

# Morgan Note for Bang Robins

In this section, let the temporally ordered observed data be  $\mathbf{O} = (\mathbf{L}_1, \mathbf{L}_2, A_2, \dots, \mathbf{L}_K, A_K, \mathbf{L}_{K+1})$  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  $\bar{\mathbf{a}} = (a_1, \dots, a_K)$ , there is a counterfactual random variable  $\mathbf{L}_{\bar{\mathbf{a}}} = \bar{\mathbf{L}}_{\bar{\mathbf{a}}, K+1}$  recording a subject's response history if treatment regime  $\bar{\mathbf{a}}$  was followed. We link the counterfactual data to the observed data through the consistency assumption  $\bar{\mathbf{L}}_{\bar{\mathbf{a}}, m} = \bar{\mathbf{L}}_m$  if  $\bar{\mathbf{A}}_{m-1} = \bar{\mathbf{a}}_{m-1}$  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  $\bar{\mathbf{a}}$  and  $m$

$$L_{\bar{\mathbf{a}}} \amalg A_m \mid \bar{\mathbf{L}}_m, \bar{\mathbf{A}}_{m-1} = \bar{\mathbf{a}}_{m-1} \quad (1)$$

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  $A_m$ ,

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

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

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

DR Estimation of  $E[Y_{\bar{\mathbf{a}}}]$

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

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



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  $m = K + 1, \dots, 2$ ,



4. a): specify and fit by IRLS a parametric regression model  $h_{m-1}(\bar{\mathbf{L}}_{m-1}, \bar{\mathbf{A}}_{m-1}; \beta_{m-1}, \phi_{m-1}) =$



$\Psi\{s_{m-1}(\bar{\mathbf{L}}_{m-1}, \bar{\mathbf{A}}_{m-1}; \beta_{m-1}) + \phi_{m-1}\bar{\pi}_{m-1}^{-1}(\hat{\alpha})\}$  for the conditional expectation  $E\{\hat{T}_m \mid \bar{\mathbf{A}}_{m-1}, \bar{\mathbf{L}}_{m-1}\}$ , where  $s_{m-1}(\bar{\mathbf{L}}_{m-1}, \bar{\mathbf{A}}_{m-1}; \beta_{m-1})$  is a known function with the unknown parameter  $\beta_{m-1}$ ,  $\Psi$  is the canonical link function of a given GLM, and  $\bar{\pi}_m(\hat{\alpha}) = \prod_{j=1}^m f(A_j \mid \bar{\mathbf{L}}_j, \bar{\mathbf{A}}_{j-1}; \hat{\alpha})$ .

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

$$\mathbf{0} = \tilde{E}\left[\left[\hat{T}_m - \Psi\{s_{m-1}(\bar{\mathbf{L}}_{m-1}, \bar{\mathbf{A}}_{m-1}; \hat{\beta}_{m-1}) + \hat{\phi}_{m-1}\bar{\pi}_{m-1}^{-1}(\hat{\alpha})\}\right] \left\{\partial s(\bar{\mathbf{L}}_{m-1}; \hat{\beta}_{m-1}) / \partial \beta'_{m-1}, \bar{\pi}_{m-1}^{-1}(\hat{\alpha})\right\}\right]$$

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

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

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

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

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

$$5. \text{ Finally } \hat{E}[Y_{\bar{a}}] = \tilde{E}[\hat{T}_1] = \tilde{E}[\hat{T}_1^{\bar{a}}]$$

