Estimating SNMs in the Danish population

Ella Jacobsen, 13-10-2022

UNIVERSITY OF COPENHAGEN

Motivation

- Estimating causal effects of DTRs in Danish population
- Applying SNMs in a setting with complex longitudinal data

Structural nested mean models

Causal effects for single binary treatment $A = a$ parameterized by

$$g\{E(Y^a|L = l, A = a)\} - g\{E(Y^0|L = l, A = a)\} = \gamma^\star(l, a; \psi^\star),$$

where

- g is a known link function
- Y^a is potential outcome under treatment a
- $L = l$ is a covariate vector
- $\gamma^\star$ is a known function, smooth in $\psi^\star$, satisfying $\gamma^\star(l, 0; \psi) = 0$ for all l, ψ
- $\psi^\star$ is the true parameter

Structural nested mean models

We can construct a variable $U^\star(\psi)$ whose mean equals the mean if treatment had been removed,

$$U^\star(\psi) = Y - \gamma^\star(L, A; \psi)$$

(E.g. if g is the identity link, then $U^\star(\psi) = Y \exp\{-\gamma^\star(L, A; \psi)\}$)

Structural nested mean models

SNMMs can be extended to repeated measures:

$$g\{E(\underline{Y}_{m+1}^{\overline{a}_m,0}|\overline{L}_m=\overline{l}_m,\overline{A}_m=\overline{a}_m)\}-g\{E(\underline{Y}_{m+1}^{\overline{a}_{m-1},0}|\overline{L}_m=\overline{l}_m,\overline{A}_m=\overline{a}_m)\}$$
$$=\gamma^\star(\overline{l}_m,\overline{a}_m;\psi^\star),$$

where underlines denote the future of a variable and overlines denote its history.

Assumptions

For identifiability we need no unmeasured confounders, formulated as

$$A \perp\!\!\!\perp Y^0 | L,$$

or for sequential treatment assignment

$$A_m \perp\!\!\!\perp \underline{Y}_{m+1}^{\overline{a}_{m-1},0} | \overline{L}_m, \overline{A}_{m-1} = \overline{a}_{m-1}$$

for $m = 0, \ldots, K$.

Structural nested mean models: estimation

The assumption of no unmeasured confounders means we can take

$$E\{U^\star(\psi^\star)|L, A\} = E\{U^\star(\psi^\star)|L\}$$

and thus estimate $\psi^\star$ by estimating equations

$$0 = \sum_{i=1}^{n} [d^\star(A_i, L_i) - E\{d^\star(A_i, L_i)|L_i\}] \cdot [U_i^\star(\psi) - E\{U_i^\star(\psi)|L_i\}].$$

Here $d^\star$ are arbitrary functions of same dimension as ψ.

Calculating empirical means

We set the empirical conditional covariance between $U^{\star}$ and $d^{\star}$ to zero - but how do we find empirical means?
Answer: Life2Vec - predictive model developed at DTU for life event embeddings

Social scientific questions

What is an appropriate treatment sequence?

- Ideally not well-predicted by covariates
- Sequence and outcome are recorded within about 10 years
- No unmeasured confounders - exogenous events?