Stage 2 simulation results

Austin Schumacher

2021-12-06

Overview

We will perform a simulation study to explore the recovery of parameters

Data generating mechanism

The likelihood for the DGM will be

$$(\hat{\hat{y}}_r^{HAZ}, \hat{\hat{y}}_r^{WAZ}) | (\mu_r^{HAZ}, \mu_r^{WAZ}), \hat{\boldsymbol{V}}_r^{DES} \sim N_2((\mu_r^{HAZ}, \mu_r^{WAZ}), \hat{V}_r^{DES})$$

with $(\hat{Y}_r^{HAZ}, \hat{Y}_r^{WAZ})$ playing the role of "observed" data, $\hat{V}_r^{DES}$ assumed fixed and known.

For WAZ, we will specify the mean as the sum of an intercept, a spatial ICAR random effect, and a nonspatial IID random effect. For HAZ, we will specify the mean as the sum of an intercept, a spatial ICAR random effect, a nonspatial IID random effect, and a coefficient times the spatial ICAR random effect for WAZ. So we have

$$\begin{aligned}
\mu_r^{HAZ} &= \alpha_1 + v_{1r} + u_{1r} + \lambda u_{2r} \\
\mu_r^{WAZ} &= \alpha_2 + v_{2r} + u_{2r} \\
v_{1r} | \sigma_{v1}^2 &\overset{iid}{\sim} N(0, \sigma_1^2) \\
v_{2r} | \sigma_{v2}^2 &\overset{iid}{\sim} N(0, \sigma_2^2) \\
\boldsymbol{u}_1 &\sim ICAR(1) \\
\boldsymbol{u}_2 &\sim ICAR(1)
\end{aligned}$$

We will use a BYM2 parameterization for the spatial and IID random effects in each model such that, for $c = 1, 2$, we let

$$v_{cr} + u_{cr} = \sigma_c(\sqrt{1 - \rho_c} v_{cr}^* + \sqrt{\rho_c} u_{cr}^*)$$

where $v_{cr}^* \overset{iid}{\sim} N(0, 1)$ is the unstructured random effect with fixed standard deviation 1 and $\boldsymbol{u}_r^*$ is the ICAR model scaled so $\mathrm{Var}(u_{cr}^*) \approx 1$, which is done by scaling the model so the geometric mean of these variances is 1, using the adjacency matrix to calculate the inverse precision of the ICAR model as in Riebler (2016).

We will present simulation results for different scenarios. Each scenario will have data generated via specified parameter values (which may or may not have a shared component), and a Bayesian model fit to the data that may or may not be correctly specified for the DGM.

All data will be the same size as the Kenya 2014 DHS (47 regions).

The DGM parameters will be set equal to those estimated from models fit to the Kenya 2014 DHS HAZ/WAZ data. The nonshared models will have $\lambda = 0$.

If a DGM has "random RE", this means that the random effects v and u were randomly generated for each simulation. If a DGM has "nonrandom RE", this means that the random effects v and u were the same across all simulations.

For each scenario, we run 500 simulations.

Results

DGM: random RE, shared component, fixed V from 2014KDHS; Model: non-shared model, beta(1,1) prior on rho

Table 1: Simulation results

param	mean bias	rel bias	cov 80	width 80	cov 95	width 95
beta[1]	0.000	0.000	0.888	0.055	0.984	0.089
beta[2]	0.001	-0.001	0.940	0.050	0.992	0.081
sigma[1]	0.053	0.280	0.308	0.080	0.482	0.124
sigma[2]	-0.029	-0.112	0.634	0.074	0.854	0.116
rho[1]	-0.112	-0.128	0.916	0.460	0.992	0.642
rho[2]	-0.194	-0.198	0.332	0.439	0.872	0.620
latent means	0.000	0.189	0.769	0.196	0.931	0.300

Table 2: Stan diagnostics

mean_pct_divergent	mean_pct_max_tree_exceeded	mean_pct_bmfi_low_chains
0	0	0

DGM: random RE, shared component, fixed V from 2014KDHS; Model: shared model, beta(1,1) prior on rho

Table 3: Simulation results

param	mean bias	rel bias	cov 80	width 80	cov 95	width 95
beta[1]	0.000	0.000	0.844	0.048	0.964	0.076
beta[2]	0.000	-0.001	0.936	0.048	0.992	0.076
sigma[1]	-0.026	-0.139	0.692	0.090	0.892	0.143
sigma[2]	-0.007	-0.028	0.788	0.082	0.952	0.127
rho[1]	-0.179	-0.205	0.894	0.567	0.992	0.764
rho[2]	-0.135	-0.138	0.490	0.300	0.864	0.443
lambda	0.022	0.067	0.776	0.196	0.922	0.302
latent means	0.000	0.153	0.799	0.205	0.950	0.313

Table 4: Stan diagnostics

mean_pct_divergent	mean_pct_max_tree_exceeded	mean_pct_bmfi_low_chains
0	0	0