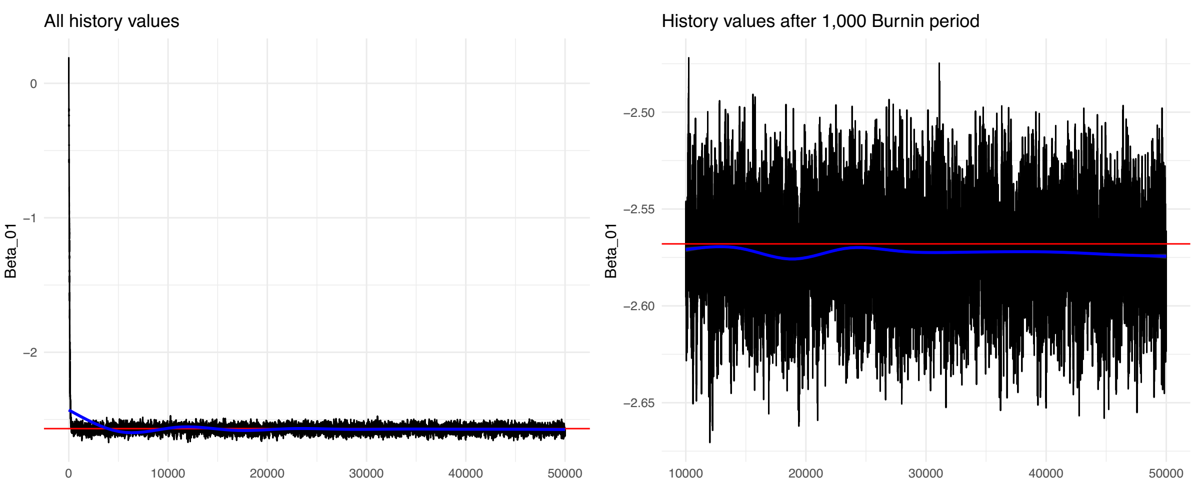

Figure 1: History plots for Beta_01 show reasonable model convergence after 1,000 burnin iterations

History plots for σ_r^2

Figure 3 and Figure 4 shows reasonable convergence. No issues to report here.

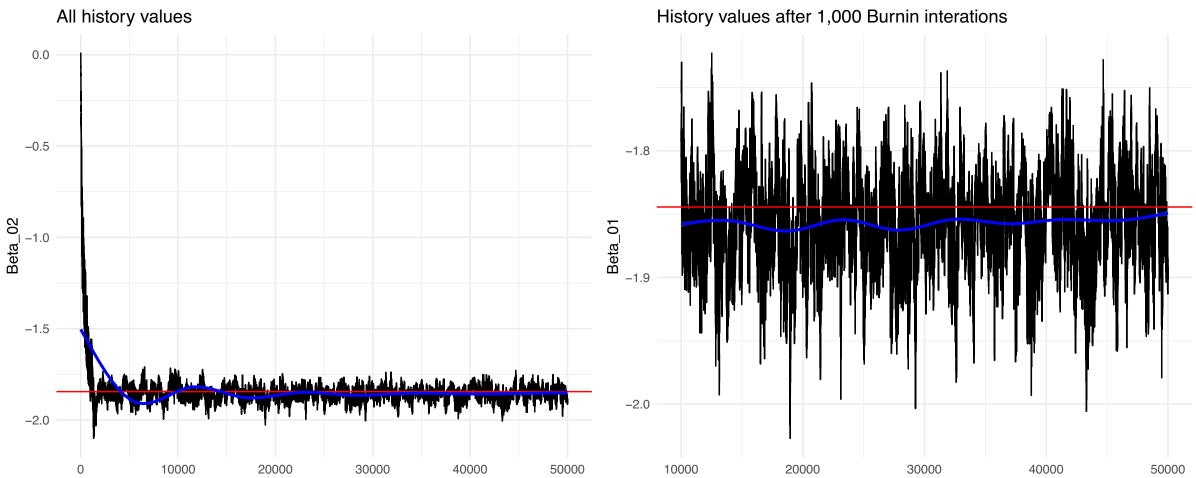

Figure 2: History plots for Beta_02 show mildly acceptable model convergence after 1,000 burnin iterations

