New Benchmarking MCMC

October 2023

Contents

1 Data Generation

1.1 Fixed spatial effect area-level, assuming all births are observed

Population settings: number of areas, m; number of clusters in each area i, C_i; average number of births per cluster, N_{births}

Parameter settings: fixed effect in each area i, b_i

1. Draw number of births, N_{ic}, for each cluster c in area i from $Pois(N_births)$
2. Draw number of deaths, Y_{ic}, for each cluster c in area i from $Pois(N_{ic}exp(b_i + \epsilon_{ic}))$ where $\epsilon_{ic} \sim \mathcal{N}(0, 0.05)$

1.2 IID spatial effect area-level, assuming all births are observed

Population settings: number of areas, m; number of clusters in each area i, C_i; average number of births per cluster, N_{births}

Parameter settings: intercept a, precision τ, hyperparameters μ_τ and σ^2_τ

1. Draw $\tilde{b}$ from $\mathcal{N}(0, \frac{1}{\tau}\mathbf{I})$
2. Draw number of births, N_{ic}, for each cluster c in area i from $Pois(N_{births})$
3. Draw number of deaths, Y_{ic}, for each cluster c in area i from $Pois(N_{ic}exp(a + b_i + \epsilon_{ic}))$ where $\epsilon_{ic} \sim \mathcal{N}(0, 0.05)$

1.3 IID spatial effect stratified cluster-level, assuming all births are observed

Population settings: number of areas, m; number of urban and rural clusters in each area i, C_{Ui} and C_{Ri}; average number of births per urban and rural cluster, $N_{Ubirths}$ and $N_{Rbirths}$

Parameter settings: precision τ, intercepts α_U and α_R, overdispersion parameter d

1. Draw $\tilde{b}$ from $\mathcal{N}(0, \frac{1}{\tau}\mathbf{I})$
2. Draw number of births, N_{ic}, for each cluster c in area i from $Pois(N_{Ubirths})$ or $Pois(N_{Rbirths})$ depending on the cluster
3. Draw number of deaths, Y_{ic}, for each cluster c in area i from $NegBin(\frac{1}{d}N_{ic}exp(\alpha_{UR[c]} + b_i + \epsilon_{ic}), \frac{1}{1+d})$ where $\epsilon_{ic} \sim \mathcal{N}(0, 0.05)$ and $\alpha_{UR[c]}$ is the intercept for the strata which the cluster belongs to

1.4 Sample observed births using 1-stage SRS

Define: number of births sampled per area i, n_i

1. Use one of above procedures to generate underlying truth, (Y_{ic}, N_{ic})
2. Draw observed number of deaths, Z_i, for each area from $Hypergeom(\sum_{c\in\delta(i)} N_{ic}, \sum_{c\in\delta(i)} Y_{ic}, n_i)$

1.5 Sample observed births using 2-stage stratified cluster sampling

TBD

2 Models assuming all births observed

2.3 Estimating $\mathbf{P}(\tilde{\eta}, \alpha, \phi, \tau|\tilde{Y})$ with random spatial effects, admin level model

2.3.1 (Alg 2.3a) Estimating $\mathbf{P}(\tilde{\eta}, \alpha, \phi, \tau|\tilde{Y})$, IID spatial effects, admin level model

Setting:

$log(r_j) = \eta_j = \alpha + b_j$

Priors: $\alpha \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$, $b_j \sim N(0, 1/\tau)$

Hyperpriors: $\tau \sim Gamma(0.01, 0.01)$

$Y_j|r_j \sim Pois(N_j r_j)$

Likelihood:

$$\mathcal{L}(\tilde{\eta}, \alpha, \tau|\tilde{Y}) \propto P(\tilde{Y}|\tilde{\eta}) \times P(\tilde{\eta}|\alpha, \tau) \times P(\alpha) \times P(\tau)$$

$$= P(\tilde{Y}|\tilde{\eta})P(\tilde{\eta}|\alpha, \tau) \times P(\alpha) \times P(\tau)$$

$$= P(\tilde{Y}|\tilde{\eta})P_b(\tilde{\eta} - \alpha|\tau) \times P(\alpha) \times P(\tau)$$

One iteration of MCMC Algorithm:

