Conditional Poisson MCMC

September 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 $\epsilon_{ic} \sim \mathcal{N}(0, 0.05)$

Assumptions: (1) each cluster contains similar number of births, (2) all births (and deaths) are observed

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, hyperparameters μ_τ and σ^2_τ

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

Assumptions: Same as 1.1

1.3 BYM2 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, neighborhood matrix Q, and hyperparameters, $\mu_\tau, \sigma^2_\tau, a_\phi, b_\phi$

1. Draw τ_b from $Gamma(a_\tau, b_\tau)$
2. Draw ϕ from $Beta(a_\phi, b_\phi)$
3. Calculate $\Sigma_b = \frac{1}{\tau}((1 - \phi)\mathbf{I} + \phi Q^{-*})$
   where Q^* is the scaled neighborhood matrix (via inla.scale.model)
4. Draw $\tilde{b}$ from $\mathcal{N}(0, \Sigma_b)$
5. Draw number of births, N_{ic}, for each cluster c in area i from $Pois(N_{births})$
6. Draw number of deaths, Y_{ic}, for each cluster c in area i from $Pois(N_{ic}exp(a + b_i + \epsilon_{ic}))$ where $\epsilon_{ic} \sim \mathcal{N}(0, 0.05)$

Assumptions: Same as 1.1

1.4 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:

1. Draw precision, τ, from $Gamma(a_\tau, b_\tau)$
2. Draw intercepts, α_U and α_R, from $N(\mu_U, \sigma^2_U)$ and $N(\mu_R, \sigma^2_R)$, respectively
3. Draw overdispersion parameter, d, from $HalfNormal(\sigma = 5)$
4. Draw $\tilde{b}$ from $\mathcal{N}(0, \frac{1}{\tau}\mathbf{I})$

5. 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
6. 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

Assumptions: Same as 1.1

1.5 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 $Hypergeom(\sum_{c\in\delta(i)} N_{ic}, \sum_{c\in\delta(i)} Y_{ic}, n_i)$

1.6 Sample observed births using 2-stage stratified cluster sampling

TBD

2 Algorithms assuming all births observed

2.1 Estimating P$(\tilde{r}|\tilde{Y})$, fixed spatial effects, admin level model

Setting:
$log(r_j) = b_j$
Priors: $b_j \sim \mathcal{N}(\mu_b, \sigma_b^2)$
$Y_j|r_j \sim Pois(N_j r_j)$
Likelihood:

$$\mathcal{L}(\tilde{r}|\tilde{Y}) \propto P(\tilde{Y}|\tilde{r})P(r) = \Pi_{i=1}^{m} P(Y_i|r_i) \times P(r_i)$$

$$= \Pi_{i=1}^{m} P(Y_i|r_i) \times |\frac{\partial b_i}{\partial r_i}|P(b_i)$$

where m is the number of areas.
One iteration of MCMC Algorithm:
For each area j,

1. Draw b_j^P from $\tilde{\pi}(b_j^P|b_j^C) = N(b_j^C, \sigma^2_{prop,b})$
2. Accept proposal with probability

$$\frac{\mathcal{L}(\tilde{r}_j^P|\tilde{Y})}{\mathcal{L}(\tilde{r}_j^C|\tilde{Y})} \times \frac{\tilde{\pi}(b_j^C|b_j^P)}{\tilde{\pi}(b_j^P|b_j^C)}$$

Note: There is also code for a single block update algorithm that is nearly the same (because the parameters are independent).

2.2 Estimating $P(\tilde{r}|\tilde{Y}, Y_+)$, fixed spatial effects, admin level model

Setting:
$log(r_j) = b_j$
Priors: $b_j \sim \mathcal{N}(\mu_r, \sigma_b^2)$
$Y_j|r_j \sim Pois(N_j r_j)$
$Y_+|\tilde{r} \sim Pois(\sum_j N_j r_j)$ 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:

$$\mathcal{L}(\tilde{r}, |\tilde{Y}, Y_+) \propto P(\tilde{Y}|Y_+, \tilde{r}) \times P(Y_+|\tilde{r}) \times \Pi_i^m |\frac{\partial b_i}{\partial r_i}|P(b_i)$$

where m is the number of areas.