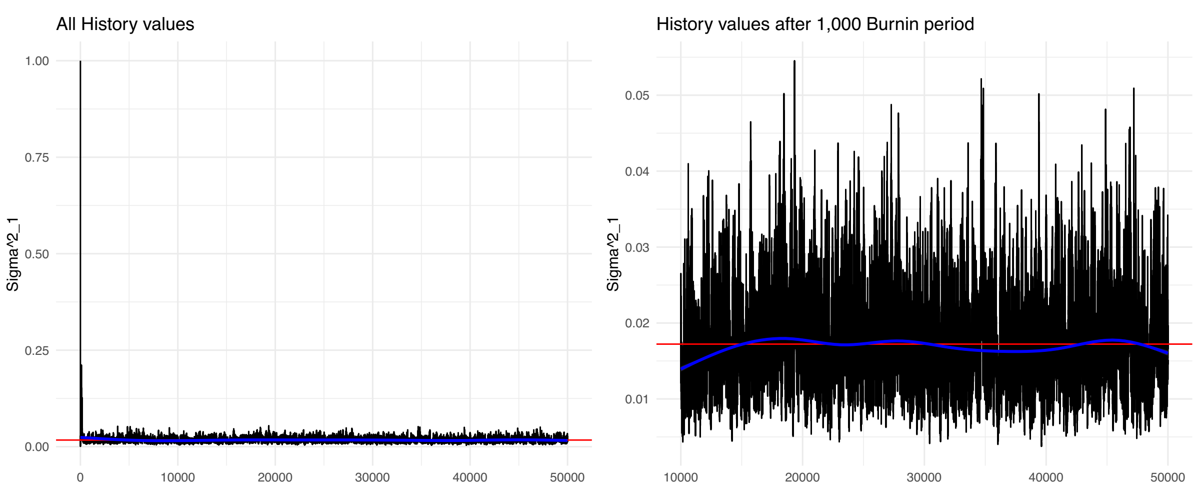

Figure 3: History plots for Sigma^2_1 show reasonable model convergence after 1,000 burnin iterations

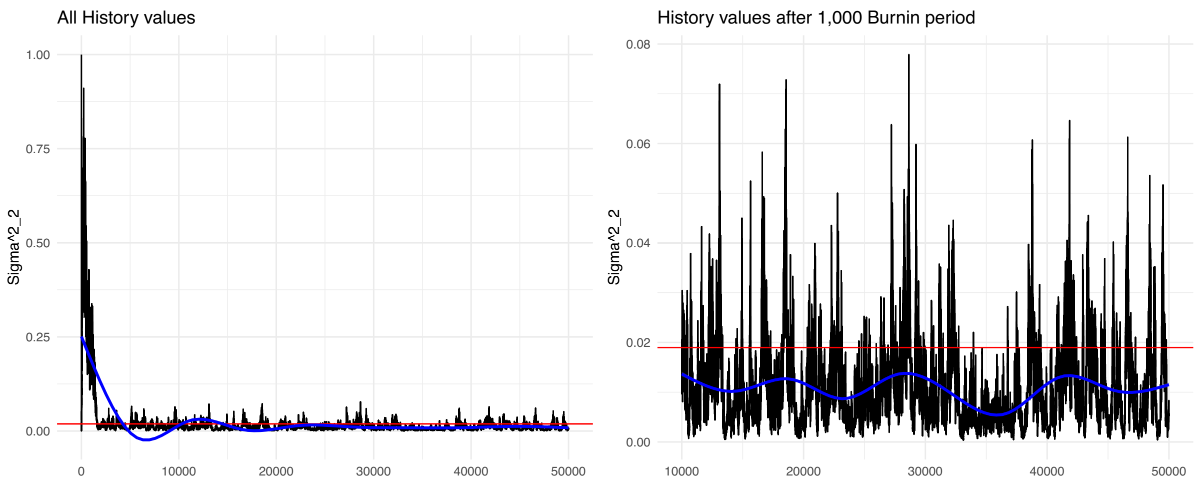

Figure 4: History plots for Sigma^2__2 show reasonable model convergence after 1,000 burnin iterations

Problem 4

Consider the hint, that $E[\theta_{ir}|\gamma, \sigma_r^2] = \gamma_0 + \gamma_1 \times (r-1)$.

based on previous information we also know that $E[\theta_{ir}|\beta_{0r}, \sigma_r^2] = \beta_{0r}$ since θ_{ir} is a normally distributed random variable.

Setting the two sides of these equations we get $\beta_{0r} = \gamma_0 + \gamma_1 \times (r-1)$.

When $r = 1$, $\beta_{01} = \gamma_0$, and when $r = 2$, $\beta_{02} = \gamma_0 + \gamma_1$.

Since θ_{ir} is the log odds using parameter π_{ir}, the log-odds ratio is given by $\beta_{02} - \beta_{01}$.

