Denis Ostroushko - PUBH 7440 - HW4 - Part 1

Problem 1

Prerequisites

In this assignment we analyze stroke-related mortality rates at the county-age-group levels in PA.

We have 67 counties, $i = 1, 2, ..., 67$, and three age groups $a = 1, 2, 3$ within each county

We assume that the number of observed death in county i and age group a is distributed by a Poisson distribution with parameter $n_{ia}\lambda_{ia}$, where $log\lambda_{ia} = \beta_{0a} + z_{ia}$.

So, the death rate for each county and age group is some function of an average effect for a given age group and an age-group-and-county specific random effect.

Given that Poisson distribution parameter is a function of two random variables, we can write pmf of Y_{ia} as:

$$Y_{ia} = \frac{e^{-(n_{ia}e^{\beta_{0a}+z_{ia}})} \times (n_{ia}e^{\beta_{0a}+z_{ia}})^{Y_{ia}}}{Y_{ia}!}$$

As mentioned previously, β_{0a} and z_{ia} are random variable {because Bayesian Analysis framework}, and therefore they have prior distributions:

$\beta_{0a}|\mu = 0, \tau_a^2 \sim N(0, \tau_a^2)$, where $\tau_a^2 = 10,000$. This equation represents three prior distributions for each age group subject to analysis. They all have identical prior distributions.

$z_{ia}|\mu = 0, \sigma_a^2 \sim N(0, \sigma_a^2)$, where σ_a^2 is also a random variable that has it's own prior distribution. Note that each county and age group (201 total data points) each have their own random effect. But, within an age group a, all random variables $z_{i,a=a}$ have the the same prior distribution with variance $\sigma_{a=a}^2$

$\sigma_a^2 \sim IG(0.001, 0.001)$, so variance comes from a non-informative Inverse Gamma (IG) distribution.

Suppressed values of deaths with county and age-groups levels

- Note: I am reusing the description of the imputation procedure given in HW3, only changing max value from 10 to 9

In order to impute missing/suppressed values of $Y_{i\alpha}$ we need to use a truncated left tail of a poisson distribution with corresponding parameter $n_{i\alpha}\lambda_{i\alpha}$. We will set a maximum value at the tail equal to 9, meaning that for our imputations we will be sampling integers from 0 to 9 from poisson distributions. In order to do that, we follow these steps:

1. For each county for each group age, determine a parameter for the poisson distribution, refer to it as $\Lambda_{i\alpha}$.
2. For each county for each age group, determine quantile corresponding to value of 10 under $\Lambda_{i\alpha}$, call this quantile q

- use `ppois()` to get this quantile

3. Sample a number from a uniform distribution between 0 and q. This will be between 0 and some number less than or equal to 1 always.

- use `runif(n=1, min = 0, max = .)`

4. Using inverse CDF of a poisson distribution with parameter $\Lambda_{i\alpha}$, obtain a value corresponding to a randomly sampled quantile

- use `qpois()` for this step

5. Impute missing value with sampled values between 0 and 9

Full hirerachical model

$$
\begin{aligned}
p(\beta_{0a}, z_{ia}, \sigma^2_{0a}|\mathbf{Y}) \propto \Pi_{i,a} \left[Pois(Y_{ia}|n_{ia} * exp(\beta_{0a} + z_{ia}))\right] \times & \quad \textit{full data likelihood} \\
I(Y_{ia} < 10)^{1-d_{ia}} \times & \quad \textit{add information about censoring of } Y_{ia} \\
Norm(\beta_{0a}|0, \tau_a^2) \times & \quad \textit{prior for } \beta_{0a} \\
Norm(z_{ia}|0, \sigma_a^2) \times & \quad \textit{prior for } z_{ia} \\
IG(\sigma_a^2|0.001, 0.001) \times & \quad \textit{prior for } \sigma_a^2
\end{aligned}
$$

Problem 2

Full conditional for β_{0a}

- Term β_{0a} appears in the likelihood of the data and its own prior distribution. Other parameters of the hierarchical model are treated as constants with respect to β_{0a}.

