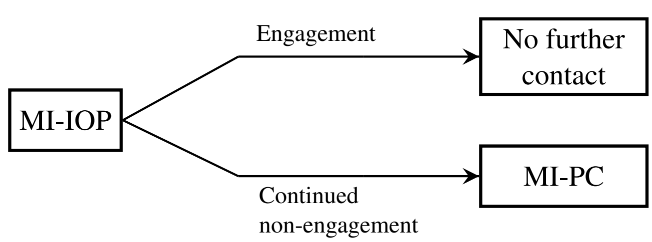
Notation Cheat Sheet for Oral Defense

Nicholas J. Seewald

May 10, 2021

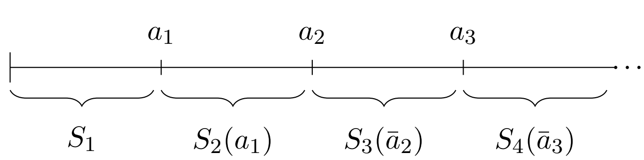
Contents

1 Dynamic Treatment Regimens (DTRs)

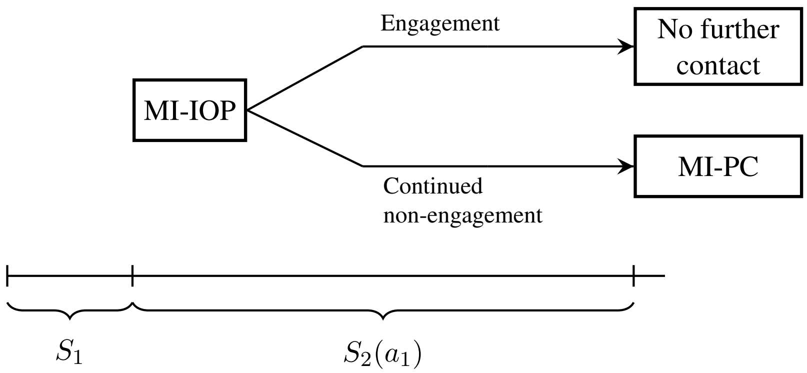


An example DTR

- **MI-IOP**: 2 motivational interviews to re-engage patient in intensive outpatient program
- **MI-PC**: 2 motivational interviews to engage patient in treatment of their choice.
- $a_1, \ldots, a_M$ indicate treatment recommendations
  - $\bar{a}_j := \{a_1, \ldots, a_j\}$
- $S_j(\bar{a}_{j-1})$ is all information collected after providing treatment a_{j-1} until just before providing a_j.
  - $\bar{S}_j(\bar{a}_{j-1}) := \{S_1, S_2(a_1), \ldots, S_j(\bar{a}_{j-1})\}$



- A **decision rule** φ_j is a function of $\bar{S}_j(\bar{a}_{j-1})$ which outputs a recommendation for subsequent treatment a_j.
  - Note that φ_j doesn't need to use *all* the information in $\bar{S}_j(\bar{a}_{j-1})$, but it can.
- An $\boldsymbol{M}$**-stage dynamic treatment regimen** is a sequence of M decision rules $\{\varphi_1, \ldots, \varphi_M\}$
- A **tailoring variable** is information contained in $S_j(\bar{a}_{j-1})$ which is used to inform recommendation to subsequent treatment
  - For example, "**response**" to prior treatment



- The above DTR can be written $\{\varphi_1, \varphi_2\}$, where

$$\begin{aligned}\varphi_1\left(S_1\right) &= \text{MI-IOP} \\ \varphi_2\left(\bar{S}_2(a_1)\right) &= R \cdot (\text{No further contact}) + (1-R) \cdot (\text{MI-PC})\end{aligned}$$

  where R is an indicator for response to MI-IOP

- For 2-stage DTRs, we can write $\{\varphi_1, \varphi_2\}$ as (a_1, a_{2R}, a_{2NR}) where a_1 is first-stage treatment, a_{2R} is second-stage treatment for *responders*, and a_{2NR} is second-stage treatment for *non-responders*.