1. Draw τ from its full conditional, $Gamma(0.01 + \frac{m}{2}, 0.01 + \frac{1}{2}\sum_i(\eta_i - \alpha)^2)$
2. Draw α from its full conditional, a Normal distribution with mean $\frac{\mu_\alpha + \sigma_\alpha^2 \tau \sum_i \eta_i}{1 + \sigma_\alpha^2 \tau m}$, and precision $m\tau + \frac{1}{\sigma_\alpha^2}$.
3. Draw a proposal for $\tilde{\eta}$ from a Taylor approximation to its full conditional, centered on the current state, $\tilde{\pi}(\tilde{\eta}^P|\tilde{\eta}^C, \alpha^C, \tau^C)$, a normal distribution with
   - mean $= \frac{\tau a + Y_i - N_i(1 - \eta_i^C)e^{\eta_i^C)}}{\tau + N_i e^{\eta_i^C}}$
   - precision $= \tau + N_i e^{\eta_i^C}$
4. Accept proposal for $\tilde{\eta}$ with probability

$$\frac{P(\tilde{Y}|\tilde{\eta}^P) \times P_b(\tilde{\eta}^P - \alpha^C|\tau^C)}{P(\tilde{Y}|\tilde{\eta}^C) \times P_b(\tilde{\eta}^C - \alpha^C|\tau^C)} \times \frac{\tilde{\pi}(\tilde{\eta}^C|\tilde{\eta}^P, \alpha^C, \tau^C)}{\tilde{\pi}(\tilde{\eta}^P|\tilde{\eta}^C, \alpha^C, \tau^C)}$$

2.4 Estimating $\mathbf{P}(\tilde{\eta}, \alpha, \phi, \tau|\tilde{Y}, Y_+)$ with random spatial effects, admin level model

2.4.1 (Alg 2.4a) Estimating $\mathbf{P}(\tilde{\eta}, \alpha, \phi, \tau|\tilde{Y}, Y_+)$, IID spatial effects

Setting:

$log(r_j) = \eta_j = \alpha + b_j$

Priors: $\alpha \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$, $b_j \sim N(0, 1/\tau)$

Hyperpriors: $\tau \sim Gamma(0.01, 0.01)$

$Y_+|\tilde{r} \sim Pois(\sum_j N_j r_j)$ where $Y_+ = \sum_j Y_j$

$\tilde{Y}|Y_+, \tilde{r} \sim Multinom(Y_+, \frac{N_j r_j}{\sum_j N_j r_j})$

Likelihood:

$$\mathcal{L}(\tilde{\eta}, \alpha, \tau|\tilde{Y}, Y_+) \propto P(\tilde{Y}|Y_+, \tilde{\eta}) \times P(Y_+|\tilde{\eta}) \times P(\tilde{\eta}|\alpha, \tau) \times P(\alpha) \times P(\tau)$$

$$= P(\tilde{Y}|Y_+, \tilde{\eta})P(Y_+|\tilde{\eta})P_b(\tilde{\eta} - \alpha|\tau) \times P(\alpha) \times P(\tau)$$

Note: MCMC Algorithm works out to be the same as Alg 2.3a.

2.6 Estimating $\mathbf{P}(\alpha_U, \alpha_R, \tilde{b}, \tau, d|\tilde{Y}, Y_+)$, IID stratified cluster level model

Setting:

$log(r_j) = \eta_j = \alpha + b_j$

Priors: $\alpha \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$, $b_j \sim N(0, 1/\tau)$

Hyperpriors: $\tau \sim Gamma(0.01, 0.01)$

$Y_+|\tilde{r}, d \sim Negbin(\frac{1}{d}\sum_i\sum_j N_{ij}r_{ij}, \frac{1}{1+d})$ where $Y_+ = \sum_i\sum_j Y_{ij}$ and $d \sim logNormal(\mu_d, \sigma_d^2)$

$\tilde{Y}|Y_+, \tilde{r} \sim Multinom(Y_+, \frac{N_{ij}r_{ij}}{\sum_i\sum_j N_{ij}r_{ij}})$