Full conditional for β_{0a}, holding a fixed. This full conditional generalizes to three terms for each age group:

$$
\begin{aligned}
p(\beta_{0a}|.) &\propto \Pi_{i,a}\left[Pois(Y_{ia}|n_{ia} * exp(\beta_{0a} + z_{ia}))\right] \times Norm(\beta_{0a}|0, \tau_a^2) \\
&\propto exp(-\sum_i n_{ia} exp(\beta_{0a} + z_{ia})) \times exp(\sum_i Y_{ia}(\beta_{0a} + z_{ia})) \times exp(-\frac{1}{2} * \frac{\beta_{0a}^2}{\tau_a^2})
\end{aligned}
$$

I do not recognize this kernel as a known distribution, so we are not able to use Gibbs sampling technique to obtain posterior distribution of β_{0a}. We will use Metropolis algorithm to obtain samples for the posterior distribution. More on this later, but we will use Metropolis update because we assume a symmetric candidate density for β_{ia}

Full conditional for z_{ia}

- Term z_{ia} appears in the likelihood of the data and its own prior distribution. Other parameters of the hierarchical model

are treated as constants with respect to z_{ia}. Since each z_{ia} is specific to its county and age group, we will use pmf of Y_{ia} given by the Poisson distribution instead of the full likelihood of Y_{ia} for age group a, which is different from the full conditional of β_{0a}.

This derivation gives a general version of a full conditional distribution for county i and age group a, we obtain 201 posterior distributions in our analysis.

$$
\begin{aligned}
p(z_{ia}|.) &\propto Pois(Y_{ia}|n_{ia} * exp(\beta_{0a} + z_{ia})) * Norm(z_{ia}|0, \sigma_a^2) \\
&\propto exp(-n_{ia} * exp(\beta_{0a} + z_{ia})) \times exp(Y_{ia} * (\beta_{0a} + z_{ia})) \times exp(-\frac{1}{2}\frac{z_{ia}^2}{\sigma_a^2})
\end{aligned}
$$

I do not recognize this kernel as a known distribution, so we are not able to use Gibbs sampling technique to obtain posterior distribution of z_{ia}. We will use Metropolis algorithm to obtain samples for the posterior distribution. More on this later, but we will use Metropolis update because we assume a symmetric candidate density for z_{ia}

Full conditional for σ_a^2

- Term σ_a^2 appears in the prior distribution of random effects z_{ia}. All other models terms are treated as constants with respect to σ_a^2, and will make up a normalizing constant for the proper density of the full conditional distribution.
- Because σ_a^2 appears in each of 67 z_{ia} for a given age group a, we need to use a joint likelihood of random effects. Assuming independence of z_{ia}, we can obtain join prior distribution as a product of 67 prior densities.
- Similar to β_{0a}, we derive a general form of the full conditional distribution for this parameter. In our analysis, we will obtain three posterior distributions for σ_a^2 for each age group.
- Inverse Gamma is a conjugate prior to a normal distribution, so we should expect to obtain an Inverse Gamma posterior distribution. We will therefore use Gibbs sampling to get samples of σ_a^2

Full conditional:

$$\begin{aligned} p(\sigma_a^2|.) &\propto \prod_i \frac{1}{\sqrt{2\pi\sigma_a^2}} exp(-\frac{1}{2\sigma_a^2}\sum_i z_{ia}^2) \times (\sigma_a^2)^{-a-1} exp(-b/\sigma_a^2) \\ &\propto (\sigma_a^2)^{-n/2-a-1} \times exp(-(b + \frac{1}{2}\sum_i z_{ia}^2) * 1/\sigma_a^2) \end{aligned}$$

Full conditional for σ_a^2 is proportional to the kernel of the Inverse Gamma with parameters $\alpha = n/2 + a$ and $\beta = b + \frac{1}{2}\sum_i z_{ia}^2$

Problem 3

Code to fit the model and obtain posterior distributions for parameters of interest is attached in the appendix after comparison with the HW3 results.