Figure 5 shows the distribution of the log-odds ratio. 95% credible internal does not include zero, suggesting that the in the state of Pennsylvania black mothers are more likely to give birth to children who are lower in weight.

About 0% of sampled log-odds ratios are below 0, suggesting overwhelming evidence that black mothers are at much higher risk of low-birth-weight events.

Problem 5

Figure 6, Figure 7, Figure 8 show desired maps for rates and ratio of rates at the county level in Pennsylvania. It appears that some counties have a very extreme disparity between Black and White moms. Lowest bucket cutoff was 1.33, implying that at the lower end the rate of low-birthweight events for Black moms is about 33% higher than White moms at the county level.

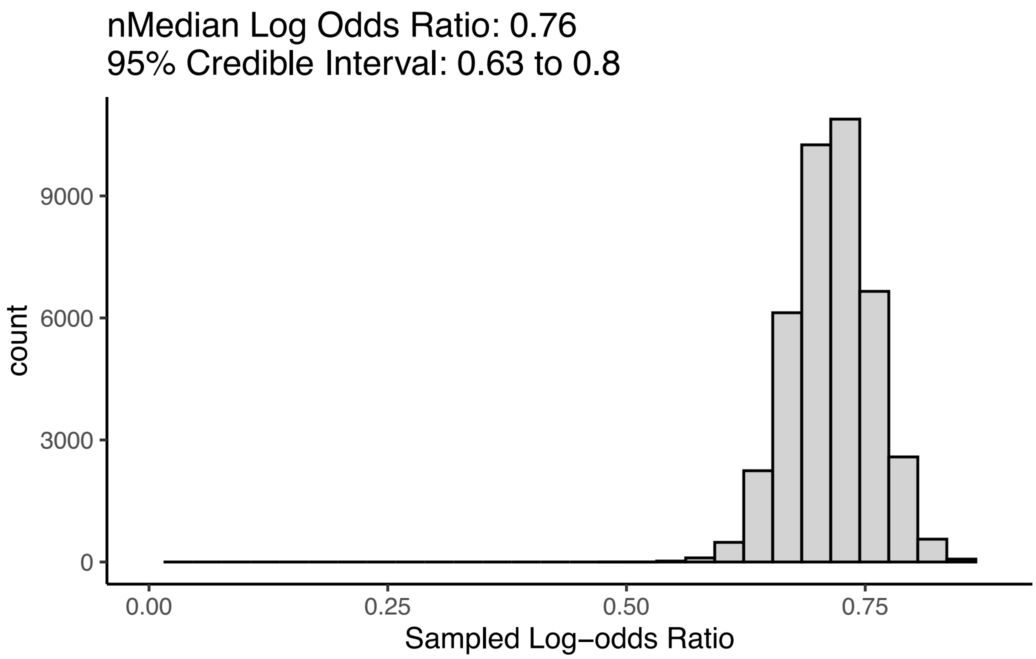

Figure 5: Log-odds ratio distribution suggesting there are racial disparities in terms of proportions of low-wight births