One iteration of MCMC Algorithm:
For each area j,

1. Draw b_j^P from $\tilde{\pi}(b_j^P|b_j^C) = N(b_j^C, \sigma^2_{prop,b})$
2. Accept proposal with probability

$$\frac{\mathcal{L}(r_j^P, \tilde{r}_{-j}^C, |\tilde{Y}, Y_+)}{\mathcal{L}(\tilde{r}^C, |\tilde{Y}, Y_+)} \times \frac{\tilde{\pi}(b_j^C|b_j^P)}{\tilde{\pi}(b_j^P|b_j^C)}$$

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

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 $\frac{\mu_\alpha + \sigma^2_\alpha \tau \sum_i \eta_i}{1 + \sigma^2_\alpha \tau m}$, and precision $m\tau + \frac{1}{\sigma^2_\alpha}$.
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

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

2.3.2 (Alg 2.3) Estimating $\mathbf{P}(\tilde{\eta}, \alpha, \phi, \tau | \tilde{Y})$, BYM2 spatial effects, admin level model

Setting:
$log(r_j) = \eta_j = \alpha + b_j$
Priors: $\alpha \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$, $b \sim BYM2(\tau, \phi)$
Hyperpriors: $\tau \sim Gamma(a_\tau, b_\tau)$, $\phi \sim Beta(a_\phi, b_\phi)$
$Y_j | r_j \sim Pois(N_j r_j)$

Likelihood:

$$\mathcal{L}(\tilde{\eta}, a, \phi, \tau | \tilde{Y}) \propto P(\tilde{Y} | \tilde{\eta}) \times P(\tilde{\eta} | a, \phi, \tau) \times P(a) \times P(\phi) P(\tau)$$
$$= \mathcal{L}(\tilde{\eta}, a, \phi, \tau | \tilde{Y}) \propto P(\tilde{Y} | \tilde{\eta}) P_b(\tilde{\eta} - a | \phi, \tau) \times P(a) \times P(\phi) P(\tau)$$

One iteration of MCMC Algorithm:

1. Draw a proposal for $logit(\phi)$ from $N(logit(\phi^C), \sigma_\phi^2)$
   Note: This implies the proposal distribution, $\pi(\phi^P | \phi^C) = \frac{1}{\phi^P(1-\phi^P)} \times \mathcal{N}(logit(\phi^C), \sigma_\phi^2)$

2. Draw a proposal for τ from $logNormal(log(\tau^C), \sigma_\tau^C)$

3. Draw proposal for $\tilde{\eta}$ from a Taylor approximation to its full conditional, centered on the the current state of η, $\tilde{\pi}(\tilde{\eta}^P | \tilde{\eta}^C, \phi^P, \tau^P, Y)$, which is a multivariate Gaussian distribution with

   - mean vector, $[Y^T + a\tau\mathbf{1}^T W^{-1}(\phi) - C(\eta^C)](D(\eta^C) + \tau W^{-1}(\phi))^{-1}$
   - precision matrix, $D(\eta^C) + \tau W^{-1}(\phi)$

   where $C(\eta^*)$ is a $1 \times m$ vector with i^{th} entry, $N_i e^{\eta_i^*}(1 - \eta_i^*)$
   and $D(\eta^*)$ is a $m \times m$ diagonal matrix with i^{th} diagonal element, $N_i e^{\eta_i^*}$

4. Accept proposal for (ϕ, τ, η) with probability

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

5. Draw a from its full conditional, a normal distribution with mean, $\frac{\mu_a + \sigma_a^2 \tau \tilde{\eta}^T W^{-1}(\phi)\mathbf{1}}{1 + \sigma_a^2 \tau \mathbf{1}^T W^{-1}(\phi)\mathbf{1}}$, and precision, $\frac{1}{\sigma_a^2} + \tau\mathbf{1}^T W^{-1}(\phi)\mathbf{1}$, where $W(\phi) = (1-\phi)I + \phi Q^-$

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.