To fit the model, I used the following parameters and candidate densities:

- Assume a symmetric candidate density $\beta_0 \sim Norm(\grave{\beta_0}, q)$ where $q = 0.075$
- Assume a symmetric candidate density $z_{ia} \sim Norm(\grave{z_{ia}}, q_{ia})$ where q_{ia} is proportional to the data-driven point estimate for the county-specific effect on the observed log-rate of stroke related mortality rate

Ratio for Updates β_{0a}

Using full conditional for β_{0a}, and omitting candidate density due to symmetrical property, we obtain the following equation to test each proposed sample for acceptance/rejection:

$$r(\beta_{0a}^{\star}) = exp(-\frac{1}{2\tau_a^2}((\beta_{0a}^{\star})^2 - (\grave{\beta}_{0a})^2))\times$$
$$exp(\sum_i Y_{ia}(\beta_{0a}^{\star} - \grave{\beta}_{0a})) \times exp([\sum_i n_{ia}exp(z_{ia})] * (exp(\grave{\beta}_{0a}) - exp(\beta_{0a}^{\star})))$$

Ratio for Updates z_{ia}

Using full conditional for z_{ia}, and omitting candidate density due to symmetrical property, we obtain the following equation to test each proposed sample for acceptance/rejection. Recall that each z_{ia} is specific to the county-age-group level, so we do not aggregate the data, like with the β_{0a} example.

$$r(z_{ia}^{\star}) = exp(-n_{ia}(exp(\beta_{0a} + z_{ia}^{\star}) + exp(\beta_{0a} + \grave{z}_{ia})))\times$$
$$exp(Y_{ia}(z_{ia}^{\star} - \grave{z}_{ia})) \times exp(-\frac{1}{2\sigma_a^2}((z_{ia}^{\star})^2 - (\grave{z}_{ia})^2))$$

Ratios for Update σ_a^2 is not needed because we use Gibbs sampling

Notes for the derivation are attached at the end of the document after the code section

Results for β_{0a}

Figure 1 shows posterior distributions for the age-group overall effect on deaths associated with stroke. Since $\log \lambda_{ia} = \beta_{0a} + z_{ia}$, presented values are on the logarithmic scale. We can make an observation that as overall age increases, the age-group overall ‘average’ death rate increases, which is something that we would expect to observe.

All posterior distributions have a nice symmetric shape, fitting a normal candidate distribution.

Figure 2 presents county-specific age-adjusted rates. Overall, the map looks similar to what we observed under the Poisson-Gamma model (HW3).