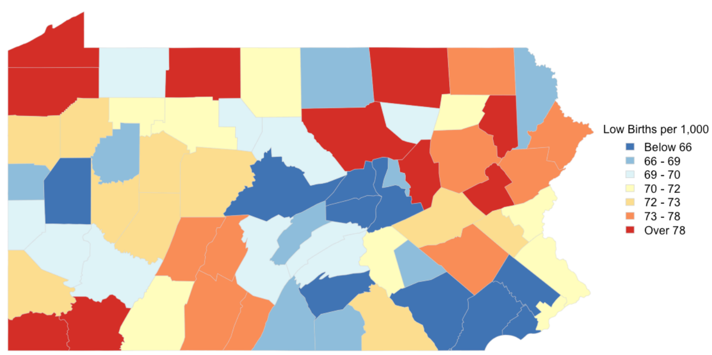

Figure 6: Reginoal rates of low-birthweight events for White moms in PA

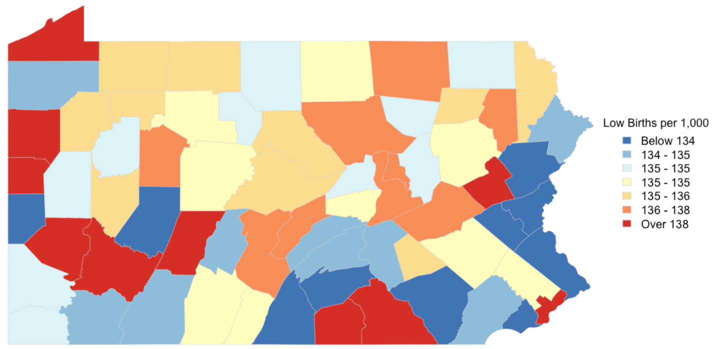

Figure 7: Reginoal rates of low-birthweight events for Black moms in PA

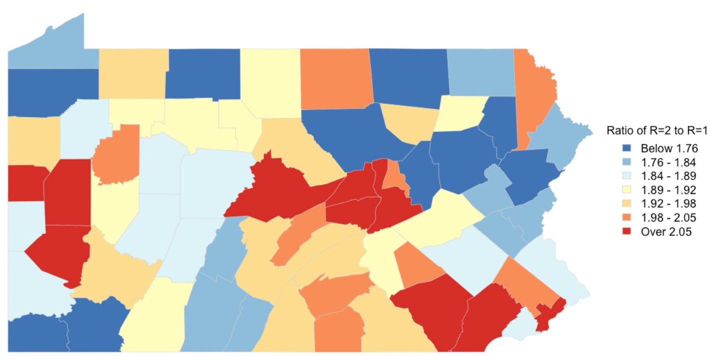

Figure 8: Reginoal ratio of Black to White mom rates in low-birthweight events in PA

Table 1: Summary statistics for Philadelphia county

Racial Group	Y	N	Crude Rate	Estiamted Rate	State-Wide Crude Rate	State-Wide Est. Rate
1	440	6230	0.07	0.08	0.07	0.07
2	1251	8557	0.15	0.14	0.14	0.14

Problem 6

Philadelphia County

Table 1 presents summary statistics from raw data (crude rates) and estimated rates. Figure 9 for the ratio of the estimated rates. It appears that the estimated ratios are slightly higher than those estimated directly from the data. I would be comfortable presenting these results. White estimated and crude ratios are different, it stems from the minor differences between race-wise rates, which appear small on paper.

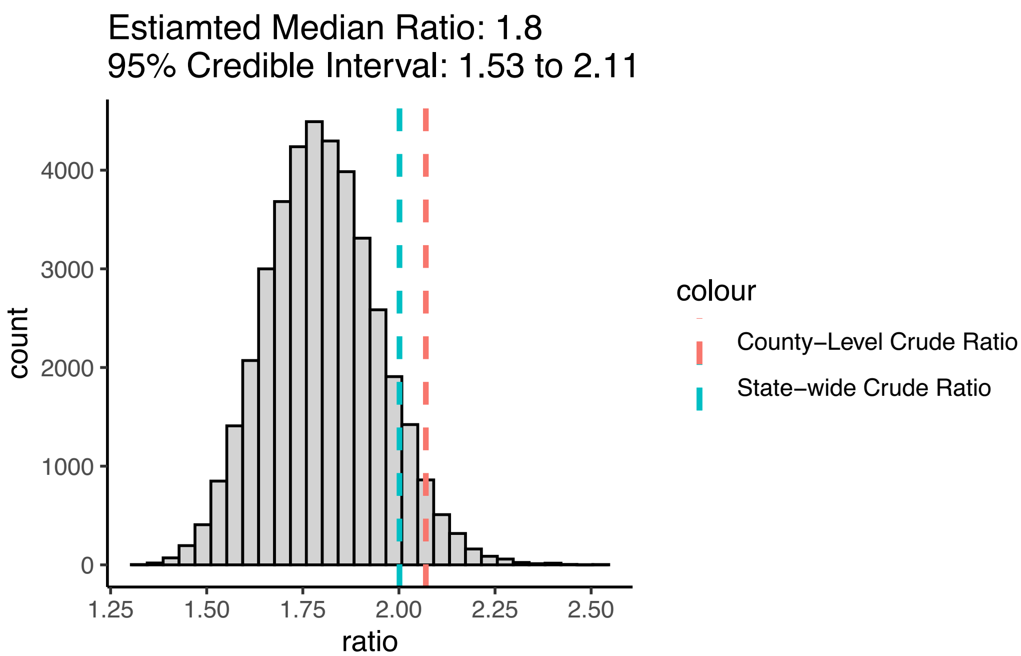

Figure 9: Theoretical distribution of low-birthweight events between Black and White moms

Sullivan County

Table 2 presents summary statistics from raw data (crude rates) and estimated rates. Figure 10 for the ratio of the estimated rates. We can see that there are no events of any birth recorded

Table 2: Summary statistics for Philadelphia county

Racial Group	Y	N	Crude Rate	Estiamted Rate	State-Wide Crude Rate	State-Wide Est. Rate
1	2	40	0.05	0.07	0.07	0.07
2	0	0		0.14	0.14	0.14

for this county for Black moms. So, we have an estimated rate, which is essentially imputed data. We can see that the crude rate is almost half of the estimated rate, and estimated rates are being pulled towards what the state-wide rates are. I would be comfortable presenting these results, but I would also advise that the estimates are based on the small amount of data and therefore future observations can vary greatly from what we have here. This is also supported by the wide 95% credible interval. Additionally, I would advise that future observations and repeated reporting would be subject to highly variable estimates because the distribution in Table 2 shows a heavy tail, implying a chance of high value extreme values, but not outliers by any means.