Likelihood:

$$\mathcal{L}(\alpha_U, \alpha_R, \tilde{b}, \tau, d|\tilde{Y}, Y_+) \propto P(\tilde{Y}|Y_+, \tilde{b}, \alpha_U, \alpha_R) \times P(Y_+|\tilde{b}, \alpha_U, \alpha_R, d) \times P(\tilde{b}|\tau) \times \pi(\alpha, \tau, d)$$

One iteration of MCMC Algorithm:

1. Draw τ from its full conditional, $Gamma(m/2 + a_\tau, \frac{1}{2}\sum_i^m b_i^2 + b_\tau)$
2. Draw proposal for $\tilde{b}$ from $\mathcal{N}(\tilde{b}^C + k_b\nabla\log\pi(\tilde{b}^C|\alpha_{UR}^C, d^C, \tau^C, \tilde{Y}, Y_+), diag(2k_b))$
3. Draw proposal for d from $logNormal(\log(d^C) + k_{logd}\nabla\log\pi(logd^C|\alpha_{UR}^C, \tilde{b}^C, \tilde{Y}, Y_+), 2k_{logd})$
4. Draw proposal for α_U from $\mathcal{N}(\alpha_U^C + k_\alpha\nabla\log\pi(\alpha_U^C|\alpha_R^C, \tilde{b}^C, d^C, \tilde{Y}, Y_+), 2k_\alpha)$ and proposal for α_R from $\mathcal{N}(\alpha_R^C + k_\alpha\nabla\log\pi(\alpha_R^C|\alpha_U^C, \tilde{b}^C, d^C, \tilde{Y}, Y_+), 2k_\alpha)$
5. Accept proposal for $(\tilde{b}, d, \alpha_U, \alpha_R)$ with probability

$$\frac{P(\tilde{Y}|Y_+, \alpha_U^P, \alpha_R^P, \tilde{b}^P)P(Y_+|\alpha_U^P, \alpha_R^P, d^P)P(\tilde{b}^P|\tau^C)\pi(\alpha_U^P, \alpha_R^P, d^P)}{P(\tilde{Y}|Y_+, \alpha_U^C, \alpha_R^C, \tilde{b}^C)P(Y_+|\alpha_U^C, \alpha_R^C, d^C)P(\tilde{b}^C|\tau^C)\pi(\alpha_U^C, \alpha_R^C, d^C)} \times \frac{\pi(\alpha_{UR}^C, \tilde{b}^C, d^C|\alpha_{UR}^P, \tilde{b}^P, d^P, \tau^C, \tilde{Y}, Y_+)}{\pi(\alpha_{UR}^P, \tilde{b}^P, d^P|\alpha_{UR}^C, \tilde{b}^C, d^C, \tau^C, \tilde{Y}, Y_+)}$$

Note:

- $\nabla\log\pi(b_i|b_{-i}, \alpha_{UR}, d, \tau, \tilde{Y}, Y_+) =$
  $\sum_{j\in C_i} Y_{ij} + \sum_{j\in C_i} N_{ij}r_{ij}\{\frac{1}{d}[\psi(Y_+ + \frac{1}{d}\sum_{i,j} N_{ij}r_{ij}) - \psi(\frac{1}{d}\sum_{i,j} N_{ij}r_{ij}) - \log(d+1)] - \frac{Y_+}{\sum_{i,j} N_{ij}r_{ij}}\} - \tau b_i$
- $\nabla\log\pi(\alpha_U|\tilde{b}, \alpha_R, d, \mu_U, \sigma_U^2, \tilde{Y}, Y_+) =$
  $\sum_{i,j\in U} Y_{ij} + \sum_{i,j\in U} N_{ij}r_{ij}\{\frac{1}{d}[\psi(Y_+ + \frac{1}{d}\sum_{i,j} N_{ij}r_{ij}) - \psi(\frac{1}{d}\sum_{i,j} N_{ij}r_{ij}) - \log(d+1)] - \frac{Y_+}{\sum_{i,j} N_{ij}r_{ij}}\} - \frac{\alpha_U - \mu_U}{\sigma_U^2}$
- $\nabla\log\pi(d|\alpha_{UR}, \tilde{b}, \tilde{Y}, Y_+, \sigma_d)$
  $= \frac{1}{d+1}(Y_+ - \sum_{i,j} N_{ij}r_{ij}) + \frac{1}{d}\sum_{i,j} N_{ij}r_{ij}\{\psi(\frac{1}{d}\sum_{i,j} N_{ij}r_{ij}) - \psi(Y_+ + \frac{1}{d}\sum_{i,j} N_{ij}r_{ij}) + \log(1+d)\} - \frac{log(d) - \mu_d}{\sigma_d^2}$