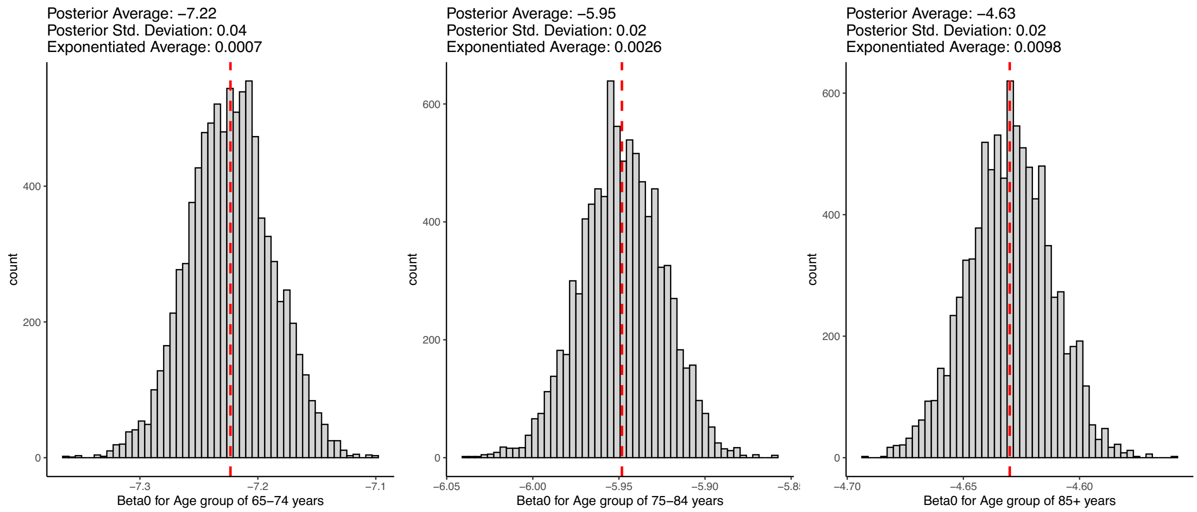

Figure 1: Posterior Distribtuions of Beta0 for the three age groups

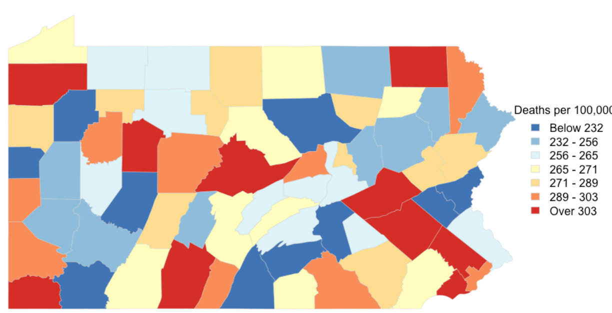

Figure 2: Final Map of Rates

Appendix

Comparison with Poisson-Gamma model

Figure 3 shows the differences for the county specific estimates between the two approaches. It it evident that under my analysis, mixed-effects regression model tends to estimate much higher stroke related mortality rates for counties with smaller population size.

In some cases the differences are as high as 20%. I suspect that the primary difference between the results are due to the use of random effects.

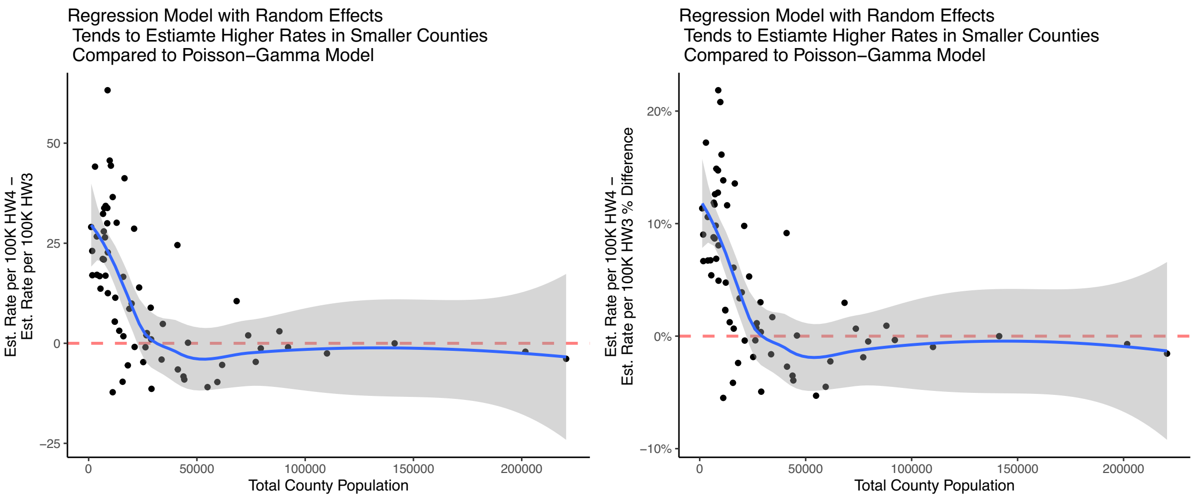

Figure 3: Comparion of Age Adjusted Rates Uisng Mixed-Effects Regression Model (HW4) and Poisson-Gamma Model (HW3)

