Chapter 10. Pairwise Randomized Experiments

이종진

Seoul National University

ga0408@snu.ac.kr

July 19, 2022

Table of Contents

10.1 Introduction

- Pairwise Randomized Expereiments
  - A special case of the stratified randomized experiments
  - Each strata contain two units (1 unit for treatment group and 1 unit for control group)
- Two special features of this designed
  - Only a single unit in each treatment and control group.
  - Each stratum has the same proportion of treated units.

10.2 The Children's Television Workshop Experiment data

- Test for the effect of educuation television program aimed at improving reading skills.
- Several schools, a pair of two classes
- Pair / Treatment / Pre-test scores / Post-test scores

Pair G_i	Treatment W_i	Pre-Test Score X_i	Post-Test Score Y_i^{obs}	Normalized Rank Post-Test Score R_i
1	0	12.9	54.6	−7.5
1	1	12.0	60.6	2.5
2	0	15.1	56.5	−4.5
2	1	12.3	55.5	5.5

10.3 Pairwise Randomized Experiments

- N: the number of units
- J: the number of strata
- $N(j)$: the number of units in j-th stratum, $j \in \{1, \ldots, J\}$
- $N_t(j)$: the number of treated unit in j-th stratum, $j \in \{1, \ldots, J\}$
- $N_c(j)$: the number of controlled unit in stratum j-th stratum, $j \in \{1, \ldots, J\}$
- G_i: the variable indicating pair where i-th units in, $G_i \in \{1, \ldots, J\}$

Pairwise Randomized Experiments

- $J = N/2$
- $N(j) = 2, N_t(j) = N_c(j) = 1,\ j \in \{1, \ldots, J\}$
- $G_i \in \{1, \ldots, N/2\}$
- The probability for any assignment vector W

$$p(\mathrm{W} \mid \mathrm{X}, \mathrm{Y}(0), \mathrm{Y}(1)) = \prod_{j=1}^{N/2} \begin{pmatrix} N(j) \\ N_t(j) \end{pmatrix}^{-1} = \prod_{j=1}^{N/2} \frac{1}{2} = 2^{-N/2}, \quad \text{for } \mathrm{W} \in \mathbb{W}^+,$$

where $\mathbb{W}^+ = \left\{\mathrm{W} \mid \sum_{i:G_i=j} W_i = 1 \text{ for } j = 1, \ldots, N/2\right\}$.

Pairwise Randomized Experiments

- Denote arbitrary label of the two units within the pair as A and B.
- $W_{j,A}$, $W_{j,B}$: treatment indicators for these unit, $W_{j,A} = 1 - W_{j,B}$
- Define the pair of observed variables.
- $(Y_{j,A}(0), Y_{j,A}(1))$, $(Y_{j,B}(0), Y_{j,B}(1))$ potential outcomes for units A and B in pair j

Pairwise Randomized Experiments

- Observed variable (A/B)

$$Y^{\textsf{obs}}_{j,A} = \begin{cases} Y_{j,A}(0) & \text{if } W_{j,A} = 0, \\ Y_{j,A}(1) & \text{if } W_{j,A} = 1, \end{cases} \quad \text{and} \quad Y^{\textsf{obs}}_{j,B} = \begin{cases} Y_{j,B}(0) & \text{if } W_{j,A} = 1, \\ Y_{j,B}(1) & \text{if } W_{ji,A} = 0. \end{cases}$$

- Observed variable $(trt/ctrl)$

$$Y^{\textsf{obs}}_{j,c} = \begin{cases} Y^{\text{obs}}_{j,A} & \text{if } W_{i,A} = 0, \\ Y^{\text{obs}}_{j,B} & \text{if } W_{i,A} = 1, \end{cases} \quad \text{and} \quad Y^{\text{obs}}_{j,t} = \begin{cases} Y^{\text{obs}}_{j,B} & \text{if } W_{i,A} = 0, \\ Y^{\text{obs}}_{j,A} & \text{if } W_{i,A} = 1 \end{cases}$$

Pairwise Randomized Experiments

- The average treatment effect within pair j

$$\tau_{\text{pair}}(j) = \frac{1}{2} \sum_{i:G_i=j} (Y_i(1) - Y_i(0)) = \frac{1}{2} ((Y_{j,A}(1) - Y_{j,A}(0)) + (Y_{j,B}(1) - Y_{j,B}(0)))$$

- The finite-sample average treatment effect

$$\tau_{\text{fs}} = \frac{1}{N} \sum_{i=1}^{N} (Y_i(1) - Y_i(0)) = \frac{2}{N} \sum_{j=1}^{N/2} \tau_{\text{pair}}(j)$$