2.4.2 (Alg 2.4) Estimating $P(\tilde{\eta}, \alpha, \phi, \tau|\tilde{Y}, Y_+)$, BYM2 spatial effects

Setting:
$log(r_j) = \eta_j = \alpha + b_j$
Priors: $\alpha \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$, $b \sim BYM2(\tau, \phi)$
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 $Y_+ = \sum_j Y_j$
$\tilde{Y}|Y_+, \tilde{r} \sim Multinom(Y_+, \frac{N_j r_j}{\sum_j N_j r_j})$

Likelihood:

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

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

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

2.5 Estimating $\mathbf{P}(\alpha_U, \alpha_R, \tilde{b}, \tau, d|\tilde{Y})$, IID spatial model accounting for stratification and clustering

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

Setting:

$log(r_{ij}) = \eta_{ij} = \alpha_U I(j \in urban) + \alpha_R I(j \in rural) + b_i$
Priors: $\alpha_U \sim \mathcal{N}(\mu_U, \sigma_U^2), \alpha_R \sim \mathcal{N}(\mu_R, \sigma_R^2), b \sim^{IID} N(0, 1/\tau)$
Hyperpriors: $\tau \sim Gamma(a_\tau, b_\tau)$
$Y_{ij}|r_{ij}, d \sim Negbin(\frac{1}{d}N_{ij}r_{ij}, \frac{1}{1+d})$ where $d \sim Half\mathcal{N}(0, \sigma_d^2)$

Likelihood:

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

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 $\mathcal{N}(\tilde{b}^C, diag(\sigma_b^2))$
3. Draw proposal d from $logNormal(log(d^C), \sigma_d^2)$
4. Draw proposals for α_U and α_R from $\mathcal{N}(\alpha_U^C, \sigma_U^2)$ and $\mathcal{N}(\alpha_R^C, \sigma_R^2)$, respectively
5. Accept proposal for $(\tilde{b}, d, \alpha_U, \alpha_R)$ with probability

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

3 Admin 1 level models with 1-stage SRS sampling

3.1 Estimating $\mathbf{P}(\tilde{Y}, \alpha, \tilde{\eta}, \phi, \tau | \tilde{Z})$ (not conditioning on total deaths) BYM2

Setting:
$log(r_j) = \eta_j = \alpha + b_j$
Priors: $\alpha \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$, $b \sim BYM2(\tau, \phi)$
Hyperpriors: $\tau \sim Gamma(0.01, 0.01)$, $\phi \sim Beta(2, 3)$
$Y_j | r_j \sim Pois(N_j r_j)$
$Z_j | Y_j \sim Hypergeom(N_j, Y_j, n_j)$
Likelihood:

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

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

$$= \Pi_i^m \{P(Z_i | Y_i) P(Y_i | \eta_i)\} \times P_b(\tilde{\eta} - \alpha | \phi, \tau) \pi(\alpha) \pi(\phi) \pi(\tau)$$

One iteration of MCMC Algorithm:
Steps 1-5 are the same as Alg 2.3, except that we use Y^C instead of Y given by the data.

6. Draw proposal for each Y_i from discrete uniform distribution with minimum $Y_i^C - f_Y$ and maximum $Y_i^C + f_Y$, where f_Y is some natural number.