Metropolis-Hastings Sampling Algorithm R-code

```
set.seed(1234)

reps = 10000

lambda_ia <-
  cbind(
    stroke_clean %>% select(lambda_0) %>% unlist(),
    matrix(data = NA,
           nrow =   stroke_clean %>% select(lambda_0) %>% unlist() %>% length(),
           ncol = (reps-1)
             )
```

```
  )
## get guesses for beta_oa as the group average from data
# first, if there are missing values, impute with prior guess for lambda0

stroke_clean %>%
  mutate(final_y = ifelse(is.na(deaths), lambda_0 * population, deaths),
         log_rate = log(final_y/population)
         ) %>%
  group_by(age.group) %>%
  summarize(b0a = mean(log_rate)) %>%
  ungroup() %>%
  select(b0a) %>%
  unlist() -> boa_guess

beta_0a <-
  cbind(
    boa_guess,
    matrix(data = NA,
           nrow =   3,
           ncol = (reps-1)
             )
  ) ## these are some pretty bad guesses for the betas, but it will work for now

## get initial guesses for z_ia as the difference between observed Y minus age_group avera
# first, if there are missing values, impute with prior guess for lambda0

stroke_clean %>%
  mutate(final_y = ifelse(is.na(deaths), lambda_0 * population, deaths),
         log_rate = log(final_y/population)
         ) %>%
  group_by(age.group) %>%
  mutate(b0a = mean(log_rate)) %>%
  ungroup() %>%
  mutate(z_ia = log_rate -b0a) %>%
  select(z_ia) %>%
  unlist() -> zi_guess

z_ia <-
  cbind(
    zi_guess,
   # rep(0, length(zi_guess)),
```

```
    matrix(data = NA,
           nrow =   length(zi_guess),
           ncol = (reps-1)
             )
  ) ## these are some pretty bad guesses for the z_i's, but it will work for now

sigma_0a <-
  cbind(
    c(1,1,1),
    matrix(data = NA,
           nrow =   3,
           ncol = (reps-1)
             )
  )## these are some pretty bad guesses for the z_i's, but it will work for now

tau2 = 10000
a = 0.001
b = 0.001

q_norm_b  = 0.075
# q_norm_zi = 0.005
q_norm_zi = abs(zi_guess)*5 # make it such that the step size for each z_ia is poportional
#                              ratio is 1/1
q_ig = 1

n = 67

for(i in 2:reps){

  if(i %% 1000 == 0){print(i)}
  ######################
  # DATA IMPUTATION STEP
  ######################

  lambda_ia[,(i-1)] * stroke_clean$population -> poisson_lambdas_iter

  ppois(9.5, poisson_lambdas_iter) -> limits_detection_iter

  # using these numbers between 0 and somewhere less than 1, sample from uniform distribut
  runif(n = length(limits_detection_iter), min = 0, max = limits_detection_iter) -> sample
```

```
# get imputed values by putting unifrom random samples into 'inverse' CDF
qpois(sampled_u, lambda = poisson_lambdas_iter) -> imp

# get final imputed vector of the observed data
stroke_clean$deaths -> final_ys_iter
final_ys_iter[which(is.na(final_ys_iter))] <- imp[which(is.na(final_ys_iter))]

############
# UPDATE sigma_a

# sample new sigma from the candidate density
## sample new values of sigma from posterior


for(SIGMA in 1:3){

  # s_prop = sig_proposed[SIGMA]
  # s_curr = sigma_0a[,(i-1)][SIGMA]
  # identify what rows of random effects to grab
  z_rows <- seq(from = SIGMA,
                to = length(final_ys_iter) - (3- SIGMA),
                length.out = 67)

  # data for ratio
  z_ia[z_rows, (i-1)] -> random_effs

  # (s_prop/s_curr)^(2*q_ig - a - n/2) *
  #
  #   exp(-1/2 * sum(random_effs^2) * (1/s_prop - 1/s_curr)) *
  #
  #   exp(-b * ((1/s_prop - 1/s_curr)) *
  #
  #   exp(q_ig * (s_prop/s_curr - s_curr/s_prop)) -> ratio

  sigma_0a[,(i)][SIGMA] = 1 / rgamma(n = 1, n/2 + a, sum(random_effs^2)/2+b )

}

############
# UPDATE Z_ia
```

```
b_0a = beta_0a[,(i-1)]
b_0a_calc = rep(b_0a, n)

sig2 = sigma_0a[,(i)]
sig2_calc = rep(sig2, n)

z_ia_curr = z_ia[,(i-1)]
z_ia_prop = rnorm(n = n*3, mean = z_ia_curr, sd = q_norm_zi)

(-stroke_clean$population *
      (exp(b_0a_calc + z_ia_prop) + exp(b_0a_calc +z_ia_curr ))
    ) +

  (final_ys_iter*(z_ia_prop - z_ia_curr)) +

  (-1/(2 * sig2_calc) * (z_ia_prop^2 - z_ia_curr^2)) -> ratio

z_ia[,i] <- ifelse(exp(ratio) > runif(n = length(ratio)), z_ia_prop, z_ia_curr)
############
# UPDATE B_0a

for(POP in 1:3){

  most_recent_beta0a <- beta_0a[POP, (i-1)]
  sampled_beta0a <- rnorm(n = 1, mean = most_recent_beta0a, sd = q_norm_b)

  POP_rows <- seq(from = POP,
                  to = length(final_ys_iter) - (3- POP),
                  length.out = 67)

  Y_ipop = final_ys_iter[POP_rows]
  n_ipo = stroke_clean$population[POP_rows]
  z_ia_calc = z_ia[,i][POP_rows]

  ratio  <-
    (sum(Y_ipop) * (sampled_beta0a - most_recent_beta0a)) +

    (sum(n_ipo * exp(z_ia_calc)) * (exp(most_recent_beta0a) - exp(sampled_beta0a)) ) +

    (-1/(2 * tau2) * (sampled_beta0a^2 - most_recent_beta0a^2))
```

```r
    beta_0a[POP, (i)] <- ifelse(exp(ratio) > runif(n=1), sampled_beta0a, most_recent_beta0
  }

  ## update lambda based on beta and zeta
  beta_0a_for_lambda = rep(beta_0a[,i], n)

  lambda_ia[,(i)] = exp(beta_0a_for_lambda + z_ia[,i])

}

res <- list(lambda_ia,
                   sigma_0a,
                   z_ia,
                   beta_0a
                   )

names(res) <- c("lambdas", "sigmas", "zs", "betas")

write_rds(x = res,
          file = "./MHresults/MH results7.rds"
          )
```