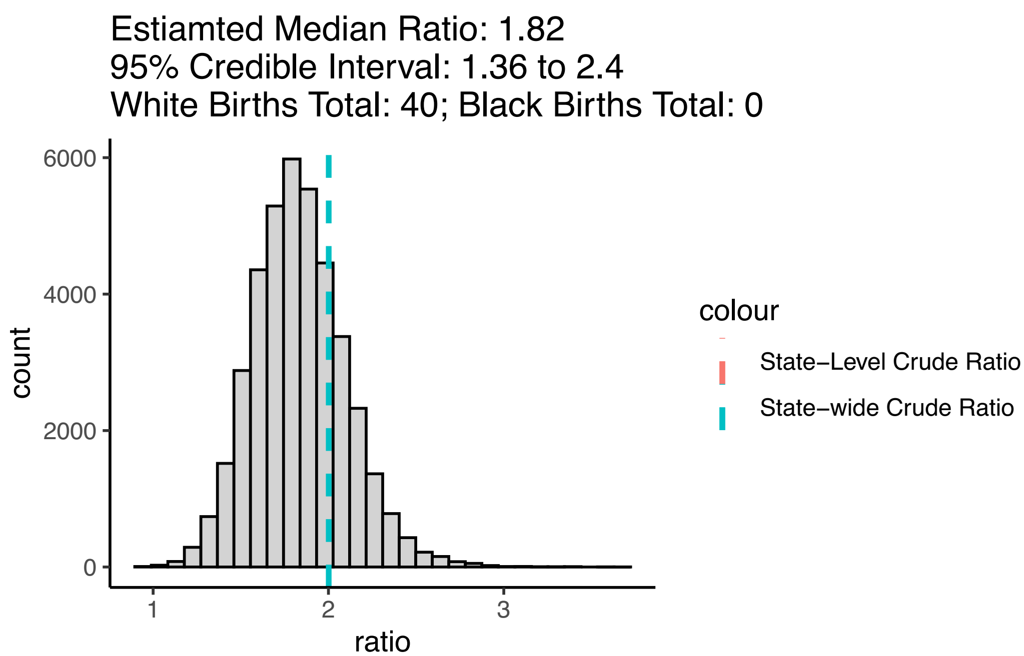

Figure 10: Theoretical distribution of low-birthweight events between Black and White moms

PUBH 7440: Intro to Bayesian Analysis
Midterm (Take-Home Portion) — Due March 14

Incidence of low weight births in PA: [Insert text saying why looking at the incidence of low birth weight is important]. Here, we let y_{ir} denote the number of low weight births from mothers of race r ($r = 1$ white, $r = 2$ black) in county i out of a total of n_{ir} births. To model these data, we will assume:

$$y_{ir} \sim \text{Bin}\left(n_{ir}, \pi_{ir}\right), \text{ where logit}\left(\pi_{ir}\right) = \theta_{ir} \sim \text{Norm}\left(\beta_{0r}, \sigma_r^2\right),$$

and where π_{ir} represents the incidence rate. Assuming standard priors for $\beta_{0r} \sim \text{Norm}\left(0, \tau^2\right)$ and $\sigma_r^2 \sim \text{IG}\left(0.001, 0.001\right)$, with $\tau^2 = 10{,}000$, answer the following questions:

1. Write the full hierarchical model.
2. Derive the full-conditional distributions for β_{0r}, π_{ir}, and σ_r^2. Which parameters have full-conditional distributions we can sample from directly, and which parameters require Metropolis steps to sample?
3. Write code to fit the model, and use $\beta_{0r} = 0$ and $\sigma_r^2 = 1$ as initial values.
    - Make history plots of β_{0r} and σ_r^2 for both races and assess model convergence. Is burn-in required? If so, how much?
4. Suppose we're interested in investigating racial disparities in the incidence of low weight births. Using the β_{0r} terms, make a histogram of the posterior distribution of the log odds ratio. Does this indicate evidence of a "significant" racial disparity? (Hint: The log odds ratio is represented by γ_1 in the conventional regression model parameterization, $E\left[\theta_{ir} \,|\, \boldsymbol{\gamma}, \sigma_r^2\right] = \gamma_0 + \gamma_1 \times (r-1)$ where $r = 1, 2$, so you'll need to first write γ_1 as a function of the β_{0r} parameters.)
5. Now suppose we're interested in *geographic* trends in the incidence of low weight births by race and in their racial disparities. Using the mapping code from HW3/HW4, make the following maps:
    - The incidence of low weight births for white mothers.
    - The incidence of low weight births for black mothers.
    - The black/white ratio of the incidence of low weight births.
6. Finally, make histograms of posterior distribution of the black/white ratio of the incidence of low weight births in Philadelphia County ($i = 51$) and Sullivan County ($i = 57$) and compare these to their respective crude estimates (i.e., the ratio of the crude incidence rates, y_{ir}/n_{ir}, for black and white mothers in both counties) and the statewide averages (i.e, the ratio of $\sum_i y_{ir} / \sum_i n_{ir}$ for black and white mothers). Are the posterior distributions consistent with either/both of these estimates based on the data? From a statistical perspective, would you have any reservations about presenting these results?

General Details:

1) we have $i = 1, 2, \ldots, 67$ indexing counties
2) within each county we have two racial groups

$$y_{ir} \sim Bin(n_{ir}, \pi_{ir})$$

$$logit(\pi_{ir}) = log\left(\frac{\pi_{ir}}{1-\pi_{ir}}\right) = \theta_{ir} \sim Norm(\beta_{0r}, \theta_r^2)$$

$$\pi_{ir} = \frac{exp(\theta_{ir})}{1+exp(\theta_{ir})}$$

3) we have the same mean and variance for all θ_{ir} within a fixed r.

4) So, we need to estimate:

2 total β_{0r}

2 total θ^2_r

67×2 total θ_{ir}

Problem 1:

Full model:

$$p(\theta_{ir}, \beta_{0r}, \sigma^2_r \mid y_{ir}, n_{ir}) \propto$$

$$\prod_{i=1}^{N=67} \prod_{j=1}^{r=2} Bin(n_{ir}, \pi_{ir}) \times$$

$$Norm(\beta_{0r}, \sigma^2_r) \times$$

$$Norm(0, \tau^2) \times$$

$$IG(0.001, 0.001)$$

$$\tau^2 = 10{,}000$$

Problem 2

1) Full conditional β_{0r}.

we have a total of 2 β_{0r} for $r=1$ and $r=2$

- β_{0r} appears in a distribution of θ_{ir}.

- for each β_{0r} we have 67 θ_{ir}'s within each level of r.

$$P(\beta_{0r} \mid \cdot) \propto \left[\prod_{i=1}^{67} Norm(\theta_{ir} \mid \beta_{0r}, \sigma_r^2)\right] \times$$

$$Norm(\beta_{0r} \mid 0, \tau^2) =$$

$$(2\pi\sigma_r^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma_r^2} \sum (\theta_{ir} - \beta_{0r})^2\right) \times$$

$$\times (-2\pi\tau^2)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\tau^2}\beta_{0r}^2\right) \propto$$

$$\exp\left(-\frac{1}{2\sigma^2}\sum(\theta_{ir}-\beta_{0r})^2-\frac{1}{2\tau^2}\beta_{0r}^2\right)=$$

$$\exp\left(-\frac{1}{2\sigma^2}\cancel{\sum\theta_{ir}^2}+\frac{1}{2\sigma^2}\cdot 2\beta_{0r}\sum\theta_{ir}-\frac{1}{2\sigma^2}\sum\beta_{0r}^2-\frac{1}{2\tau^2}\beta_{0r}^2\right)\propto$$

$$\exp\left(\frac{\beta_{0r}}{\sigma^2}\sum\theta_{ir}-\frac{1}{2}\beta_{0r}^2\left(\frac{n}{\sigma^2}+\frac{1}{\tau^2}\right)\right)$$

After considering Metropolis,

I worked out that this update can be done w/ normal distribution.

