Morgan Note for Bang Robins

In this section, let the temporally ordered observed data be $\mathbf{O} = (\mathbf{L}_1, A_1, \mathbf{L}_2, A_2, ..., \mathbf{L}_K, A_K, \mathbf{L}_{K+1})$ where A_k is a treatment given at time k and $\mathbf{L}_k$ are other variables measured just prior to treatment. Associated with each treatment history $\overline{\mathbf{a}} = (a_1, ..., a_K)$, there is a counterfactual random variable $\mathbf{L}_{\overline{\mathbf{a}}} = \overline{\mathbf{L}}_{\overline{\mathbf{a}},K+1}$ recording a subject's response history if treatment regime $\overline{\mathbf{a}}$ was followed. We link the counterfactual data to the observed data through the consistency assumption $\overline{\mathbf{L}}_{\overline{\mathbf{a}},m} = \overline{\mathbf{L}}_m$ if $\overline{\mathbf{A}}_{m-1} = \overline{\mathbf{a}}_{m-1}$ which states that the observed and counterfactual response through m will be equal if the observed and counterfactual treatments agree through $m-1$. That is to say, the future cannot determine the past. We impose the assumption of sequential ignorability (i.e., no unmeasured confounders) that for all $\overline{\mathbf{a}}$ and m

$$L_{\overline{\mathbf{a}}} \coprod A_m \mid \overline{\mathbf{L}}_m, \overline{\mathbf{A}}_{m-1} = \overline{\mathbf{a}}_{m-1} \tag{1}$$

which implies that sufficient covariates have been recorded in the $\mathbf{L}_m$ so that, as in a sequential randomized trial, the treatment A_m is independent of the counterfactuals given the observed past. Further we assume that, for all a_m in the support of A_m,

$$\text{if } f(\overline{\mathbf{A}}_{m-1}, \overline{\mathbf{L}}_m) > 0 \;\; \text{then } f(a_m \mid \overline{\mathbf{A}}_{m-1}, \overline{\mathbf{L}}_m) > 0, \tag{2}$$

which says that there is a positive probability that, in the observed study, any regime $\overline{\mathbf{a}}$ may be followed by a given subject.

We shall consider inference concerning the parameter $E\left[Y_{\overline{a}}\right] = E\left[L_{K+1,\overline{a}}\right]$

DR Estimation of $E\left[Y_{\overline{a}}\right]$

1. .Compute the MLE $\hat{\boldsymbol{\alpha}}$ of $\boldsymbol{\alpha}$ from the observed data by pooled (over persons i and times m) logistic regression

$$\log it\left\{pr\left(A_{m,i}=1 \mid \bar{\mathbf{l}}_{m,i}, \bar{\mathbf{a}}_{m-1,i}; \boldsymbol{\alpha}\right)\right\} = w_m\left(\bar{\mathbf{l}}_{m,i}, \bar{\mathbf{a}}_{m-1,i}; \boldsymbol{\alpha}\right)$$

   where w_m is a user specified function.

   Do three times to test double robustness and efficiency: (i) for the true propensity model generating the data, (ii) a correct super model, and (iii) an incorrect super model.

2. Set $\hat{T}_{K+1} = \mathbf{L}_{K+1} = Y.$

3. Recursively, for $m = K+1, \ldots, 2,$

4. a): specify and fit by IRLS a parametric regression model $h_{m-1}(\overline{\mathbf{L}}_{m-1}, \overline{\mathbf{A}}_{m-1}; \boldsymbol{\beta}_{m-1}, \phi_{m-1}) = \Psi\{s_{m-1}(\overline{\mathbf{L}}_{m-1}, \overline{\mathbf{A}}_{m-1}; \boldsymbol{\beta}_{m-1}) + \phi_{m-1}\overline{\pi}_{m-1}^{-1}(\hat{\boldsymbol{\alpha}})\}$ for the conditional expectation $E\{\hat{T}_m \mid \overline{\mathbf{A}}_{m-1}, \overline{\mathbf{L}}_{m-1}\}$, where $s_{m-1}(\overline{\mathbf{L}}_{m-1}, \overline{\mathbf{A}}_{m-1}; \boldsymbol{\beta}_{m-1})$ is a known function with the unknown parameter $\boldsymbol{\beta}_{m-1}$, Ψ is the canonical link function of a given GLM, and $\overline{\pi}_m(\hat{\boldsymbol{\alpha}}) = \prod_{j=1}^{m} f(A_j \mid \overline{\mathbf{L}}_j, \overline{\mathbf{A}}_{j-1}; \hat{\boldsymbol{\alpha}})$.

   b): Let $\hat{h}_{m-1}\left(\overline{\mathbf{L}}_{m-1}, \overline{\mathbf{A}}_{m-1}\right) = \Psi\{s_{m-1}(\overline{\mathbf{L}}_{m-1}, \overline{\mathbf{A}}_{m-1}; \hat{\boldsymbol{\beta}}_{m-1}) + \hat{\phi}_{m-1}\overline{\pi}_{m-1}^{-1}(\hat{\boldsymbol{\alpha}})\}$ be the predicted value from IRLS fit of the model. This implies that $(\hat{\boldsymbol{\beta}}'_{m-1}, \hat{\phi}_{m-1})'$ is a solution of

$$\mathbf{0} = \widetilde{E}\Big[[\widehat{T}_m - \Psi\{s_{m-1}(\overline{\mathbf{L}}_{m-1}, \overline{\mathbf{A}}_{m-1}; \hat{\boldsymbol{\beta}}_{m-1}) + \hat{\phi}_{m-1}\overline{\pi}_{m-1}^{-1}(\hat{\boldsymbol{\alpha}})\}]$$
$$\{\partial s(\overline{\mathbf{L}}_{m-1}; \hat{\boldsymbol{\beta}}_{m-1})/\partial \boldsymbol{\beta}'_{m-1}, \overline{\pi}_{m-1}^{-1}(\hat{\boldsymbol{\alpha}})\}\Big]$$

   where $\widetilde{E}(V) = n^{-1}\sum_{i=1}^{n} V_i$.

c): Set $\hat{T}_{m-1} = \hat{h}_{m-1}\left(\overline{\mathbf{L}}_{m-1}, \overline{\mathbf{A}}_{m-2}, a_{m-1}\right) =$

$\Psi\{s_{m-1}(\overline{\mathbf{L}}_{m-1}, \overline{\mathbf{A}}_{m-2}, a_{m-1}; \hat{\boldsymbol{\beta}}_{m-1})$

$+\hat{\phi}_{m-1} \prod_{j=1}^{m-2} f(A_j \mid \overline{\mathbf{L}}_j, \overline{\mathbf{A}}_{j-1}; \hat{\boldsymbol{\alpha}}) f(a_{m-1} \mid \overline{\mathbf{L}}_{m-1}, \overline{\mathbf{A}}_{m-2}; \hat{\boldsymbol{\alpha}})$

Note we could write $\hat{T}_{m-1}$ more precisely as $\hat{T}_{m-1}^{a_{m-1},\ldots,a_K}$

5. Finally $\widehat{E}\left[Y_{\overline{a}}\right] = \widetilde{E}\left[\hat{T}_1\right] = \widetilde{E}\left[\hat{T}_1^{\overline{a}}\right]$