Problem 1

Full hierarchical model:

- $Y_{ia} \sim Pois(n_{ia} \lambda_{ia})$

  $\log \lambda_{ia} = \beta_{0a} + Z_{ia}$

  $\lambda_{ia} = \exp(\beta_{0a} + Z_{ia}) \Rightarrow$

  $Y_{ia} \sim Pois(n_{ia} \exp(\beta_{0a} + Z_{ia}))$

- Full model

$$p(\beta_{0a}, Z_{ia}, \sigma_a^2 \mid Y_{ia}) \propto h(\vec{\theta}) =$$

$$\left[\prod_{i,a} Pois(Y_{ia} \mid \beta_{0a}, Z_{ia})\right]^{1-d_{ia}} \times \mathbb{1}(Y_{ia} < 10) \times$$

$$\times Norm(\beta_{0a} \mid 0, \tau_0^2) \times$$

$$\times Norm(Z_{ia} \mid 0, \sigma_a^2) \times$$

$$\times IG(\sigma_a^2 \mid a, b)$$

where $\tau_0^2 = 10{,}000$

$a = 0.001$

$b = 0.001$

Problem 3

likelihood of data is:

$$L = \prod_{i,a} \left[\exp\left(-n_{ia} \exp(\beta_{0a} + z_{ia})\right) \times \left(n_{ia} \exp(\beta_{0a} + z_{ia})\right)^{y_{ia}} \times \frac{1}{y_{ia}!} \right]$$