Details Below ↓

Alternatively for β_0 consider a general normal distribution:

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right) \propto$$

$$\exp\left(-\frac{1}{2\sigma^2}\left(x^2 - 2x\mu + \mu^2\right)\right) \propto$$

$$\exp\left(-\frac{1}{2\sigma^2}x^2 + \frac{1}{\sigma^2}x \cdot \mu\right)$$

This is exactly what we have above:

Change notation of original distribution:

$\sigma \longrightarrow S$, Then, by matching the terms we get

$$\begin{cases} \frac{1}{2\sigma^2} = \frac{1}{2}\left(\frac{n}{S^2} + \frac{1}{\tau^2}\right), \\ \frac{\mu}{\sigma^2} = \frac{\sum \theta_{ir}}{S^2}. \end{cases}$$

$$\frac{1}{\sigma^2} = \left(\frac{n}{S^2} + \frac{1}{\tau^2}\right) \Rightarrow \sigma^2 = \frac{1}{\left(\frac{n}{S^2} + \frac{1}{\tau^2}\right)}$$

Plug this into the second equation in the system:

$$\frac{\mu}{\theta^2} = \mu \times \left(\frac{n}{s^2} + \frac{1}{\tau^2} \right) = \frac{\sum \theta_{ir}}{s^2},$$

$$\mu = \frac{\sum \theta_{ir}}{s^2 \left(\frac{n}{s^2} + \frac{1}{\tau^2} \right)} = \frac{\sum \theta_{ir}}{n + s^2/\tau^2}$$

Switch notation back to $s \to \theta$

Potentially we can sample β_{0r} from $N\left(\frac{\sum \theta_{ir}}{n + \theta^2/\tau^2}, \frac{1}{\left(\frac{n}{\theta^2} + \frac{1}{\tau^2}\right)} \right)$

2) Full conditional for π_{ir}

$$Y_{ir} \sim \binom{n_{ir}}{y_{ir}} \pi_{ir}^{y_{ir}} (1-\pi_{ir})^{n_{ir}-y_{ir}}$$

$$\log\left(\frac{\pi_{ir}}{1-\pi_{ir}}\right) = \theta_{ir} \sim Norm(\beta_{0r}, \sigma_i^2)$$

$$\pi_{ir} = \frac{\exp(\theta_{ir})}{1+\exp(\theta_{ir})} \Longrightarrow$$

$$\boxed{Y_{ir} \sim \binom{n_{ir}}{y_{ir}} \left[\frac{\exp(\theta_{ir})}{1+\exp(\theta_{ir})}\right]^{y_{ir}} \left[\frac{1}{1+\exp(\theta_{ir})}\right]^{n_{ir}-y_{ir}}}$$

π_{ir} has a deterministic relationship with θ_{ir}, so we obtain a full conditional for θ_{ir}, and obtain π_{ir} using expit(·) function

$$p(\theta_{ir} \mid \cdot) \propto$$

$$\cancel{\binom{n_{ir}}{y_{ir}}} \left[\frac{\exp(\theta_{ir})}{1+\exp(\theta_{ir})} \right]^{y_{ir}} \left[\frac{1}{1+\exp(\theta_{ir})} \right]^{n_{ir}-y_{ir}} \times$$

$$\exp\left(-\frac{1}{2\sigma^2} (\theta_{ir} - \beta_{0r})^2 \right) =$$

$$\exp(\theta_{ir})^{y_{ir}} \times (1+\exp(\theta_{ir}))^{-y_{ir}} \times$$

$$(1+\exp(\theta_{ir}))^{-n_{ir}+y_{ir}} \times \exp\left(-\frac{1}{2\sigma^2}(\theta_{ir}-\beta_{0r})^2\right)$$

$$= \exp(\theta_{ir})^{y_{ir}} \times (1+\exp(\theta_{ir}))^{-n_{ir}} \times$$

$$\exp\left(-\frac{1}{2\sigma^2} (\theta_{ir} - \beta_{0r})^2 \right)$$

Will use metropolis updates
for this one

3) full conditional for σ^2

As known before, this should be a conjugate prior, so full conditional must be IG as well.

$$p(\sigma^2 \mid \cdot) \propto \prod_{i=1}^{n=67} Norm(\theta_{ir} / \beta_{0r}, \sigma^2) \times$$

$$\times IG(\sigma^2 / 0.001, 0.001)$$

$$\propto (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum (\theta_{ir} - \beta_{0r})^2\right) \times$$

$$(\sigma^2)^{-0.001-1} \times \exp(-0.001/\sigma^2)$$

Let $\sum_i (\theta_{ir} - \beta_{0r})^2 = A$

$$\propto (\sigma^2)^{-\frac{n}{2}-0.001-1} \times$$

$$\exp\left(-\left\{\frac{1}{2} A\right\} \cdot \frac{1}{\sigma^2} - 0.001/\sigma^2\right)$$

$$= (\theta^2)^{-\left(\frac{n}{2}+0.001\right)-1} \times$$

$$\exp\left(-\left\{\frac{1}{2}A+0.001\right\}\cdot\frac{1}{\theta^2}\right)$$

$$\Rightarrow p(\theta^2 \mid \cdot) \sim IG\left(\frac{n}{2}+0.001,\ \frac{1}{2}A+0.001\right)$$

Problem 3

Notes for code: we need to get ratios for Metropolis updates.