3 Admin 1 level models with 1-stage SRS sampling

3.2 Estimating P($\tilde{Y}, \alpha, \tilde{\eta}, \phi, \tau | \tilde{Z}, Y_+$) (assuming we know total deaths) BYM2

Setting:

$log(r_j) = \eta_j = \alpha + b_j$

Priors: $\alpha \sim \mathcal{N}(\mu_\alpha, \sigma^2_\alpha)$, $b \sim BYM2(\tau, \phi)$

Hyperpriors: $\tau \sim Gamma(0.01, 0.01)$, $\phi \sim Beta(2, 3)$

$Y_+ | \tilde{r} \sim Pois(\sum_j N_j r_j)$ where $Y_+ = \sum_j Y_j$

$\tilde{Y} | Y_+, \tilde{r} \sim Multinom(Y_+, \frac{\tilde{N}\tilde{r}}{\sum_j N_j r_j})$

$Z_j | Y_j \sim Hypergeom(N_j, Y_j, n_j)$

Likelihood:

$$\mathcal{L}(\tilde{Y}, \tilde{\eta}, \alpha, \phi, \tau | \tilde{Z}, Y_+) \propto P(\tilde{Z}, Y_+ | \tilde{Y}, \tilde{\eta}) \times P(\tilde{Y} | \tilde{\eta}) \times P(\tilde{\eta} | \alpha, \phi, \tau) \times \pi(\alpha)\pi(\phi)\pi(\tau)$$

$$= P(\tilde{Z} | Y_+, \tilde{Y}, \tilde{\eta}) \times P(Y_+ | \tilde{Y}, \tilde{\eta}) \times P(\tilde{Y} | \tilde{\eta}) \times P(\tilde{\eta} | \alpha, \phi, \tau) \times \pi(\alpha)\pi(\phi)\pi(\tau)$$

$$= P(\tilde{Z} | \tilde{Y}) \times P(Y_+, \tilde{Y} | \tilde{\eta}) \times P(\eta | \alpha, \phi, \tau) \times \pi(\alpha)\pi(\phi)\pi(\tau)$$

$$= P(\tilde{Z} | \tilde{Y}) \times P(\tilde{Y} | Y_+, \tilde{\eta}) \times P(Y_+ | \tilde{\eta}) \times P_b(\tilde{\eta} - \alpha | \phi, \tau) \times \pi(\alpha)\pi(\phi)\pi(\tau)$$

One iteration of MCMC Algorithm:

Steps 1-5 are the same as Alg 3.1.

6. Draw proposal for each Y_i
   - Randomly pair all areas with another area (without replacement)
   - For each pair (Y_{k1}, Y_{k2}), draw a natural number J from $\{0, ..., f_Y\}$ and then increase and decrease Y_{k1} and Y_{k2} by that J, respectively
7. Accept proposal for $\tilde{Y}$ with probability

$$\frac{\Pi_i P(Z_i | Y_i^P) P(\tilde{Y}^P | Y_+, \eta^C)}{\Pi_i P(Z_i | Y_i^C) P(\tilde{Y}^C | Y_+, \eta^C)}$$

4 Models with 1-stage stratified cluster sampling

4.1 Estimating $P(\alpha_U, \alpha_R, \tilde{b}, \tau, d|Y_+, \tilde{Y}^{(s)}, \delta)$ IID spatial effect

Setting:

$log(r_{ij}) = \eta_{ij} = \alpha_U I(j \in urban) + \alpha_R I(j \in rural) + b_i$

Priors: $\alpha_U \sim \mathcal{N}(\mu_U, \sigma_U^2), \alpha_R \sim \mathcal{N}(\mu_R, \sigma_R^2), \tilde{b} \sim N(0, \frac{1}{\tau}\mathbf{I})$

Hyperpriors: $\tau \sim Gamma(0.01, 0.01)$

$Y_+|\tilde{r}, d \sim Negbin(\frac{1}{d}\sum N_{ij}r_{ij}, \frac{1}{1+d})$ where $d \sim logNormal(\mu_d, \sigma_d^2)$