$$= \exp\left(-\sum_{i,a} n_{ia} \exp(\beta_{0a} + z_{ia})\right) \times \prod_{i,a} \times n_{ia}^{y_{ia}} \times \exp\left(y_{ia}(\beta_{0a} + z_{ia})\right) \times \frac{1}{y_{ia}!}$$

$$= \exp\left(-\sum_{i,a} n_{ia} \exp(\beta_{0a} + z_{ia})\right) \times \exp\left(\sum y_{ia}(\beta_{0a} + z_{ia})\right) \times \prod \frac{n_{ia}^{y_{ia}}}{y_{ia}!}$$

$$\propto \exp\left(\sum y_{ia}(\beta_{0a} + z_{ia}) - \sum n_{ia} \exp(\beta_{0a} + z_{ia})\right)$$

◦ Given a new version of data likelihood, derive a new ratio form

$$h(\beta_{0a}) = \exp\left(\sum y_{ia}(\beta_{0a} + z_{ia}) - \sum u_{ia}\exp(\beta_0 + z_{ia})\right) \times \frac{1}{\sqrt{2\pi\tau_a^2}} e^{-\frac{1}{2}\frac{\beta_{0a}^2}{\tau_a^2}}$$

ratio: $r(\beta^*_{0a}) =$

$$\frac{\exp\left(\sum y_{ia}(\beta^*_{0a} + z_{ia}) - \sum u_{ia}\exp(\beta^*_{0a} + z_{ia})\right) \times e^{-\frac{1}{2}\beta^{*2}_{0a}/\tau_a^2}}{\exp\left(\sum y_{ia}(\beta'_{0a} + z_{ia}) - \sum u_{ia}\exp(\beta'_{0a} + z_{ia})\right) \times e^{-\frac{1}{2}\beta'^2_{0a}/\tau_a^2}}$$

$$= \exp\left(\sum y_{ia}(\beta^a_{0a} - \beta'_{0a})\right) \times$$

$$\exp\left(\sum u_{ia}\exp(\beta'_{0a} + z_{ia}) - u_{ia}\exp(\beta^*_{ia} + z_{ia})\right) \times$$

$$\exp\left(\frac{1}{2\tau_a^2}\left(\beta_{0a}^{*2} - \beta_{0a}^{'2}\right)\right)$$

$$= \exp\left(\sum y_{ia}\left(\beta_{0a}^{a} - \beta_{0a}^{'}\right) \times\right.$$

$$\exp\left(\sum u_{ia}\exp(\beta_{0a}^{'})\exp(z_{ia}) - u_{ia}\exp(\beta_{ia}^{*})\exp(z_{ia})\right) \times$$

$$\exp\left(-\frac{1}{2\tau_a^2}\left(\beta_{0a}^{*2} - \beta_{0a}^{'2}\right)\right)$$

$$= \exp\left(\sum y_{ia}\left(\beta_{0a}^{*} - \beta_{0a}^{'}\right)\right) \times$$

$$\exp\left(\left[\sum u_{ia}\exp(z_{ia})\right]\left(\exp(\beta_{a}^{'}) - \exp(\beta_{ia}^{*})\right)\right) \times$$

$$\exp\left(-\frac{1}{2\tau_a^2}\left(\beta_{0a}^{*2} - \beta_{0a}^{'2}\right)\right) = R$$

- Translate to the log scale
- But, the result will need to be exponentiated anyway.

$$\log R =$$

$$\sum y_{ia} \left(\beta^{*}_{0a} - \beta'_{0a} \right) +$$

$$\left[\sum u_{ia} \exp(z_{ia}) \cdot \left(\exp(\beta'_{0a}) - \exp(\beta^{*}_{0a}) \right) \right] -$$

$$\frac{1}{2\tau^2_a} \left(\beta^{*2}_{0a} - \beta'^{2}_{0a} \right)$$

Update Z_{ia}.

Z_{ia} is specific to each county and age group. Therefore, I don't think we need to use likelihood, just the dens.

