PUBH 7440: Intro to Bayesian Analysis
Homework from Week 5 — Due Feb 22nd/27th

Stroke mortality in PA: Using the same data as HW 3, we want to fit the following two models:

- $Y_{ia} \sim \text{Pois}\left(n_{ia}\lambda_{ia}\right)$ where $\log\lambda_{ia} = \beta_{0a} + z_{ia}$ and $z_{ia} \sim \text{Norm}\left(0, \sigma_a^2\right)$.
- $Y_{ia} \sim \text{Pois}\left(n_{ia}\lambda_{ia}\right)$ where $\log\lambda_{ia} = \theta_{ia} \sim \text{Norm}\left(\beta_{0a}, \sigma_a^2\right)$.

In both cases, use standard $\beta_{0a} \sim \text{Norm}\left(0, \tau_a^2\right)$ and $\sigma_a^2 \sim \text{IG}\left(0.001, 0.001\right)$ priors, with $\tau_a^2 = 10{,}000$, and answer the following questions:

1. Write the full hierarchical model. As with HW 3, account for the suppression of $Y_{ia} < 10$.
2. Derive the full-conditional distributions for $\{\boldsymbol{\beta}_0, \mathbf{z}, \sigma_a^2\}$ and $\{\boldsymbol{\beta}_0, \boldsymbol{\theta}, \sigma_a^2\}$ for the two models, respectively. 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 two models. How do the results compare? Examples of "results" to compare include:
    - Posterior summaries for β_{0a} from the two models.
    - Posterior summaries for $\lambda_{i\cdot}$, the age-adjusted rates (you can use the same code as in HW 3 to make maps of these rates).

    Note: Use symmetric candidate densities — e.g., $\theta^* \sim \text{Norm}\left(\theta^{(\ell-1)}, q\right)$ — to generate proposed values for all Metropolis steps. Don't worry too much about acceptance rates, though the closer you can get to 44% acceptance rates, the better your convergence will be. If you're having issues with this, let me know.

My plan will be to have the *first model* (the one with z_{ia}) due on Thursday Feb 22nd and the second model due Tuesday Feb 27th. I suppose that means that when you turn in the results for Model 1, you need to report the various summaries for #3, but there won't be a "compare" part of that question until you've done Model 2.

Problem 3

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

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

λ_{ia} is not a random variable because it is deterministically defined through β_{0a} and z_{ia}

$\beta_{0a} \sim Norm(0, \tau_0^2) \leftarrow$ prior: $\tau_a^2 = 10,000$ in our example

$z_{ia} \sim Norm(0, \theta_a^2) \leftarrow$ prior: θ_0^2 has a distribution

$\theta_a^2 \sim IG(a, b) \leftarrow$ prior.

$\tau_a^2 = 10,000$

$a = b = 0.01$

- since $\beta_{0a}, z_{ia}, \theta_0^2$ have priors, we need posteriors for them.

we also need candidate densities for them.

Let's say β_{0a} will have $N(\beta_{0a}', q_{0a})$

z_{ia} will have $N(z_{ia}', q_{ia})$

θ^2 will have $IG(q_\theta, q_\theta \theta_{t-1}^2)$

- i think within the loops we need to update β_{00} within each age group, so we might want to loop over them within M-H loop

This term here carries over from the last iteration of MH algo

∘ Since $\log \lambda_{ia}$ is deterministic, rewrite Poisson pmf.

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

$$P(Y_{ia} | \beta_{0a}, z_{ia}) =$$
$$(n_{ia} \exp(\beta_{0a} + z_{ia}))^{Y_{ia}} \times$$
$$\exp(-n_{ia} \exp(\beta_{0a} + z_{ia})) \times \frac{1}{(Y_{ia}!)}$$

2) Define data likelihood

$$L = \prod_{i,a} P(Y_{ia} | \beta_{0a}, z_{ia}) =$$

$$= \prod \Big[n_{ia}^{Y_{ia}} \cdot e^{(\beta_{0a} + z_{ia})^{Y_{ia}}} \times$$
$$\times e^{-n_{ia} e^{\beta_{0a} + z_{ia}}} \times \frac{1}{(Y_{ia})!} \Big]$$

$$= \prod \Big[n_{ia}^{Y_{ia}} \times e^{Y_{ia}(\beta_{0a} + z_{ia})} \times e^{-n_{ia} e^{\beta_{0a} + z_{ia}}} \times \frac{1}{Y_{ia}} \Big]$$

$$(1) = \prod_{i,a} [n_{ia}^{Y_{ia}}] \times \boxed{e^{\beta_{0a} \sum_{i,a} Y_{ia}}} \times e^{\sum_{i,a} Y_{ia} z_{ia}} \times$$

$$\times \boxed{e^{-\sum u_{i\alpha} e^{\beta_{0\alpha} + z_{i\alpha}}}} \times \frac{1}{\prod (y_{i\alpha}!)}$$