2 Sequential, Multiple-Assignment Randomized Trials (SMARTs)

2.1 Observed Data

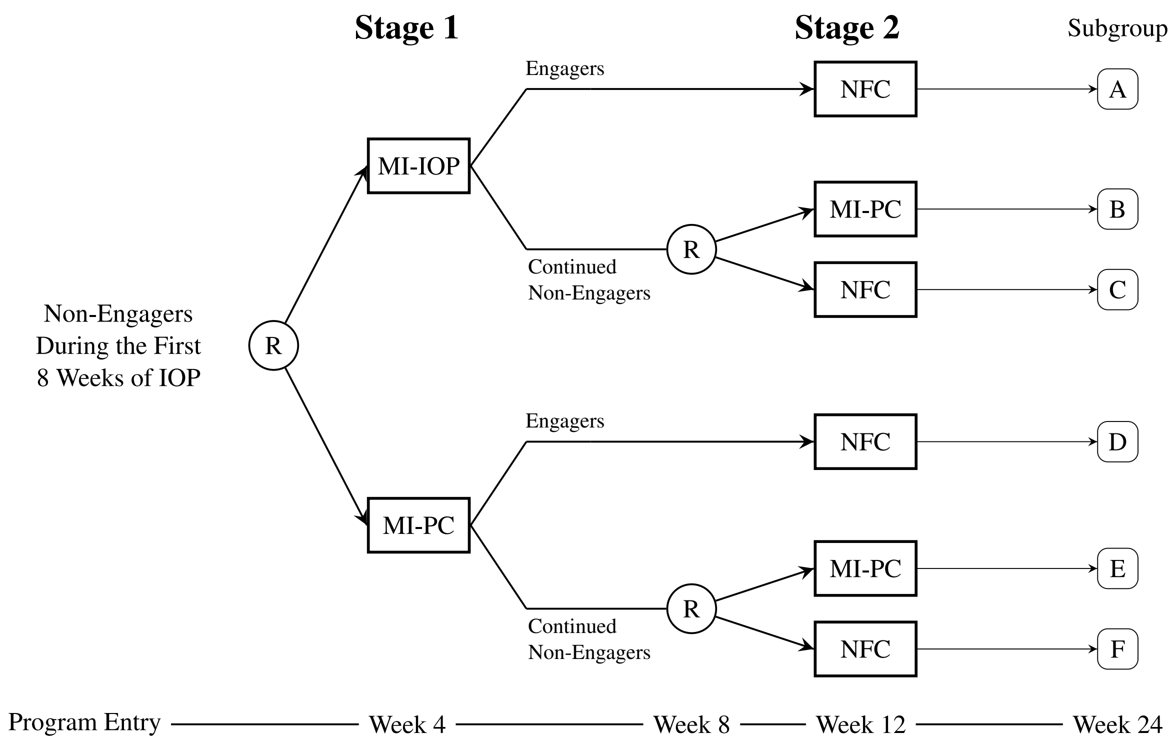
- We observe the outcome at T measurement occasions, with T_1 measurements in the first stage (including baseline) and T_2 in the second stage.
- For the ith individual, we collect

$$(\boldsymbol{X}_i, Y_{1,i}, A_{1,i}, \boldsymbol{Y}_{2:T_1,i}, R_i, A_{2,i}, \boldsymbol{Y}_{T_1+1:T,i})$$

  where
  - $\boldsymbol{X}_i$ is a vector of baseline covariates
  - $Y_{1,i}$ is the baseline outcome measurement

- $A_{1,i}$ is the randomly-assigned first-stage treatment
- $\boldsymbol{Y}_{2:T_1,i}$ is a vector of outcome measurements in the first stage (for times $2, \ldots, T_1$)
- R_i is response status (1 if responder, 0 if non-responder); the tailoring variable
- $A_{2,i}$ is the randomly-assigned second-stage treatment
- $\boldsymbol{Y}_{T_1+1:T,i}$ is a vector of outcome measurements in the first stage (for times $T_1 + 1, \ldots, T$)

