Diabetes-dementia updated simulation results

Andrew Mertens

2022-11-17

To change in organization cgnage method to "Estimator" and put as first column change estimator to "algorithm" and put as second column

NA for IPTW Qint and detq, but pick just detq (or not for qint)

Figure out why GLM Q-int and no Q-int is the same— rerun

Add in different trunctation

-Fix lasso pre-screening labeling

Scenario 1: Realistic simulation, null outcome

True RD: 0

filenames	Estimator	Method	Q-int	Det. Q	Bias	Varia
EN_ic_T11_iptw	Elastic Net	IPTW	NA	Yes	0.00121	0.000
EN_ic_T11	Elastic Net	TMLE	No	Yes	-0.00012	0.000
noDetQ_ic_glm_T11_iptw	GLM	IPTW	NA	No	0.00157	0.000
ic_glm_T11_iptw	GLM	IPTW	NA	Yes	0.00140	0.000
noDetQ_ic_glm_T11	GLM	TMLE	No	No	0.00217	0.000
ic_glm_T11	GLM	TMLE	No	Yes	0.00369	0.001
ic_Qint_noDetQ_lasso_prescreen_T11_iptw	GLM, LASSO prescreen	IPTW	NA	No	0.00144	0.000
ic_Qint_noDetQ_lasso_prescreen_T11	GLM, LASSO prescreen	TMLE	Yes	No	0.00031	0.000
noDetQ_tmle_T11_iptw	LASSO	IPTW	NA	No	-0.00006	0.000
noDetQ_Qint_ic_T11_iptw	LASSO	IPTW	NA	No	0.00137	0.000
ic_T11_iptw	LASSO	IPTW	NA	Yes	0.00120	0.000
noDetQ_tmle_T11	LASSO	TMLE	No	No	-0.00006	0.000
noDetQ_ic_T11	**LASSO**	**TMLE**	**No**	**No**	**-0.00005**	**0.000**
ic_T11	LASSO	TMLE	No	Yes	-0.00011	0.000
noDetQ_Qint_ic_T11	LASSO	TMLE	Yes	No	-0.00005	0.000
noDetQ_Qint_tmle_T11	LASSO	TMLE	Yes	No	-0.00003	0.000
rf_ic_T11_iptw	Random Forest	IPTW	NA	Yes	0.00206	0.000
rf_ic_T11	Random Forest	TMLE	No	Yes	-0.00364	0.000
ridge_ic_T11_iptw	Ridge	IPTW	NA	Yes	0.00094	0.000
ridge_ic_T11	Ridge	TMLE	No	Yes	-0.00014	0.000

Performance of difference variance estimators on null data

Notes:

- Only showing LASSO estimator results-all estimator performances assessed in the realistic simulated data below.
- Sanity-check on estimation performance on data with a known null association between GLP1 and dementia.
- The IC variance estimator is anti-conservative and the TMLE variance estimator is conservative.
- The bootstrap is anti-conservative but less so than the IC variance estimator.
- The TMLE estimator is very conservative, with CI widths 8-10X that of the bootstrap.
- The IPTW estimator is uniformly biased with overly-wide confidence intervals in all simulations (not shown).

variance_estimator	coverage	mean_ci_width	filenames
ic	31.5	0.00294	old_null_sim_res_noDetQ_ic_T11
tmle	100.0	0.11535	old_null_sim_res_noDetQ_tmle_T11
bootstrap	91.5	0.01288	bootstrap
bootstrap-clustered ID	100.0	0.02023	bootstrap_iptw-IPTW

Scenario 2: Realistic simulation, protective effect of GLP1 on dementia

True Risk Difference: -0.009683665

Comparison of different estimators' performance

Notes:

- Based on these results, we chose the LASSO estimator with Q-prediction and no deterministic Q function
- Several of the estimators have comparable performance, but the chosen estimator performs best in both RR and RD estimation
- Ridge regressions have lower MSE but not perfect 95% oracle coverage
- Including the deterministic Q function marginally decreases bias/variance, so we should use in the bootstrap estimator

