# Model-Based Inference for Completely Randomized Experiments

DongYoon Yang

July 11, 2022

Seoul National University

# Table of Contents

# Table of Contents

# Completely randomized experiments

A completely randomized experiment is a classical randomized experiment with an assignment mechanism satisfying b

$$\mathbb{W}^{+} = \left\{ \mathbb{W} \in \mathbb{W} \mid \sum_{i=1}^{N} W_i = N_t \right\} \tag{1}$$

for some present $N_t \in \{1, 2, \cdots, N-1\}$

- $\left( \begin{array}{c} N \\ N_{\mathrm{t}} \end{array} \right)$ assignment vectors in $\mathbb{W}^{+}$ are equally likely.
- $Pr(\boldsymbol{W}|\boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)) = \begin{cases} \left( \begin{array}{c} N \\ N_{\mathrm{t}} \end{array} \right)^{-1}, & \text{if} \sum_{i=1}^{n} = N_t \\ 0, & \text{otherwise} \end{cases}$

# Missing values for estimating treatment effect

$$\tau_{\text{fs}} = \tau(\boldsymbol{Y}(0), \boldsymbol{Y}(1)) = \frac{1}{N} \cdot \sum_{i=1}^{N} (Y_i(1) - Y_i(0)) \quad (2)$$

$$Y_i(0) = \begin{cases} Y_i^{\text{mis}} & \text{if } W_i = 1, \\ Y_i^{\text{obs}} & \text{if } W_i = 0, \end{cases} \quad \text{and} \quad Y_i(1) = \begin{cases} Y_i^{\text{mis}} & \text{if } W_i = 0 \\ Y_i^{\text{obs}} & \text{if } W_i = 1 \end{cases} \quad (3)$$

$$\begin{aligned} \tau_{\text{fs}} &= \tau(\boldsymbol{Y}(0), \boldsymbol{Y}(1)) \\ &= \frac{1}{N} \cdot \sum_{i}^{N} \left( \left( W_i \cdot Y_i^{\text{obs}} + (1 - W_i) \cdot Y_i^{\text{mis}} \right) - \left( (1 - W_i) \cdot Y_i^{\text{obs}} + W_i \cdot Y_i^{\text{mis}} \right) \right) \\ &= \frac{1}{N} \cdot \sum_{i=1}^{N} \left( (2 \cdot W_i - 1) \cdot \left( Y_i^{\text{obs}} - Y_i^{\text{mis}} \right) \right) \\ &= \tau \left( \boldsymbol{Y}^{\text{obs}}, \boldsymbol{Y}^{\text{mis}}, \boldsymbol{W} \right) \end{aligned} \quad (4)$$

$$\hat{\tau} = \hat{\tau} \left( \boldsymbol{Y}^{\text{obs}}, \hat{\boldsymbol{Y}}^{\text{mis}}, \boldsymbol{W} \right) = \frac{1}{N} \cdot \sum_{i=1}^{N} \left( (2 \cdot W_i - 1) \cdot \left( Y_i^{\text{obs}} - \hat{Y}_i^{\text{mis}} \right) \right) \quad (5)$$

# Table of Contents

# Single Imputation

| Unit | Potential Outcomes | | Treatment | Observed Outcome |
|---|---|---|---|---|
| | $Y_i(0)$ | $Y_i(1)$ | $W_i$ | $Y_i^{\text{obs}}$ |
| 1 | 0 | (12.8) | 0 | 0 |
| 2 | (4.13) | 9.9 | 1 | 9.9 |
| 3 | 12.4 | (12.8) | 0 | 12.4 |
| 4 | (4.13) | 3.6 | 1 | 3.6 |
| 5 | 0 | (12.8) | 0 | 0 |
| 6 | (4.13) | 24.9 | 1 | 24.9 |
| Average | 4.13 | 12.8 | | |
| Diff (ATE): | 8.67 | | | |

# Multiple Imputation

| Unit | Potential Outcomes | | Treatment | Observed Outcome |
|---|---|---|---|---|
| | $Y_i(0)$ | $Y_i(1)$ | $W_i$ | $Y_i^{\text{obs}}$ |
| Panel A: First draw | | | | |
| 1 | 0 | (3.6) | 0 | 0 |
| 2 | (12.4) | 9.9 | 1 | 9.9 |
| 3 | 12.4 | (9.9) | 0 | 12.4 |
| 4 | (12.4) | 3.6 | 1 | 3.6 |
| 5 | 0 | (9.9) | 0 | 0 |
| 6 | (0) | 24.9 | 1 | 24.9 |
| Average | 6.2 | 10.3 | | |
| Diff (ATE): | 4.1 | | | |
| Panel B: Second draw | | | | |
| 1 | 0 | (9.9) | 0 | 0 |
| 2 | (0) | 9.9 | 1 | 9.9 |
| 3 | 12.4 | (24.9) | 0 | 12.4 |
| 4 | (0) | 3.6 | 1 | 3.6 |
| 5 | 0 | (3.6) | 0 | 0 |
| 6 | (0) | 24.9 | 1 | 24.9 |
| Average | 2.1 | 12.8 | | |
| Diff (ATE): | 10.7 | | | |

