

ipwstep8

lqy

February 13, 2018

Overview

In this step:

Setting for data generation:

- 1. cluster sizes: 50 clusters in each intervention arm; 50 individuals (varies)
- 2. Covariate X generation: do not assign a large variance.
- 3. Outcome model S1 i.e. no interaction and a single ICC for this i.e. 0.05
- Model: $\pi_{ijl} = \text{expit}(1 + 1.36i + x_{ijl} + \delta_{ij})$
 - $\delta \sim N(0, 0.2)$, $x_{ijl} = \alpha_{ij} + u_{ijl}$
 - $\alpha_{ij} \sim N(0, 0.18)$
 - $u_{ijl} \sim N(0, 0.2)$
- 4. Missing model
 - a. Miss 2 i.e. no clustering: $\text{logit}(R_{ijl} = 0|Y_{ij}, X_{ij}) = \text{intercept} + X_{ijl} + i$ (intercept varies to control missing percentage)
 - b. Miss 5 with different ICC of missingness: $\text{logit}(R_{ijl} = 0|Y_{ij}, X_{ij}) = -1.8 + X_{ijl} + i + \theta_{ij}$, θ_{ij} changes with the changes of missingness ICC, from 0.05, 0.1, 0.2 to 0.9.
- 5. True effect calculation: Calculate the truth using Hosain's method i.e. analyzing the complete data set

Analysis of each data set:

- 1. Use geeDREstimation
- 2. Run W-GEE using both independence and exchangeable in each scenario
- 4. IPW versions
 - a. Weights calculated with no clustering and correct model i.e. A+X
 - b. Weights calculated with clustering and correct model i.e. A+X +theta

Results

Table 1: Results for independent working correlation matrix

ICC	EST		MCSD		SD		Coverage		non_con
	IPW	IPW_CLU	IPW	IPW_CLU	IPW	IPW_CLU	IPW	IPW_CLU	
0.1	0.021	0.024	0.144	0.148	0.154	0.160	0.945	0.950	0.000
0.2	0.015	0.019	0.143	0.154	0.155	0.171	0.962	0.978	0.000
0.3	0.007	0.010	0.144	0.162	0.154	0.184	0.968	0.975	0.000
0.4	0.004	0.011	0.133	0.167	0.148	0.198	0.975	0.970	0.000
0.5	0.006	0.017	0.141	0.192	0.158	0.231	0.972	0.968	0.000
0.6	0.001	0.018	0.131	0.212	0.148	0.248	0.982	0.972	0.000
0.7	0.000	0.027	0.152	0.249	0.160	0.284	0.962	0.972	0.005
0.8	0.003	0.020	0.137	0.263	0.154	0.289	0.975	0.957	0.002
0.9	0.003	0.035	0.128	0.243	0.146	0.260	0.975	0.962	0.000

Table 2: Results for exchangeable working correlation matrix

ICC	EST		MCSD		SD		Coverage		non_con
	IPW	IPW_CLU	IPW	IPW_CLU	IPW	IPW_CLU	IPW	IPW_CLU	
0.1	0.021	0.024	0.145	0.149	0.154	0.160	0.952	0.952	0.000
0.2	0.014	0.019	0.144	0.155	0.155	0.170	0.958	0.975	0.000
0.3	0.006	0.010	0.145	0.162	0.153	0.183	0.962	0.975	0.000
0.4	0.004	0.013	0.134	0.169	0.148	0.198	0.975	0.968	0.000
0.5	0.006	0.018	0.143	0.191	0.157	0.231	0.962	0.968	0.000
0.6	0.001	0.020	0.133	0.214	0.147	0.248	0.978	0.972	0.000
0.7	0.000	0.030	0.155	0.253	0.159	0.284	0.960	0.967	0.005
0.8	0.002	0.018	0.140	0.272	0.152	0.291	0.967	0.957	0.002
0.9	0.004	0.037	0.130	0.255	0.145	0.261	0.960	0.948	0.000

- IPW: no cluster effects
- IPW_CLU: contains cluster effects

In those scenarios, with different ICCs, the results do not have big differences with each other. and with large ICC, they slightly overcoverage. And with out consideration of clusters, the results seems to be better.

Change variance

Since the variances in the data generation may affects the results a lot. I also tried to run the simulation with the different variances.

variances: 0.2, 0.5, 0.8, 1.1, 2, 3, 5

ICC: fixed as 0.5

Repeat: 200 times

Table 3: Results for exchangeable working correlation matrix

Var	EST		MCSD		SD		Coverage		non_con
	IPW	IPW_CLU	IPW	IPW_CLU	IPW	IPW_CLU	IPW	IPW_CLU	
0.2	0.008	0.000	0.145	0.191	0.157	0.233	0.975	0.970	0.000
0.5	0.003	0.010	0.155	0.223	0.162	0.255	0.980	0.940	0.000
0.8	0.002	0.029	0.158	0.244	0.168	0.277	0.955	0.935	0.000
1.1	0.008	0.021	0.164	0.258	0.172	0.296	0.965	0.955	0.005
2.0	0.005	0.073	0.175	0.315	0.182	0.338	0.965	0.945	0.000
3.0	0.003	0.100	0.187	0.369	0.195	0.374	0.944	0.934	0.015
5.0	0.028	0.157	0.213	0.452	0.225	0.426	0.934	0.867	0.020

Table 4: Results for exchangeable working correlation matrix

Var	EST		MCSD		SD		Coverage		non_con
	IPW	IPW_CLU	IPW	IPW_CLU	IPW	IPW_CLU	IPW	IPW_CLU	
0.2	0.009	-0.001	0.148	0.190	0.156	0.233	0.955	0.975	0.000
0.5	0.005	0.007	0.156	0.218	0.161	0.256	0.970	0.945	0.000
0.8	0.000	0.029	0.159	0.240	0.168	0.276	0.960	0.940	0.000
1.1	0.006	0.026	0.170	0.252	0.172	0.294	0.950	0.950	0.005
2.0	0.006	0.065	0.177	0.310	0.184	0.336	0.965	0.950	0.000
3.0	0.004	0.085	0.191	0.385	0.198	0.379	0.934	0.934	0.015
5.0	0.033	0.165	0.216	0.429	0.228	0.424	0.949	0.893	0.020

By using CRTgeeDR, even with large variance, there do not have a lot of nonconvergence.