Risk difference

Estimator	Method	Q-int	Det. Q	Bias	Variance	mse	Bias/SE	Oracle coverage
Elastic Net	IPTW	No	Yes	0.003509	9.0e-06	2.1e-05	1.196808	85.0
Elastic Net	TMLE	No	Yes	-0.002305	1.4e-05	1.9e-05	-0.620157	92.0
Elastic Net	IPTW	Yes	Yes	0.003509	9.0e-06	2.1e-05	1.196808	85.0
Elastic Net	TMLE	Yes	Yes	-0.002305	1.4e-05	1.9e-05	-0.620157	92.0
GLM	IPTW	No	Yes	0.003525	9.0e-06	2.1e-05	1.171211	87.5
GLM	TMLE	No	Yes	0.002048	6.4e-05	6.8e-05	0.256147	93.0
LASSO	IPTW	No	No	0.005853	6.0e-06	4.1e-05	2.309559	37.5
LASSO	TMLE	No	No	-0.000766	2.5e-05	2.5e-05	-0.153535	94.0
LASSO	IPTW	No	Yes	0.003485	9.0e-06	2.1e-05	1.185158	85.0
LASSO	**TMLE**	**No**	**Yes**	**0.000267**	**1.4e-05**	**1.4e-05**	**0.070843**	**94.5**
LASSO	IPTW	Yes	No	0.005853	6.0e-06	4.1e-05	2.309559	37.5
LASSO	TMLE	Yes	No	-0.000766	2.5e-05	2.5e-05	-0.153535	94.0

LASSO	IPTW	Yes	Yes	0.003485	9.0e-06	2.1e-05	1.185158	85.0
LASSO, AUC fit	IPTW	No	Yes	0.003585	6.0e-06	1.9e-05	1.431366	71.5
LASSO, AUC fit	TMLE	No	Yes	-0.002860	8.0e-06	1.6e-05	-1.009325	81.0
LASSO, AUC fit	IPTW	Yes	Yes	0.003585	6.0e-06	1.9e-05	1.431366	71.5
LASSO, AUC fit	TMLE	Yes	Yes	-0.002860	8.0e-06	1.6e-05	-1.009325	81.0
LASSO, Lambda: 1se	IPTW	No	Yes	0.003240	6.0e-06	1.6e-05	1.357844	77.5
LASSO, Lambda: 1se	TMLE	No	Yes	-0.002463	9.0e-06	1.5e-05	-0.802549	89.5
LASSO, Lambda: 1se	IPTW	Yes	Yes	0.003240	6.0e-06	1.6e-05	1.357844	77.5
LASSO, Lambda: 1se	TMLE	Yes	Yes	-0.002463	9.0e-06	1.5e-05	-0.802549	89.5
Ridge	IPTW	No	Yes	0.003830	6.0e-06	2.1e-05	1.521015	71.5
Ridge	TMLE	No	Yes	0.000446	1.1e-05	1.1e-05	0.134349	94.0
Ridge, AUC fit	IPTW	No	Yes	0.003943	5.0e-06	2.0e-05	1.802503	53.5
Ridge, AUC fit	TMLE	No	Yes	-0.002833	6.0e-06	1.4e-05	-1.174903	76.5

Comparison of different variance estimators

Notes:

- Showing LASSO estimator results with modeled Q (rather than intercept-only)
- The IC variance estimator is anti-conservative and the TMLE variance estimator is conservative
- The bootstrap is anti-conservative but less so than the IC variance estimator
- The IPTW estimator is uniformly biased with overly-wide confidence intervals in all simulations (not shown)

Risk difference coverage

variance_estimator	coverage	mean_ci_width	power	bias_se_ratio_emp	filenames
ic, Det-Q	41.0	0.00522	95.0	-0.20338	sim_res_DetQ_
ic	36.5	0.00523	96.0	-0.57351	sim_res_noDetQ
ic, Det-Q	67.0	0.00736	92.0	0.14223	sim_res_DetQ_
tmle	90.0	0.01751	68.5	-0.16649	sim_res_noDetQ
tmle	99.5	0.02129	49.0	-0.05020	sim_res_noDetQ
NA	46.5	0.00144	100.0	-0.09835	sim_res_rf
Bootstrap, Det Q function	87.0	0.01346	68.5	NA	DetQ
Bootstrap, Det Q function, 500 iterations	89.0	0.01338	69.5	NA	DetQ, 500 iter
Bootstrap-Ridge	87.5	0.01289	72.0	NA	Ridge
Bootstrap	85.5	0.01454	68.5	NA	no DetQ
Bootstrap- IPTW	100.0	0.02341	0.0	NA	no DetQ_iptw-I