2.2 The ENGAGE Study



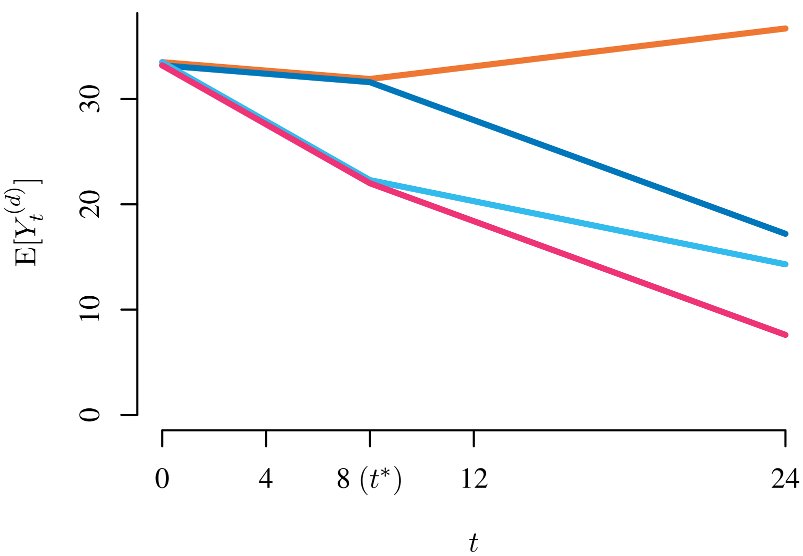
The ENGAGE SMART. **MI-IOP** is two motivational interviews to re-engage patient in *intensive outpatient program*; **MI-PC** is two motivational interviews to engage patient in treatment of *patient's choice*; **NFC** is *no further contact*.

2.3 Modeling and Estimation

- An example marginal structural mean model for ENGAGE is

$$\begin{aligned}\mu_t^{(d)}(\boldsymbol{X};\boldsymbol{\theta}) = \boldsymbol{\eta}^\top \boldsymbol{X} + \beta_0 + \mathbb{1}_{\{t \le t^*\}} \left(\beta_1 t + \beta_2 a_1 t\right) \\ + \mathbb{1}_{\{t > t^*\}} \left(\beta_1 t^* + \beta_2 t^* a_1 + \beta_3 (t - t^*) + \beta_4 (t - t^*) a_1 \right. \\ \left. + \beta_5 (t - t^*) a_{2NR} + \beta_6 (t - t^*) a_1 a_{2NR}\right),\end{aligned}$$

  where $\mathbb{1}_{\{E\}}$ is the indicator function for the event E.



Visualization of example marginal structural mean model for ENGAGE

- An M-estimator for the regression parameters $\boldsymbol{\theta} = (\boldsymbol{\eta}^\top, \boldsymbol{\beta}^\top)^\top$ solves

$$0 = \sum_{i=1}^{N} \sum_{d} \left[\frac{I^{(d)}(A_{1,i}, R_i, A_{2,i})}{P(A_{1,i} = a_1) P(A_{2,i} = a_2 \mid A_{1,i} = a_1, R_i)} \cdot \left(\boldsymbol{D}^{(d)}(\boldsymbol{X}_i) \right)^\top \cdot \boldsymbol{V}^{(d)} \left(\boldsymbol{X}_i; \boldsymbol{\tau} \right)^{-1} \cdot \left(\boldsymbol{Y}_i - \boldsymbol{\mu}^{(d)}(\boldsymbol{X}_i; \boldsymbol{\eta}, \boldsymbol{\beta}) \right) \right],$$