- $$Y_{ia} \sim \exp(-n_{ia}\exp(\beta_{0a}+Z_{ia})) \times (n_{ia}\exp(\beta_{0a}+Z_{ia}))^{Y_{ia}} \times \frac{1}{(Y_{ia}!)}$$

$$Z_{ia} \sim \frac{1}{\sqrt{2\pi\theta_a^2}}\, e^{-\frac{1}{2}\frac{Z_{ia}^2}{\theta_a^2}}$$

- ratio:

$$r(Z^*_{ia}) = \frac{\exp(-n_{ia}\exp(\beta_{0a}+Z^*_{ia})) \times (\cancel{n_{ia}}\exp((\beta_{0a}+Z^*_{ia})))^{Y_{ia}} \times \cancel{\frac{1}{Y_{ia}!}} \times \cancel{\frac{1}{\sqrt{2\pi\theta_a^2}}}\, e^{-Z^{*2}/2}}{\exp(-n_{ia}\exp(\beta_{0a}+Z'_{ia})) \times (\cancel{n_{ia}}\exp((\beta_{0a}+Z'_{ia})))^{Y_{ia}} \times \cancel{\frac{1}{Y_{ia}!}} \times \cancel{\frac{1}{\sqrt{2\pi\theta}}}\, e^{-Z'^2/\sigma}}$$

$$= \exp\left(-u_{ia}\left(\exp(\beta_{0a} + z^*_{ia}) + \exp(\beta_{0a} + z'_{ia})\right)\right) \times$$

$$\exp\left(y_{ia}\left(z^*_{ia} - z'_{ia}\right)\right) \times$$

$$\exp\left(-\frac{1}{2\sigma^2_a}\left(z^{*2}_{ia} - z'^2_{ia}\right)\right) = R$$

$$\log R =$$

$$\left(-u_{ia}\left[\exp(\beta_{0a} + z^*_{ia}) + \exp(\beta_{0a} + z'_{ia})\right]\right) +$$

$$y_{ia}\left(z^*_{ia} - z'_{ia}\right) +$$

$$\left(-\frac{1}{2\sigma^2_a}\left(z^{*2}_{ia} - z'^2_{ia}\right)\right)$$

Update σ^2

$\sigma^2 \sim IG$, and it shows up in z_i with unknown variance. So, we can treat IG as a conjugate prior

Lesson learned: before lengthy calculation always look for the fact if something is a conj. prior

σ_a^2 is observed in $p(z_{ia} | a, \sigma_a^2)$

σ_a^2 is a term in its own prior IG( )

Write out densities:

$$z_{ia} \sim \frac{1}{\sqrt{2\pi\sigma_a^2}} e^{-\frac{1}{2}\frac{z_{ia}^2}{\sigma_a^2}} \Rightarrow \ell = \left(\frac{1}{\sqrt{2\pi\sigma_a^2}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_i z_{ia}^2}$$

$$\sigma_a^2 \sim \frac{b^a}{\Gamma(a)} (\sigma_a^2)^{-a-1} e^{-b/\sigma_a^2}$$

---

rewrite ℓ:

$$(2\pi)^{-\frac{n}{2}} (\sigma_a^2)^{-\frac{n}{2}} e^{-\frac{\sum z_{ia}^2}{2\sigma^2}} \propto (\sigma_a^2)^{-\frac{n}{2}} e^{-\frac{\sum z^2}{2}\cdot\frac{1}{\sigma^2}}$$

So, full conditional

$$p(\sigma^2 \mid \cdot) \propto (\sigma_a^2)^{-\frac{n}{2}} e^{-\frac{\sum z_i^2}{2}\cdot\frac{1}{\sigma_a^2}} \times$$

$$\times (\sigma_a^2)^{-a-1} e^{-b/\sigma_a^2} =$$

$$= (\sigma_a^2)^{-\underbrace{\left(\frac{n}{2}+a\right)}_{a^*}-1} e^{-\underbrace{\left(b+\frac{\sum z_i^2}{2}\right)}_{b^*}\frac{1}{\sigma_a^2}}$$