1) Ratios for β_{0r}

$r(\beta_{0r}') =$

$$\frac{\exp\left(\frac{\beta_{0r}^*}{\sigma^2}\sum\theta_{ir} - \frac{1}{2}\beta_{0r}^{*2}\left(\frac{n}{\sigma^2}+\frac{1}{\tau^2}\right)\right)}{\exp\left(\frac{\beta_{0r}'}{\sigma^2}\sum\theta_{ir} - \frac{1}{2}\beta_{0r}'^2\left(\frac{n}{\sigma^2}+\frac{1}{\tau^2}\right)\right)} =$$

$$\exp\left(\frac{\sum\theta_{ir}}{\sigma^2}\left(\beta_{0r}^* - \beta_{0r}'\right)\right) \times$$

$$\exp\left(-\frac{1}{2}\left(\frac{n}{\sigma^2}+\frac{1}{\tau^2}\right)\left(\beta_{0r}^{*2} - \beta_{0r}'^2\right)\right)$$

On the log Scale

$$\frac{\sum \theta_{ir}}{\sigma^2}\left(\beta_{0r}^{*} - \beta_{0r}'\right) - \frac{1}{2}\left(\frac{n}{\sigma^2} + \frac{1}{\tau^2}\right)\left(\beta_{0r}^{*2} - \beta_{0r}'^{2}\right)$$

2) Ratios for θ_{ir}

$$r(\theta_{ir}^{*}) = \frac{\exp(\theta_{ir}^{*})^{y_{ir}} \times (1+\exp(\theta_{ir}^{*}))^{-u_{ir}} \times \exp\left(-\frac{1}{2\sigma_\theta^2}(\theta_{ir}^{*}-\beta_{0r})^2\right)}{\exp(\theta_{ir}^{'})^{y_{ir}} \times (1+\exp(\theta_{ir}^{'}))^{-u_{ir}} \times \exp\left(-\frac{1}{2\sigma_\theta^2}(\theta_{ir}^{'}-\beta_{0r})^2\right)} =$$

$$= \left[\frac{\exp(\theta_{ir}^{*})}{\exp(\theta_{ir}^{'})}\right]^{y_{ir}} \times \left[\frac{1+\exp(\theta_{ir}^{*})}{1+\exp(\theta_{ir}^{'})}\right]^{-u_{ir}} \times$$

$$\times \exp\left(-\frac{1}{2\sigma_\theta^2}\left((\theta_{ir}^{*}-\beta_{0r})^2 - (\theta_{ir}^{'}-\beta_{0r})^2\right)\right)$$

On the log Scale:

$$y_{ir}(\theta_{ir}^{*} - \theta_{ir}^{\prime}) - n_{ir}\left[\log(1+\exp(\theta_{ir}^{*})) - \log(1+\exp(\theta_{ir}^{\prime}))\right] -$$

$$-\frac{1}{2}\left((\theta_{ir}^{*} - \beta_{0r})^2 - (\theta_{ir}^{\prime} - \beta_{0r})^2\right)$$

Problem 5

$$\text{logit}(\pi_{ir}) = \log\left(\frac{\pi_{ir}}{1-\pi_{ir}}\right) = \theta_{ir}.$$

$$E[\theta_{ir} \mid \beta_{0r}, \sigma^2_{0r}] = \beta_{0r}$$

$$E[\log(\pi_{ir}) - \log(1-\pi_{ir})] =$$

$$E[\log(\pi_{ir})] - E[\log(1-\pi_{ir})] = E[\theta_{ir}] = \beta_{0r}$$

$$E\left[\log\left(\frac{\pi_{ir}}{1-\pi_{ir}}\right)\right] = \beta_{0r}$$

let p_{0r} be the rate of events within group r.

Then, log odds ratio is given by

$$\log\left[\frac{\frac{p_{01}}{1-p_{01}}}{\frac{p_{02}}{1-p_{02}}}\right] = \log\left[\frac{p_{01}}{1-p_{01}}\right] - \log\left[\frac{p_{02}}{1-p_{02}}\right]$$

So this basically boils down to

$\underline{\beta_{01} - \beta_{02}}$. |