  where

  - $I^{(d)}(A_{1,i}, R_i, A_{2,i}) = \mathbb{1}_{\{A_{1,i}=a_1\}} \left(R_i + (1 - R_i)\, \mathbb{1}_{\{A_{2,i}=a_2\}} \right)$ for ENGAGE
  - $\boldsymbol{D}^{(d)}(\boldsymbol{X}_i) = \frac{\partial}{\partial(\boldsymbol{\eta}^\top, \boldsymbol{\beta}^\top)^\top} \boldsymbol{\mu}^{(d)}(\boldsymbol{X}_i; \boldsymbol{\eta}, \boldsymbol{\beta})$
  - $\boldsymbol{V}^{(d)} \left(\boldsymbol{X}_i; \boldsymbol{\tau} \right)$ is a working model for $\mathbf{Var} \left(\boldsymbol{Y}^{(d)} - \boldsymbol{\mu}^{(d)}(\boldsymbol{X}_i; \boldsymbol{\eta}, \boldsymbol{\beta}) \right)$
  - d specifies an embedded DTR

- $W^{(d)}(A_{1,i}, R_i, A_{2,i}) = \frac{I^{(d)}(A_{1,i}, R_i, A_{2,i})}{P(A_{1,i}=a_1) P(A_{2,i}=a_2 | A_{1,i}=a_1, R_i)}$ for a DTR d

Proposition. *Suppose* $\boldsymbol{\mu}^{(d)} \left(\boldsymbol{X}; \boldsymbol{\theta} \right)$ *is a correctly-specified model for* $\mathrm{E} \left[\boldsymbol{Y}^{(d)} \mid \boldsymbol{X} \right]$*, then* $\boldsymbol{\theta}_n$ *is consistent for* $\boldsymbol{\theta}^*$*, the true parameter value.*

Proposition. $\sqrt{n} \left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta} \right)$ *converges in distribution to* $\mathcal{N} \left(\mathbf{0}, \boldsymbol{B}^{-1} \boldsymbol{M} \boldsymbol{B}^{-1} \right)$*, where*

$$\boldsymbol{B} := \mathrm{E} \left[\sum_{d \in \mathcal{D}} W^{(d)} \left(A_{1,i}, R_i, A_{2,i} \right) \boldsymbol{D}^{(d)}(\boldsymbol{X}_i)^\top \boldsymbol{V}^{(d)}(\boldsymbol{X}_i; \boldsymbol{\tau})^{-1} \boldsymbol{D}^{(d)}(\boldsymbol{X}_i) \right] \in \mathbb{R}^{p \times p}$$

and

$$\boldsymbol{M} := \mathrm{E}\left[\left(\sum_{d\in\mathcal{D}} W^{(d)}\left(A_{1,i}, R_i, A_{2,i}\right) \boldsymbol{D}^{(d)}(\boldsymbol{X}_i)^\top \boldsymbol{V}^{(d)}(\boldsymbol{X}_i;\boldsymbol{\tau})^{-1}\left(\boldsymbol{Y}_i - \boldsymbol{\mu}^{(d)}(\boldsymbol{X}_i;\boldsymbol{\theta})\right)\right)^{\otimes 2}\right] \in \mathbb{R}^{p\times p},$$

with $\boldsymbol{Z}^{\otimes 2} = \boldsymbol{Z}\boldsymbol{Z}^\top$.

3 Sample Size for Comparing DTRs in Longitudinal SMARTs

- **Simplifications:**
  - No baseline covariates (conservative)
  - Measurement occasions are equally spaced in both stages
    - T_1 measurements in stage 1 (includes baseline)
    - T_2 measurements in stage 2
    - $T = T_1 + T_2$ total measurements
- **Estimand of interest:** $\mathrm{E}\left[Y_T^{(1,a_{2R},a_{2NR})} - Y_T^{(-1,a'_{2R},a'_{2NR})}\right]$
- **Hypothesis test of interest:**

$$H_0 : \mathrm{E}\left[Y_T^{(1,a_{2R},a_{2NR})} - Y_T^{(-1,a'_{2R},a'_{2NR})}\right] = 0 \quad \text{vs.} \quad H_1 : \mathrm{E}\left[Y_T^{(1,a_{2R},a_{2NR})} - Y_T^{(-1,a'_{2R},a'_{2NR})}\right] = \Delta$$