7. Accept proposal for $\tilde{Y}$ with probability

$$\frac{\Pi_i P(Z_i | Y_i^P) P(Y_i^P | \eta_i^C)}{\Pi_i P(Z_i | Y_i^C) P(Y_i^C | \eta_i^C)}$$

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

Setting:
$log(r_j) = \eta_j = \alpha + b_j$
Priors: $\alpha \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$, $b \sim BYM2(\tau, \phi)$
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

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

3.3 Estimating $\mathbf{P}(\tilde{Y}, Y_+, \alpha, \tilde{\eta}, \phi, \tau | \tilde{Z}, \hat{Y}_+)$ (assuming we have an estimate of total deaths) BYM2

Setting:

$log(r_j) = \eta_j = \alpha + b_j$
Priors: $\alpha \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$, $b \sim BYM2(\tau, \phi)$
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 $Y_+ = \sum_j Y_j$
$\tilde{Y} | Y_+, \tilde{r} \sim Multinom(Y_+, \frac{\tilde{N}\tilde{r}}{\sum_j N_j r_j})$
$Z_j | Y_j \sim Hypergeom(N_j, Y_j, n_j)$
$\hat{Y}_+ | Y_+ \sim \mathcal{N}(Y_+, \hat{\sigma}^2_{Y_+})$

Calculation of HT estimator, $\hat{Y}_+$, and estimated variance, $\hat{\sigma}^2_{Y_+}$

Note the inclusion probability of an individual in area i is $\pi_i = \frac{n_i}{N_i}$ and the inclusion probability of two individuals in area i is $\pi_{ii} = \frac{n_i(n_i-1)}{N_i(N_i-1)}$. Then it follows,

$\hat{Y}_+ = \sum_i^m \sum_j^{n_i} \frac{Z_{ij}}{\pi_{ij}} = \sum_i^m \frac{N_i}{n_i} Z_i$
$\hat{\sigma}^2_{Y_+} = \sum_i^m \sum_j^{n_i} \frac{1-\pi_i}{\pi_i^2} Z_{ij}^2 + 2 \sum_i^m \sum_j^{n_i} \sum_{k>j}^{n_i} \frac{\pi_{ii}-\pi_i^2}{\pi_i^2 \pi_{ii}} Z_{ij} Z_{ik}$
$= \sum_i^m \{ \frac{1-\pi_i}{\pi_i^2} Z_i + 2 \frac{\pi_{ii}-\pi_i^2}{\pi_i^2 \pi_{ii}} \sum_j^{n_i} \sum_{k>j}^{n_i} Z_{ij} Z_{ik} \} = \sum_i^m \{ \frac{1-\pi_i}{\pi_i^2} Z_i + \frac{\pi_{ii}-\pi_i^2}{\pi_i^2 \pi_{ii}} Z_i (Z_i - 1) \}$
$= \sum_i^m \frac{N_i(N_i-n_i)}{n_i^2(n_i-1)} Z_i (n_i - Z_i)$

Likelihood:

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

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

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

One iteration of MCMC Algorithm:

Steps 1-5 are the same as Alg 3.1.

6. Draw Y_+ from its full conditional, which is simply $\mathcal{N}(\hat{Y}_+, \hat{\sigma}^2_{Y_+})$
7. Draw proposal for each Y_i
   - Randomly choose m numbers which add up to the difference between the previous and current Y_+ and add those amounts to the respective Y_is so that they add up to the newly drawn Y_+
   - 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
8. 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 Admin 2 level models with 1-stage SRS sampling

4.1 Estimating $\mathbf{P}(\tilde{Y}, \alpha, \tilde{\eta}, \phi, \tau | \tilde{Z}, \tilde{Y}_+)$ (assuming we know total deaths in each admin 1 area) BYM2

Setting:

$log(r_j) = \eta_j = \alpha + b_j$
Priors: $\alpha \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$, $b \sim BYM2(\tau, \phi)$
Hyperpriors: $\tau \sim Gamma(0.01, 0.01)$, $\phi \sim Beta(2, 3)$
$Y_{+i}|\tilde{r}_i \sim Pois(\sum_{j\in\delta(i)} N_j r_j)$ where $Y_{+i} = \sum_{j\in\delta(i)} Y_j$
$\tilde{Y}_i|Y_{+i}, \tilde{r}_i \sim Multinom(Y_{+i}, \frac{N_j r_j}{\sum_{j\in\delta(i)} N_j r_j})$
$Z_j|Y_j \sim Hypergeom(N_j, Y_j, n_j)$
where i indexes admin1, j indexes admin2, and $\tilde{Y}_i$ is the vector of counts for each admin2 region in admin1 region i.

