Supplementary Material for Estimation of Causal Peer Influence Effects

Panos Toulis, Edward Kao
Department of Statistics, Harvard University, 1 Oxford Street, Cambridge, MA 02138, USA

Appendix A: MCMC Inference Steps

This section describes the two-step inference in Figure 3, under the linear effects model.

Bayesian regression step:

In order to sample σ^2 more efficiently, away from zero, a weakly informative Normal-Inverse-gamma conjugate prior is used. This may introduce a slight bias but we will show experimentally that the estimates are reasonable. A large variance on the Normal is used to approximate a flat prior:

$$P_0(\boldsymbol{\beta}, \sigma^2) = \mathbb{N}(\boldsymbol{\beta}; \boldsymbol{\mu_0} = \mathbf{0}, \boldsymbol{V_0} = 100\boldsymbol{I}) \times \text{Inv-Gamma}(\sigma^2; a_0 = 2, b_0 = 2) \quad (1)$$

The resulting posterior can be sampled by first drawing σ^2 then $\boldsymbol{\beta}$

$$\sigma^2 | \boldsymbol{S}, \mathbf{z}, \boldsymbol{y} \sim \text{Inv-Gamma}(a^*, b^*) \quad (2)$$

$$\boldsymbol{\beta} | \sigma^2, \boldsymbol{S}, \mathbf{z}, \boldsymbol{y} \sim \mathbb{N}(\boldsymbol{\mu^*}, \boldsymbol{V^*}) \quad (3)$$

where $\boldsymbol{X}_{(N\times 3)} = \begin{bmatrix}\mathbf{z} & \boldsymbol{S} & \mathbf{1}\end{bmatrix}$, $\boldsymbol{V^*} = (\boldsymbol{V_0^{-1}} + \boldsymbol{X^T X})^{-1}$, $\boldsymbol{\mu^*} = (\boldsymbol{V_0^{-1}} + \boldsymbol{X^T X})^{-1}(\boldsymbol{V_0^{-1}\mu_0} + \boldsymbol{X^T y})$, $a^* = a_0 + n/2$, $b^* = b_0 + \frac{1}{2}(\boldsymbol{\mu^T V_0^{-1} \mu_0} + \boldsymbol{y^T y} - \boldsymbol{\mu^{*T} V^{*-1} \mu^*})$

Metropolis-Hastings step:

Conditioning on everything else, a random scan is performed through each S_i by drawing from the following unnormalized closed-form posterior:

$$P(S_i | \mathbf{z}, \boldsymbol{y}, \boldsymbol{\beta}, \sigma^2, \kappa_i) \propto P(y_i | S_i, \mathbf{z}, \boldsymbol{\beta}, \sigma^2) \times P_0(S_i | \kappa_i) = \mathbb{N}(y_i; \boldsymbol{X_{i\cdot}\beta}, \sigma^2) \times \text{Pois}(S_i; \kappa_i) \quad (4)$$

A random walk proposal is used with a boundary at 0 (since the amount of peer influence effects is non-negative). The acceptance probability is the following (note the Hastings correction is only effective at the boundary):

$$P_{acc} = min\left(1, \frac{P(S_i^{t+1}|\cdot)}{P(S_i^t|\cdot)} \frac{P(S_i^{t+1} \to S_i^t)}{P(S_i^t \to S_i^{t+1})}\right) \quad (5)$$

1. Appendix B: Proof of Theorem 1

This section presents the proof for Theorem 1 in section 3.1.

Proof.
Define as W_i the membership of i in A of INR, i.e. $W_i = 1$ iff $i \in A$. For convenience denote, $S_0 = \sum_i Y_i(\mathbf{0})$ and $N_c = N - N_t$. It holds that: $E[\hat{\delta}_{INR}] = E[\sum_{i\in A}(\frac{1}{N_t}Y_i(\mathbf{z}_i) - \frac{1}{N_c}Y_i(\mathbf{0}))] = \frac{1}{N_t} \cdot E[\sum_i W_i \cdot Y_i(\mathbf{z}_i)] - \frac{1}{N_c} \cdot E[\sum_i (1 - W_i) \cdot Y_i(\mathbf{0})]$

By unconfoundedness $W_i \perp Y_i$, and also $E[W_i] = N_t/N$ and $E[1 - W_i] = N_c/N$ so we have:

$$E[\sum_i W_i \cdot Y_i(\mathbf{z}_i)] = \sum_i E[W_i] \cdot E[Y_i(0, \mathbf{z}_i)] = \frac{N_t}{N} \sum_i \rho_i \sum_{\mathbf{z}\in\mathbf{Z}_0(\mathcal{N}_i;k)} Y_i(0, \mathbf{z}) = N_t(\delta_0 + S_0/N)$$

Furthermore,

$$\frac{1}{N_c} \cdot E[\sum_i (1 - W_i) Y_i(\mathbf{0})] = \frac{1}{N_c} \cdot E[1 - W_i] \cdot \sum_i Y_i(\mathbf{0}) = S_0/N$$

Substituting back into the original equation:

$$E[\hat{\delta}_{INR}] = \delta_0 + S_0/N - S_0/N = \delta_0$$

By definition we have that,

$$\delta = \frac{1}{N}\Big(\sum_i \rho_i \cdot \sum_{\mathbf{z}\in\mathbf{Z}(\mathcal{N}_i;k)} Y_i(0, \mathbf{z})\Big) - \frac{1}{N} S_0$$

Define for convenience $\delta_{0,i} = \rho_{0,i} \cdot (\sum_{\mathbf{z}\in\mathbf{Z}_0(\mathcal{N}_i;k)} Y_i(0, \mathbf{z})) - Y_i(\mathbf{0})$ and $\delta_{1,i} = \rho_{1,i} \cdot (\sum_{\mathbf{z}\in\mathbf{Z}_1(\mathcal{N}_i;k)} Y_i(0, \mathbf{z})) - Y_i(\mathbf{0})$

Notice that, $\delta_0 = \frac{1}{N}\sum_i \delta_{0,i}$ and $\delta_1 = \frac{1}{N}\sum_i \delta_{1,i}$. It is straightforward then to see that,

$$\delta = \frac{1}{N}\sum_i (\frac{\rho_i}{\rho_{0,i}}\delta_{0,i} + \frac{\rho_i}{\rho_{1,i}}\delta_{1,i}) \quad (6)$$

Notice that, by our assumptions, $\rho_i/\rho_{0,i} = \alpha$ and $\rho_i/\rho_{1,i} = 1-\alpha$, thus we simplify as

$$\delta = \alpha\delta_0 + (1-\alpha)\delta_1 = \delta_0 - (1-\alpha)(\delta_0 - \delta_1) \quad (7)$$

Using the fact we proved earlier that $E[\hat{\delta}_{INR}] = \delta_0$, we finally get:

$$E[\hat{\delta}_{INR}] = \delta + (1-\alpha)\cdot(\delta_0 - \delta_1)$$

This completes the proof. ■