- Under our example model, $\mathrm{E}\left[Y_T^{(-1,a_{2R},a_{2NR})} - Y_T^{(-1,a'_{2R},a'_{2NR})}\right] = \boldsymbol{c}^\top\boldsymbol{\beta}$
- **Working Assumptions:**
  1. *Constrained conditional variability:*
     (a) For all embedded DTRs d,

$$\mathrm{E}\left[\left(\boldsymbol{Y}_i^{(d)} - \boldsymbol{\mu}^{(d)}\right)^{\otimes 2} \mid R_i^{(d)} = 1\right] - \mathrm{E}\left[\left(\boldsymbol{Y}_i^{(d)} - \boldsymbol{\mu}^{(d)}\right)^{\otimes 2}\right]$$

     is positive semi-definite.

     (b) For all embedded DTRs d,

$$\frac{1}{P\left(R_i^{(d)} = 1\right)}\mathrm{E}\left[\left(\boldsymbol{Y}_i^{(d)} - \boldsymbol{\mu}^{(d)}\right)^{\otimes 2}\right] - \mathrm{E}\left[\left(\boldsymbol{Y}_i^{(d)} - \boldsymbol{\mu}^{(d)}\right)^{\otimes 2} \mid R_i^{(d)} = 1\right]$$

     is positive semi-definite.

  2. *Constrained Conditional Means.* For every embedded DTR d and and all embedded DTRs d' such that $a_1^{(d)} \neq a_1^{(d')}$,

$$\left(\mathrm{E}\left[\boldsymbol{Y}_i^{(d)} \mid R_i^{(d)} = 1\right] - \boldsymbol{\mu}^{(d)}\right)\left(\boldsymbol{\mu}^{(d)} - \boldsymbol{\mu}^{(d')}\right)^\top$$

     is "small".

- **Sample Size Formula:**

$$n \geq \frac{4\left(z_{1-\alpha/2} + z_{1-\beta}\right)^2}{\delta^2} \cdot \mathrm{DE}(\boldsymbol{r}) \cdot \omega(\rho, T, T_2)$$

where

- $\delta = \Delta/\sigma$ is the target standardized effect size
- α is the desired type-I error
- $1 - \gamma$ is the desired power
- $\rho = \mathrm{cor}(Y_t, Y_{t'})$ for $t \neq t'$
- T_1 is the number of measurements in stage 1
- T_2 is the number of measurements in stage 2
- $\mathrm{DE}(\boldsymbol{r})$ is a SMART design-specific inflation factor whcih depends on response probabilities $\boldsymbol{r}$
  - For an ENGAGE-type SMART, $\mathrm{DE}(\boldsymbol{r}) = 2 - (r_1 + r_{-1})/2$
  - For a SMART in which all participants are re-randomzized, $\mathrm{DE}(\boldsymbol{r}) = 2$
  - For a SMART in which only responders to first-stage treatment 1 are re-randomized, $\mathrm{DE}(\boldsymbol{r}) = (3 - r_1)/2$
- $\omega(\rho, T, T_2) = \frac{f(\rho,T,T_2)}{g(\rho,T,T_2)}$ is a deflation factor that accounts for within-person correlation where

$$f(\rho, T, T_2) = 6(1-\rho)(T-1)\left(\rho(T-1)\left((T-1)T_2 - T_2^2 + 2\right) + 4T_2(T - T_2 - 1) + 2\right)$$

and

$$g(\rho, T, T_2) = (T_2 + 1)\left(2\left(T^2(4T_2 + 2) - T(T_2(5T_2 + 9) + 1) + T_2(T_2 + 2)^2\right) + \rho(T-1)(T - T_2 - 2)(2TT_2 + T - 2T_2(T_2 + 2))\right)$$