$(\tilde{Y}^{(s)}, Y_+ - Y_+^{(s)})|Y_+, \tilde{r}, \tilde{\delta} \sim Multinom(Y_+, (\frac{N_{ij}r_{ij}}{\sum_{i,j} N_{ij}r_{ij}}, \frac{\sum_{i,j:\delta_{ij}=0} N_{ij}r_{ij}}{\sum_{i,j} N_{ij}r_{ij}}))$

$P(\tilde{Y}^{(s)}|Y_+, \tilde{r}, \tilde{\delta}) = \sum_w P(\tilde{Y}^{(s)}, w|Y_+, \tilde{r}, \tilde{\delta}) = P(\tilde{Y}^{(s)}, Y_+ - Y_+^{(s)}|Y_+, \tilde{r}, \tilde{\delta})$

Likelihood:

$$\mathcal{L}(\alpha_{UR}, \tilde{b}, \tau, d|Y_+, \tilde{Y}^{(s)}, \delta) \propto P(\tilde{Y}^{(s)}|Y_+, \alpha_{UR}, \tilde{b}, \tilde{\delta})P(Y_+|\alpha_{UR}, \tilde{b}, d)P(\tilde{b}|\tau)\pi(\alpha_{UR}, \tau, d)$$

One iteration of MCMC Algorithm:

1. Draw τ from its full conditional, $Gamma(m/2 + a_\tau, \frac{1}{2}\sum_i^m b_i^2 + b_\tau)$
2. Draw proposal for $\tilde{b}$ from $\mathcal{N}(\tilde{b}^C + k_b \nabla \log \pi(\tilde{b}^C|\alpha_{UR}^C, d^C, \tau^C, \tilde{Y}^{(s)}, Y_+), diag(2k_b))$
3. Draw proposal for d from $logNormal(\log(d^C) + k_{logd} \nabla \log \pi(logd^C|\alpha_{UR}^C, \tilde{b}^C, \tilde{Y}^{(s)}, Y_+), 2k_{logd})$
4. Draw proposal for α_U from $\mathcal{N}(\alpha_U^C + k_\alpha \nabla \log \pi(\alpha_U^C|\alpha_R^C, \tilde{b}^C, d^C, \tilde{Y}^{(s)}, Y_+), 2k_\alpha)$ and proposal for α_R from $\mathcal{N}(\alpha_R^C + k_\alpha \nabla \log \pi(\alpha_R^C|\alpha_U^C, \tilde{b}^C, d^C, \tilde{Y}^{(s)}, Y_+), 2k_\alpha)$
5. Accept proposal for $(\tilde{b}, d, \alpha_U, \alpha_R)$ with probability

$$\frac{P(\tilde{Y}^{(s)}|Y_+, \alpha_U^P, \alpha_R^P, \tilde{b}^P)P(Y_+|\alpha_U^P, \alpha_R^P, d^P)P(\tilde{b}^P|\tau^C)\pi(\alpha_U^P, \alpha_R^P, d^P)}{P(\tilde{Y}^{(s)}|Y_+, \alpha_U^C, \alpha_R^C, \tilde{b}^C)P(Y_+|\alpha_U^C, \alpha_R^C, d^C)P(\tilde{b}^C|\tau^C)\pi(\alpha_U^C, \alpha_R^C, d^C)} \times \frac{\pi(\alpha_{UR}^C, \tilde{b}^C, d^C|\alpha_{UR}^P, \tilde{b}^P, d^P, \tau^C, \tilde{Y}^{(s)}, Y_+)}{\pi(\alpha_{UR}^P, \tilde{b}^P, d^P|\alpha_{UR}^C, \tilde{b}^C, d^C, \tau^C, \tilde{Y}^{(s)}, Y_+)}$$

Note:

- $\nabla \log \pi(b_i|b_{-i}, \alpha_{UR}, d, \tau, \tilde{Y}, Y_+) =$
$\sum_{j\in C_i, \delta_{ij}=1} Y_{ij}^{(s)} - \frac{Y_+ \sum_{j\in C_i} N_{ij}r_{ij}}{\sum_{i,j} N_{ij}r_{ij}} + \frac{(Y_+ - Y_+^{(s)})\sum_{j\in C_i:\delta_{ij}=0} N_{ij}r_{ij}}{\sum_{i,j:\delta_{ij}=0} N_{ij}r_{ij}} + \frac{1}{d}\sum_{j\in C_i} N_{ij}r_{ij}\{\psi(Y_+ + \frac{1}{d}\sum_{i,j} N_{ij}r_{ij}) - \psi(\frac{1}{d}\sum_{i,j} N_{ij}r_{ij}) - \log(d+1)\} - \tau b_i$
- $\nabla \log \pi(\alpha_U|\tilde{b}, \alpha_R, d, \mu_U, \sigma_U^2, \tilde{Y}, Y_+) =$
$\sum_{j\in U, \delta_{ij}=1} Y_{ij}^{(s)} - \frac{Y_+ \sum_{j\in U} N_{ij}r_{ij}}{\sum_{i,j} N_{ij}r_{ij}} + \frac{(Y_+ - Y_+^{(s)})\sum_{j\in U:\delta_{ij}=0} N_{ij}r_{ij}}{\sum_{i,j:\delta_{ij}=0} N_{ij}r_{ij}} + \frac{1}{d}\sum_{i,j\in U} N_{ij}r_{ij}\{\psi(Y_+ + \frac{1}{d}\sum_{i,j} N_{ij}r_{ij}) - \psi(\frac{1}{d}\sum_{i,j} N_{ij}r_{ij}) - \log(d+1)\} - \frac{\alpha_U - \mu_U}{\sigma_U^2}$
- $\nabla \log \pi(d|\alpha_{UR}, \tilde{b}, \tilde{Y}, Y_+, \sigma_d)$
$= \frac{1}{d+1}(Y_+ - \sum_{i,j} N_{ij}r_{ij}) + \frac{1}{d}\sum_{i,j} N_{ij}r_{ij}\{\psi(\frac{1}{d}\sum_{i,j} N_{ij}r_{ij}) - \psi(Y_+ + \frac{1}{d}\sum_{i,j} N_{ij}r_{ij}) + \log(1+d)\} - \frac{log(d) - \mu_d}{\sigma_d^2}$

5 Models with two-stage stratified cluster sampling - Not Yet Implemented

Sampling Scheme (DHS): The population is split into strata: admin 2 area crossed with urban/rural

Stage 1: For each strata h, a_h clusters are selected, each with probability $\frac{a_h M_{h_i}}{\sum M_{h_i}}$, where M_{h_i} is the total number of households in cluster h_i and a_h is different fixed number for each strata.

Stage 2: For each urban cluster h_{Ui}, g_U households are selected with equal probability, and for each rural cluster h_{Ri}, g_R households are selected with equal probability.

Note: It is unclear how g_U, g_R, and a_hs are chosen exactly, but it is loosely related to the total number of households/clusters in sampling frame.

Some notation:

- i indexes admin 2 areas, j indexes clusters within admin 2 areas
- N_{ij} and Y_{ij} denotes total births and deaths in admin area i, cluster j, respectively
- Y_{i+} denotes the total deaths in admin area i, and Y_+ denotes the sum of these totals over all areas
- n_{ij} and Z_{ij} denotes sampled births and deaths in admin area i, cluster j, respectively
- δ_{ij} is an indicator variable denoting whether cluster j in admin area i is sampled

5.1 Estimating $\mathbf{P}(\alpha_U, \alpha_R, \tilde{b}, \tau, d, \tilde{Y}^{(s)}|Y_+, \tilde{Z}, \tilde{\delta})$ IID model

Setting:

$log(r_{ij}) = \eta_{ij} = \alpha_U I(j \in urban) + \alpha_R I(j \in rural) + b_i$

Priors: $\alpha_U \sim \mathcal{N}(\mu_U, \sigma_U^2), \alpha_R \sim \mathcal{N}(\mu_R, \sigma_R^2), \tilde{b} \sim N(0, \frac{1}{\tau}\mathbf{I})$

Hyperpriors: $\tau \sim Gamma(0.01, 0.01)$

$Y_+|\tilde{r}, d \sim Negbin(\frac{1}{d}\sum_{i,j} N_{ij}r_{ij}, \frac{1}{1+d})$ where $d \sim logNormal(\mu_d, \sigma_d^2)$