Likelihood:

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

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

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

$$= \Pi_j\{P(Z_j|Y_j)\} \times \Pi_i\{P(\tilde{Y}_i|Y_{+i}, \tilde{\eta}_i) \times P(Y_{+i}|\tilde{\eta}_i)\} \times P_b(\tilde{\eta} - \alpha|\phi, \tau) \times \pi(\alpha)\pi(\phi)\pi(\tau)$$

One iteration of MCMC Algorithm:

Steps 1-5 are the same as Alg 3.1.

6. For each admin1 area, indexed by i,
   - Draw J from $\{1, ..., \lfloor m_i/2 \rfloor\}$ where m_i is the number of admin2 areas in admin1 area i
   - Randomly choose, without replacement, J admin2 areas to decrease by f_Y and J admin2 areas to increase by f_Y, where f_Y is some natural number
7. Accept proposal for $\tilde{Y}$ with probability

$$\frac{\Pi_j P(Z_j|Y_j^P)\Pi_i P(\tilde{Y}_i^P|Y_{+i}, \eta_i^C)}{\Pi_j P(Z_j|Y_j^C)\Pi_i P(\tilde{Y}_i^C|Y_{+i}, \eta_i^C)}$$

5 Models with two-stage stratified cluster sampling

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.

5.1 Estimating $P(\alpha_U, \alpha_R, \tilde{b}, \tilde{r}, \phi, \tau, \sigma^2 | \tilde{Z}, \tilde{R})$ BYM2, admin level model

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
- n_{ij} and Z_{ij} denotes sampled births and deaths in admin area i, cluster j, respectively
- R_{ij} is an indicator variable denoting whether cluster j in admin area i is sampled

Setting:
$log(r_{ij}) = \eta_{ij} = \alpha_U I(j \in urban) + \alpha_R I(j \in rural) + b_i + e_{ij}$
Priors: $\alpha_U \sim \mathcal{N}(\mu_U, \sigma_U^2), \alpha_R \sim \mathcal{N}(\mu_R, \sigma_R^2), b \sim BYM2(\tau, \phi), e \sim^{IID} \mathcal{N}(0, \sigma^2)$
Hyperpriors: $\tau \sim Gamma(0.01, 0.01)$, $\phi \sim Beta(2, 3)$, $\sigma^2 \sim InvGamma(0.01, 0.01)$
$Y_{ij} | r_{ij} \sim Pois(N_{ij} r_{ij})$
$Z_{ij} | Y_{ij}, R_{ij} = 1 \sim Hypergeom(N_{ij}, Y_{ij}, n_{ij})$
$P(Z_{ij} | Y_{ij}, R_{ij} = 0) = I(Z_{ij} = 0)$

Likelihood:

$$\mathcal{L}(\alpha_U, \alpha_R, \tilde{b}, \tilde{\eta}, \phi, \tau, \sigma^2 | \tilde{Z}, \tilde{R}) \propto P(\tilde{Z}, \tilde{R} | \tilde{\eta}) P(\tilde{\eta} | \alpha_U, \alpha_R, \tilde{b}, \sigma^2) P(\tilde{b} | \phi, \tau) \pi(\alpha_U, \alpha_R, \tau, \phi, \sigma^2)$$

$$\propto \Pi_i^m \Pi_{j \in C_i} P(Z_{ij} | R_{ij}, \eta_{ij}) \times P(\tilde{\eta} | \alpha_U, \alpha_R, \tilde{b}, \sigma^2) P(\tilde{b} | \tau, \phi) \pi(\alpha_U, \alpha_R, \tau, \phi, \sigma^2)$$

$$= \Pi_i^m \Pi_{j \in C_i : R_{ij}=1} P(Z_{ij} | R_{ij} = 1, \eta_{ij}) \times P_e(\tilde{\eta} - \alpha_{UR} - \tilde{b} | \sigma^2) P(\tilde{b} | \tau, \phi) \pi(\alpha_U, \alpha_R, \tau, \phi, \sigma^2)$$