# Table of Contents

# The four steps of the bayesian approach

1. Derivation of $p(\boldsymbol{Y}^{\text{mis}}|\boldsymbol{Y}^{\text{obs}}, W, \theta)$
   - $p(\boldsymbol{Y}(0), \boldsymbol{Y}(1)|W, \theta) = p(\boldsymbol{Y}(0), \boldsymbol{Y}(1)|\theta)$
     $\because$ completely randomized experiment
2. Derivation of $p(\theta|\boldsymbol{Y}^{\text{obs}}, W)$
   - Assume $\theta \sim p(\theta)$
3. Derivation of $p(\boldsymbol{Y}^{\text{mis}}|\boldsymbol{Y}^{\text{obs}}, W)$
   - Use 1, 2 and integrating out $\theta$.
4. Derivation of $p(\tau|\boldsymbol{Y}^{\text{obs}}, W)$

# Simulation Example

$$\begin{pmatrix} Y_i(0) \\ Y_i(1) \end{pmatrix} \mid \mu_c, \mu_t \sim \mathcal{N}\left(\begin{pmatrix} \mu_c \\ \mu_t \end{pmatrix}, \begin{pmatrix} 100 & 0 \\ 0 & 64 \end{pmatrix}\right) \tag{6}$$

1. Get $p(\boldsymbol{Y}^{\text{mis}}|\boldsymbol{Y}^{\text{obs}}, W, \mu_c, \mu_t)$ by step 1.
2. Get $p(\mu_c, \mu_t|\boldsymbol{Y}^{\text{obs}}, W)$ by step 2.
3. Sample the $\mu_c^{(1)}$ and $\mu_t^{(1)}$ from $p(\mu_c, \mu_t|\boldsymbol{Y}^{\text{obs}}, W)$.
4. Impute the $\boldsymbol{Y}^{\text{mis},(1)}$ by 1 and 3.
5. Compute $\hat{\tau}^{(1)}$
6. Repeat $1 \sim 3$, get $\hat{\tau}^{(k)}, k = 1, \cdots, K$
7. Get sample mean and standard deviation.

# Dependence between potential outcomes

$$\begin{pmatrix} Y_i(0) \\ Y_i(1) \end{pmatrix} \mid \boldsymbol{\theta} \sim \mathcal{N}\left( \begin{pmatrix} \mu_c \\ \mu_t \end{pmatrix}, \begin{pmatrix} \sigma_t^2 & \rho\sigma_t\sigma_c \\ \rho\sigma_t\sigma_c & \sigma_c^2 \end{pmatrix} \right) \quad (7)$$

- $\boldsymbol{\theta} = (\mu_c, \mu_t, \sigma_t, \sigma_c, \rho)$
- **Step 1.**

$$\begin{aligned} & Y_i^{\mathrm{mis}} \mid Y_i^{\mathrm{obs}}, \mathsf{W}_i, \mathsf{X}_i, \boldsymbol{\theta} \\ & \sim \mathcal{N}\bigg( W_i \cdot (\mu_c + \frac{\sigma_c}{\sigma_t}(Y_i^{\mathrm{obs}} - \mu_t)) + (1 - W_i) \cdot (\mu_t + \frac{\sigma_t}{\sigma_c}(Y_i^{\mathrm{obs}} - \mu_c) \\ & \qquad (1 - \rho^2)(W_i \cdot \sigma_t^2 + (1 - W_i) \cdot \sigma_t^2) \bigg) \end{aligned}$$

# Table of Contents

# Bayesian model-based imputation with covariates

$$\begin{pmatrix} Y_i(0) \\ Y_i(1) \end{pmatrix} \Bigg| X_i, \boldsymbol{\theta} \sim \mathcal{N}\left(\begin{pmatrix} X_i\beta_c \\ X_i\beta_t \end{pmatrix}, \begin{pmatrix} \sigma_c^2 & 0 \\ 0 & \sigma_t^2 \end{pmatrix}\right) \tag{8}$$

- $\boldsymbol{\theta} = (\beta_c, \beta_t, \sigma_c, \sigma_t)$
- $Y_i^{\text{mis}} \mid Y_i^{\text{obs}}, \mathrm{W}_i, \mathrm{X}_i, \theta \sim$ $\mathcal{N}\left(W_i \cdot X_i\beta_c + (1 - W_i) \cdot X_i\beta_t, W_i \cdot \sigma_t^2 + (1 - W_i) \cdot \sigma_t^2\right)$
- Analytic solutions are difficult to obtain. In this case, we can get $\hat{\tau}$ by simulation study.