Ralph's partially-identified Gibbs sampler

This is a continuation of Ralph's partially-identified DAG. After struggling with partial identification of causal effects, Ralph is now interested in addressing inference from a sample plagued by imperfect compliance.

Ralph knows the 16 compliance-response (cr) pairs define the causal effects of interest.

$$\begin{aligned}
\Pr\left(y_1 \mid do\left(x_0\right)\right) &= \sum_{i=0}^{3} v_{c=i,r=2} + \sum_{i=0}^{3} v_{c=i,r=3}, \\
\Pr\left(y_1 \mid do\left(x_1\right)\right) &= \sum_{i=0}^{3} v_{c=i,r=1} + \sum_{i=0}^{3} v_{c=i,r=3}, \\
ACE\left(X \to Y\right) &= \sum_{i=0}^{3} v_{c=i,r=1} - \sum_{i=0}^{3} v_{c=i,r=2}
\end{aligned}$$

where v_{cr} is the probability associated with the cr pair.

Since only $p_{yx.z}$ is observed in the sample and v_{cr} is latent, Ralph decides to employ Markov-chain Monte Carlo simulation in the form of a Gibbs sampler (that is, from the full set of conditional posterior distributions). The first conditional posterior is

$$\Pr\left(cr^i \mid v_{cr^i}, data = \{p_{yx.z}\}\right) \propto f\left(x^i, y^i \mid z^i, cr^i\right) v_{cr^i}$$

where the superscript refers to individual i in the sample and $f\left(x^i, y^i \mid z^i, cr^i\right)$ is an indicator function equal to one when x, y, z agrees with cr and zero otherwise. Ralph recognizes this as a multinomial distribution and generates values for the latent variable cr.

cr	$yx.z$
00	00.0, 00.1
01	00.0, 00.1
02	10.0, 10.1
03	10.0, 10.1
10	00.0, 01.1
11	00.0, 11.1
12	10.0, 01.1
13	10.0, 11.1
20	01.0, 00.1
21	11.0, 00.1
22	01.0, 10.1
23	11.0, 10.1
30	01.0, 01.1
31	11.0, 11.1
32	01.0, 01.1
33	11.0, 11.1

As v_{cr^i} is unknown, Ralph begins with some initial value and replaces it in subsequent rounds with draws from the second conditional posterior distribution

$$\Pr\left(v_{CR} \mid cr^1, \ldots, cr^n\right) \propto \prod_{i=0}^{3} \prod_{j=0}^{3} \left(v_{cr_{ij}}\right)^{N_{ij}+N_{ij}^{'}-1}$$

where N_{ij} refers to the number of draws corresponding to cr_{ij} from the first conditional posterior and $N_{ij}^{'}$ refers to the prior concentration parameter for a Dirichlet distribution. This conditional posterior follows a Dirichlet distribution and generates v_{cr} draws.

Suppose the DGP is

$$\begin{array}{cc} p_{00.0} = 0.55 & p_{00.1} = 0.45 \\ p_{01.0} = 0.45 & p_{01.1} = 0. \\ p_{10.0} = 0. & p_{10.1} = 0. \\ p_{11.0} = 0. & p_{11.1} = 0.55 \\ \Pr\left(z_1\right) = 0.50 & \Pr\left(z_0\right) = 0.50 \end{array}$$

Suggested:

1. For the DGP, generate a sample of $n = 100$. Using a Gibbs sampler, simulate 15,000 draws and discard the first 5,000 burn-in draws. Examine plots, histograms, and descriptive statistics (mean, standard deviation, and quantiles) for the post burn-in draws.

2. Repeat 1 for a sample of $n = 1,000$. Compare inferences for the two sample sizes relative to the known DGP.

Now, suppose the DGP is

$$\begin{array}{cc} p_{00.0} = 0.919 & p_{00.1} = 0.315 \\ p_{01.0} = 0.000 & p_{01.1} = 0.139 \\ p_{10.0} = 0.081 & p_{10.1} = 0.073 \\ p_{11.0} = 0.000 & p_{11.1} = 0.473 \\ \Pr\left(z_1\right) = 0.500 & \Pr\left(z_0\right) = 0.500 \end{array}$$

3. Repeat 1 and 2 for the second DGP.