Update for candidate β_0^*

- highlight in which parts contain β_0 and have to be used for the ratio
- The rest in equation (1) remains constant w.r.t. $\beta_{0\alpha}$
- Use the fact that $\beta_{0\alpha}$ have symmetric candidate density, and therefore we do not include candidate density into update
- But we have 3 $\beta_{0\alpha}$, β_{01} for the youngest ppl, .. β_{02} and β_{03}.
- So, this update needs to be done 3 times.

$$h(\beta) = e^{\beta_{0\alpha} \sum_i y_i} \times e^{-\sum u_i e^{\beta_{0\alpha} + z_i}} \times$$

$$\times \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{\beta_{0\alpha}^2}{\tau_\alpha^2}}$$

$$r(\beta^*) = \frac{e^{\beta^* \sum y_i} \times e^{-\sum n_i e^{\beta^* + z_i}} \times e^{-\frac{\beta^{*2}}{\tau_0^2}}}{e^{\beta' \sum y_i} \times e^{-\sum n_i e^{\beta' + z_i}} \times e^{\beta'^2/\tau_0^2}}$$

$$= e^{\sum y_i (\beta^* - \beta')} \times e^{-\sum n_i \left(e^{\beta^* - \beta'}\right)} \times$$

$$\times\, e^{-\frac{1}{\tau_0^2}\left(\beta^{*2} - \beta'^2\right)}$$

I am using the fact that z^* has a symmetric (normal) candidate

Update z_{ia} density and therefore it is not included into the update

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

$$P(Y_{ia} | \beta_{0a}, z_{ia}) =$$

$$(n_{ia} \exp(\beta_{0a} + z_{ia}))^{Y_{ia}} \times$$

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

2) Define data likelihood

$$L = \prod_{i,a} P(Y_{ia} | \beta_{0a}, z_{ia}) =$$

$$= \prod \left[n_{ia}^{Y_{ia}} \cdot e^{(\beta_{0a} + z_{ia})^{Y_{ia}}} \times \right.$$

$$\left. \times e^{-n_{ia} e^{\beta_{0a} + z_{ia}}} \times \frac{1}{(Y_{ia})!} \right]$$

$$= \prod \left[n_{ia}^{Y_{ia}} \times e^{Y_{ia}(\beta_{0a} + z_{ia})} \times e^{-n_{ia} e^{\beta_{0a} + z_{ia}}} \times \frac{1}{Y_{ia}} \right]$$

$$(2) = \prod_{i,a} [n_{ia}^{Y_{ia}}] \times e^{\beta_{0a} \sum_{i,a} Y_{ia}} \times e^{\sum_{i,a} Y_{ia} z_{ia}} \times$$

$$\times e^{-\sum n_{ia} e^{\beta_{0a} + z_{ia}}} \times \frac{1}{\prod (Y_{ia}!)}$$

- define $h(z)$ using prior for $z_{i\alpha}$ and circled parts of equation (2)

$$h(z) = e^{\sum y_i z_i} \times e^{-\sum n_i e^{\beta_0 + z_i}} \times \frac{1}{\sqrt{2\pi\theta^2}} e^{-\frac{1}{2}\frac{z_i^2}{\theta_\alpha^2}}$$

- define ratio function for $h(z^*)$

$$r(z^\theta) = \frac{h(z^\theta)}{h(z')} =$$

$$= \frac{e^{\sum y_i z_i^\theta} \times e^{-\sum n_i e^{\beta_0 + z_i^\theta}} \times e^{-\frac{1}{2}\frac{z^{\theta 2}}{\theta_\alpha^2}}}{e^{\sum y_i z_i'} \times e^{-\sum n_i e^{\beta_0 + z_i'}} \times e^{-\frac{1}{2}\frac{z'^2}{\theta_\alpha^2}}}$$

$$= e^{\sum y_i (z_i^\theta - z_i')} \times e^{-\sum n_i (z_i^\theta - z_i')} \times$$

$$\times e^{-\frac{1}{2\theta^2}(z^{\theta 2} - z'^2)}$$

Update σ_α^2, we no longer have symmetric candidate density.