$\tilde{Y}^{(s)}, Y_+ - Y_+^{(s)}|Y_+, r_{ij} \sim Multinom(Y_+, (\frac{N_{ij}r_{ij}}{\sum_{i,j} N_{ij}r_{ij}}, \frac{\sum_{\delta_{ij}=0} N_{ij}r_{ij}}{\sum_{i,j} N_{ij}r_{ij}}))$

$Z_{ij}|Y_{ij}, \delta_{ij} = 1 \sim Hypergeom(N_{ij}, Y_{ij}, n_{ij})$

$P(Z_{ij}|Y_{ij}, \delta_{ij} = 0) = I(Z_{ij} = 0)$

Likelihood:

$$\mathcal{L}(\tilde{Y}^{(s)}, \alpha_U, \alpha_R, \tilde{b}, \tau, d|Y_+, \tilde{Z}, \tilde{\delta}) \propto P(\tilde{Z}|\tilde{Y}^{(s)}, \tilde{\delta})P(\tilde{Y}^{(s)}|Y_+, \alpha_U, \alpha_R, \tilde{b}, d)P(Y_+|\alpha_U, \alpha_R, \tilde{b}, d)P(\tilde{b}|\tau)\pi(\alpha_U, \alpha_R, \tau, d)$$

$$= \Pi_{i,j:\delta_{ij}=1}\{P(\tilde{Z}_{ij}|\tilde{Y}_{ij}^{(s)}, \delta_{ij} = 1)\}P(\tilde{Y}^{(s)}, Y_+ - Y_+^{(s)}|Y_+, \alpha_U, \alpha_R, \tilde{b}, d)P(Y_+|\alpha_U, \alpha_R, \tilde{b}, d)P(\tilde{b}|\tau)\pi(\alpha_U, \alpha_R, \tau, d)$$

Potential issue: This would require estimating a latent parameter, Y, for every sampled cluster (on the order of 100)

5.2 Estimating $\mathbf{P}(\alpha_U, \alpha_R, \tilde{b}, \tau, d | Y_+, \tilde{Z}, \tilde{\delta})$ IID model

Setting:

$log(r_{ij}) = \eta_{ij} = \alpha_U I(j \in urban) + \alpha_R I(j \in rural) + b_i$

Priors: $\alpha_U \sim \mathcal{N}(\mu_U, \sigma_U^2), \alpha_R \sim \mathcal{N}(\mu_R, \sigma_R^2), \tilde{b} \sim N(0, \frac{1}{\tau}\mathbf{I})$

Hyperpriors: $\tau \sim Gamma(0.01, 0.01)$

$Y_+ | \tilde{r}, d \sim Negbin(\frac{1}{d}\sum_{i,j} N_{ij} r_{ij}, \frac{1}{1+d})$ where $d \sim logNormal(\mu_d, \sigma_d^2)$

$\tilde{Y}^{(s)}, Y_+ - Y_+^{(s)} | Y_+, r_{ij} \sim Multinom(Y_+, (\frac{N_{ij} r_{ij}}{\sum_{i,j} N_{ij} r_{ij}}, \frac{\sum_{\delta_{ij}=0} N_{ij} r_{ij}}{\sum_{i,j} N_{ij} r_{ij}}))$

$Z_{ij} | Y_{ij}, \delta_{ij} = 1 \sim Hypergeom(N_{ij}, Y_{ij}, n_{ij})$

$P(Z_{ij} | Y_{ij}, \delta_{ij} = 0) = I(Z_{ij} = 0)$

Likelihood:

$$\mathcal{L}(\alpha_U, \alpha_R, \tilde{b}, \tau, d | Y_+, \tilde{Z}, \tilde{\delta}) \propto P(\tilde{Z} | \tilde{\delta}, Y_+, \alpha_U, \alpha_R, \tilde{b}, d) P(Y_+ | \alpha_U, \alpha_R, \tilde{b}, d) P(\tilde{b} | \tau) \pi(\alpha_U, \alpha_R, \tau, d)$$

where, the first term,

$$P(\tilde{Z} | \tilde{\delta}, Y_+, \alpha_U, \alpha_R, \tilde{b}, d) = \sum_{\tilde{y} \in V(Y_+)} \{\Pi_i^m \Pi_{j \in C(i):\delta_{ij}=1} \{P(Z_{ij} | \delta_{ij} = 1, y_{ij})\} \times P(\tilde{y}, Y_+ - \Sigma\tilde{y} | Y_+, \alpha_U, \alpha_R, \tilde{b})\}$$

where $V(Y_+)$ is the set of vectors, each of length equal to the number of observed clusters, whose sum is less than or equal to Y_+