$$= \Pi_i^m \Pi_{j \in C_i : R_{ij}=1} \sum_{y=Zij}^{N_{ij}} P(Z_{ij}, y | R_{ij} = 1, \eta_{ij}) \times P_e(\tilde{\eta} - \alpha_{UR} - \tilde{b} | \sigma^2) P(\tilde{b} | \tau, \phi) \pi(\alpha_U, \alpha_R, \tau, \phi, \sigma^2)$$

$$= \Pi_i^m \Pi_{j \in C_i : R_{ij}=1} \sum_{y=Zij}^{N_{ij}} P(Z_{ij}, y | R_{ij} = 1, \eta_{ij}) \times P_e(\tilde{\eta} - \alpha_{UR} - \tilde{b} | \sigma^2) P(\tilde{b} | \tau, \phi) \pi(\alpha_U, \alpha_R, \tau, \phi, \sigma^2)$$

$$= \Pi_i^m \Pi_{j \in C_i : R_{ij}=1} \sum_{y=Zij}^{N_{ij}} P(Z_{ij} | R_{ij} = 1, y) P(y | \eta_{ij}) \times P_e(\tilde{\eta} - \alpha_{UR} - \tilde{b} | \sigma^2) P(\tilde{b} | \tau, \phi) \pi(\alpha_U, \alpha_R, \tau, \phi, \sigma^2)$$

6 Computations for PC priors in stan

6.1 τ, spatial precision

Gumbell II distribution where $a = 1/2$ and $b = -ln(\alpha)/U$

6.2 ϕ, spatial variance

From Simpson, we have

$$\pi(\phi) = \lambda e^{-\lambda d(\phi)} |\frac{\partial d(\phi)}{\partial \phi}|$$

, where

$$d(\phi) = \sqrt{2KLD(\phi)} = \sqrt{\phi tr(Q^{-1}) - \phi m - ln|(1-\phi)I + \phi Q^{-1}|}$$

So it follows that

$$\frac{\partial d(\phi)}{\partial \phi} = \frac{1}{2}(2KLD(\phi))^{-1/2} \times \frac{\partial}{\partial \phi}[2KLD(\phi)]$$

$$= \frac{1}{2}(2KLD(\phi))^{-1/2} \times \{tr(Q^{-1}) - m - \frac{\partial}{\partial \phi} ln|(1-\phi)I + \phi Q^{-1}|\}$$

$$= (8KLD(\phi))^{-1/2} \times \{tr(Q^{-1}) - m - tr\{((1-\phi)I + \phi Q^{-1})^{-1}(Q^{-1} - I)\}\}$$

$$= (8KLD(\phi))^{-1/2} \times \{tr(Q^{-1} - I) - tr\{((1-\phi)I + \phi Q^{-1})^{-1}(Q^{-1} - I)\}\}$$

$$= (8KLD(\phi))^{-1/2} \times tr\{(I - ((1-\phi)I + \phi Q^{-1})^{-1})(Q^{-1} - I)\}$$

This implies the log pdf of the prior on ϕ is given by

$$log\pi(\phi) = log\lambda - \lambda\sqrt{2KLD(\phi)} - \frac{1}{2}log(8KLD(\phi)) + log|tr(W_{\phi,Q})|$$

where $W_{\phi,Q} = (I - ((1-\phi)I + \phi Q^{-1})^{-1})(Q^{-1} - I)$
or, equivalently,

$$\pi(\phi) = \frac{\lambda}{2\sqrt{2KLD(\phi)}} e^{-\lambda\sqrt{2KLD(\phi)}} |tr(W_{\phi,Q})|$$

Now we need to choose λ to calibrate the distribution s.t. $P(\phi < U) = \alpha$ (find λ which minimizes difference between closed form and output from `inla.pc.bym.phi`)