1) $Y_{i\alpha} \sim Pois(n_{i\alpha} \exp(\beta_{0\alpha} + z_{i\alpha}))$

$$P(Y_{i\alpha} | \beta_{0\alpha}, z_{i\alpha}) =$$

$$(n_{i\alpha} \exp(\beta_{0\alpha} + z_{i\alpha}))^{Y_{i\alpha}} \times$$

$$\exp(-n_{i\alpha} \exp(\beta_{0\alpha} + z_{i\alpha})) \times \frac{1}{(Y_{i\alpha}!)}$$

2) Define data likelihood

$$L = \prod_{i,\alpha} P(Y_{i\alpha} / \beta_{0\alpha}, z_{i\alpha}) =$$

$$= \prod \left[n_{i\alpha}^{Y_{i\alpha}} \cdot e^{(\beta_{0\alpha} + z_{i\alpha})^{Y_{i\alpha}}} \times \right.$$
$$\left. \times e^{-n_{i\alpha} e^{\beta_{0\alpha} + z_{i\alpha}}} \times \frac{1}{(Y_{i\alpha})!} \right]$$

$$= \prod \left[n_{i\alpha}^{Y_{i\alpha}} \times e^{Y_{i\alpha}(\beta_{0\alpha} + z_{i\alpha})} \times e^{-n_{i\alpha} e^{\beta_{0\alpha} + z_{i\alpha}}} \times \frac{1}{Y_{i\alpha}} \right]$$

$$(2) = \prod_{i,\alpha} [n_{i\alpha}^{Y_{i\alpha}}] \times e^{\beta_{0\alpha} \sum_{i,\alpha} Y_{i\alpha}} \times e^{\sum_{i,\alpha} Y_{i\alpha} z_{i\alpha}} \times$$
$$\times e^{-\sum n_{i\alpha} e^{\beta_{0\alpha} + z_{i\alpha}}} \times \frac{1}{\prod (Y_{i\alpha}!)}$$

But now I think we need to mutiply this likelihood by $\underline{p(z_{ia})}$ and $\underline{p(\sigma_a^2)}$ and candidate density to get the proper density for update

But, L does not have σ_a^2 in it anywhere. So, let's just start w/

$$p(z) \times p(\sigma) \times \text{candidate}(\sigma)$$

$$h(\sigma_a^2) = \frac{1}{\sqrt{2\pi\sigma_a^2}} e^{-\frac{1}{2}\frac{z_{ia}^2}{\sigma_a^2}} \times \quad \leftarrow z_{ia} \text{ prior}$$

$$\times \cancel{\frac{b^a}{\Gamma(a)}} (\sigma_a^2)^{-a-1} e^{-b/\sigma_a^2} \times \quad \leftarrow \sigma_a^2 \text{ prior}$$

$$\underbrace{\times \frac{\cancel{(q\sigma_{i-1}^2)^q}}{\cancel{\Gamma(q)}} \times (\sigma_a^2)^{-q-1} \times e^{-\frac{\cancel{q\sigma_{i-1}^2}}{\sigma_a^2}}}_{\sigma_a^2 \text{ here is a variable}}$$

Ratio is defined as:

$$r(\theta_a^{2\Delta}) = \frac{h(\theta_a^{2\Delta})}{h(\theta_a^{2'})} =$$

$$= \frac{\dfrac{\cancel{1}}{\sqrt{\cancel{2\pi}\,\theta_a^{2\Delta}}}\, e^{-\frac{1}{2}\cdot\frac{z_{ia}^2}{\theta_a^{2\Delta}}} \times (\theta_a^{2\Delta})^{-a-1}\, e^{-b/\theta_a^{2\Delta}} \times (\theta_a^{2\Delta})^{-q-1} \times e^{-\frac{q\theta_{t-1}^2}{\theta_a^{2\Delta}}}}{\dfrac{\cancel{1}}{\sqrt{\cancel{2\pi}\,\theta_a^{2'}}} \times e^{-\frac{1}{2}\cdot\frac{z_{ia}^2}{\theta_a^{2'}}} \times (\theta_a^{2'})^{-a-1}\, e^{-b/\theta_a^{2'}} \times (\theta_a^{2'})^{-q-1} \times e^{-\frac{q\theta_{t-1}^2}{\theta_a^{2'}}}}$$

$$= \frac{\sqrt{\theta_a^{2'}}}{\sqrt{\theta_a^{2\Delta}}} \times e^{-\frac{1}{2}\left(\frac{z_{ia}^2}{\theta_a^{2\Delta}} - \frac{z_{ia}^2}{\theta_a^{2'}}\right)} \times \left(\frac{\theta_a^{2\Delta}}{\theta_a^{2'}}\right)^{-a-1} \times e^{\left(b/\theta_a^{2'} - b/\theta_a^{2\Delta}\right)} \times \left(\frac{\theta_a^{2\Delta}}{\theta_a^{2'}}\right)^{-q-1} \times e^{-\frac{q\theta_{t-1}^2}{\theta_a^{2\Delta}} + \frac{q\theta_{t-1}^2}{\theta_a^{2'}}}$$

Is question now correct this thing is