Potential issue: $V(Y_+)$ is a very large set which grows exponentially with Y_+

5.3 Estimating $\mathbf{P}(\alpha_U, \alpha_R, \tilde{b}, \tau, d, \tilde{Y}_{i+} | Y_+, \tilde{Z}, \tilde{\delta})$ IID model

Setting:

$log(r_{ij}) = \eta_{ij} = \alpha_U I(j \in urban) + \alpha_R I(j \in rural) + b_i$

Priors: $\alpha_U \sim \mathcal{N}(\mu_U, \sigma_U^2), \alpha_R \sim \mathcal{N}(\mu_R, \sigma_R^2), \tilde{b} \sim N(0, \frac{1}{\tau}\mathbf{I})$

Hyperpriors: $\tau \sim Gamma(0.01, 0.01)$

$Y_+ | \tilde{r}, d \sim Negbin(\frac{1}{d}\sum_{i,j} N_{ij} r_{ij}, \frac{1}{1+d})$ where $d \sim logNormal(\mu_d, \sigma_d^2)$

$\tilde{Y}_{i+} | Y_+, r_{ij} \sim Multinom(Y_+, \frac{\sum_{j \in C_i} N_{ij} r_{ij}}{\sum_{i,j} N_{ij} r_{ij}})$ i.e. area-level multinomial

$\tilde{Y}_i | Y_{i+}, r_{ij} \sim Multinom(Y_{i+}, \frac{N_{ij} r_{ij}}{\sum_{j \in C_i} N_{ij} r_{ij}})$ i.e. a separate multinomial for each area

$Z_{ij} | Y_{ij}, \delta_{ij} = 1 \sim Hypergeom(N_{ij}, Y_{ij}, n_{ij})$

$P(Z_{ij} | Y_{ij}, \delta_{ij} = 0) = I(Z_{ij} = 0)$

Likelihood:

$$\mathcal{L}(\tilde{Y}_{i+}, \alpha_U, \alpha_R, \tilde{b}, \tau, d | Y_+, \tilde{Z}, \tilde{\delta}) \propto P(\tilde{Z} | \tilde{\delta}, \tilde{Y}_{i+}, \alpha_U, \alpha_R, \tilde{b}, d) \times P(\tilde{Y}_{i+} | Y_+, \alpha_U, \alpha_R, \tilde{b}, d) \times P(Y_+ | \alpha_U, \alpha_R, \tilde{b}, d) P(\tilde{b} | \tau) \pi(\alpha_U, \alpha_R, \tau, d)$$

where, the first component,

$$P(\tilde{Z} | \tilde{\delta}, \tilde{Y}_{i+}, \alpha_U, \alpha_R, \tilde{b}, d) \propto \Pi_i^m \sum_{\tilde{y}_i \in V(Y_{i+})} P(\tilde{Z}_i | \tilde{y}_i, \tilde{\delta}) P(\tilde{y}_i | Y_{i+}, \alpha_U, \alpha_R, \tilde{b}, d)$$

$$= \Pi_i^m \sum_{\tilde{y}_i \in V(Y_{i+})} \{\Pi_{j \in C_i} P(\tilde{Z}_{ij} | y_{ij}, \tilde{\delta}) \times P(\tilde{y}_i | Y_{i+}, \alpha_U, \alpha_R, \tilde{b}, d)\}$$

where $V(Y_{i+})$ is the set of vectors, each of length equal to the number of observed clusters in area i, whose sum is less than or equal to Y_{i+}

Potential issue: This setup has the same issues as 6.1 and 6.2, but by combining both methods and 'meeting in the middle' it lessens the burden of both issues because the number of latent Ys is reduced the number of areas and each $V(Y_{i+})$ is a smaller set (though I'm not sure small enough).