10.4 Fisher's Exact P-values in Pairwise Randomized Experiments

- Fisher null hypothesis

$$H_0: \quad Y_i(0) = Y_i(1), \text{ for all } i = 1, \ldots, N$$

- T^{dif}

$$T^{\mathrm{dif}} = \left| \frac{1}{J} \sum_{j=1}^{J} \left(Y_{j,t}^{\mathrm{obs}} - Y_{j,c}^{\mathrm{obs}} \right) \right|$$

$$= \left| \frac{2}{N} \sum_{j=1}^{N/2} \left(W_{i,A} \cdot \left(Y_{j,A}^{\mathrm{obs}} - Y_{j,B}^{\mathrm{obs}} \right) + (1 - W_{i,A}) \cdot \left(Y_{j,B}^{\mathrm{obs}} - Y_{j,A}^{\mathrm{obs}} \right) \right) \right| .$$

- Identical statistics; $T^{\mathbf{dif}} = \left| \bar{Y}_{\mathrm{t}}^{\mathbf{obs}} - \bar{Y}_{\mathrm{c}}^{\mathrm{obs}} \right|$, however, the p-value for this statistics will be different.

Rank-statistics

- R_i: normalized rank of Y_i^{obs}.
- $R_{j,A}$: corresponding a normalized rank for the given unit.
- T^{rank}

$$T^{\mathrm{rank}} = \left|\bar{R}_t - \bar{R}_c\right| = \left|\frac{2}{N}\sum_{j=1}^{N/2}\left(W_{j,A}\cdot(R_{j,A} - R_{j,B}) + (1 - W_{j,A})\cdot(R_{j,B} - R_{j,A})\right)\right|$$

- $T^{\text{rank,pair}}$

$$T^{\text{rank,pair}} = \left|\frac{2}{N}\sum_{j=1}^{N/2}\left(1_{Y_{j,t}^{\mathrm{obs}} > Y_{j,c}^{\mathrm{obs}}} - 1_{Y_{j,t}^{\mathrm{obs}} < Y_{j,c}^{\mathrm{obs}}}\right)\right|$$

- Substantial variation in the level of the outcomes between the pairs, $T^{\text{rank,pair}} > T^{\mathrm{rank}}$ (power)

10.5 The Analysis of Pairwise Randomized Experiments from Neyman's Repeated Sampling Perspective

- The estimator for average treatment effect in pair j,

$$\hat{\tau}^{\text{pair}}(j) = Y_{j,t}^{\text{obs}} - Y_{j,c}^{\text{obs}} = \sum_{i:G_i=j} (2 \cdot W_i - 1) \cdot Y_i^{\text{obs}}.$$

- The estimator for finite-sample average treatment effect,

$$\hat{\tau}^{\text{fs}} = \frac{1}{N/2} \sum_{j=1}^{N/2} \hat{\tau}^{\text{pair}}(j) = \bar{Y}_{\text{t}}^{\text{obs}} - \bar{Y}_{\text{c}}^{\text{obs}}$$

- The sampling variance

$$\mathbb{V}_W\left(\hat{\tau}^{\text{fs}}\right) = \frac{1}{(N/2)^2} \sum_{j=1}^{N/2} \left(S_c^2(j) + S_t^2(j) - \frac{S_{ct}^2(j)}{2} \right)$$

where $S_c^2(j) = \sum_{i:G_i=j} \left(Y_i(0) - \bar{Y}_j(0)\right)^2 = \frac{1}{2} \cdot (Y_{j,A}(0) - Y_{j,B}(0))^2$,
$S_t^2(j) = \sum_{i:G_i=j} \left(Y_i(1) - \bar{Y}_j(1)\right)^2 = \frac{1}{2} \cdot (Y_{j,A}(1) - Y_{j,B}(1))^2$,
$S_{ct}^2(j) = \frac{1}{2} \cdot ((Y_{j,A}(1) - Y_{j,A}(0)) - (Y_{j,B}(1) - Y_{j,B}(0)))^2$,

$\bar{Y}_j(0) = \frac{1}{2} \cdot (Y_{j,A}(0) + Y_{j,B}(0))$ and $\bar{Y}_j(1) = \frac{1}{2} \cdot (Y_{j,A}(1) + Y_{j,B}(1))$

Neyman sampling variance

- The Neyman sampling variance

$$\hat{\mathbb{V}}^{\text{neyman}}\left(\hat{\tau}^{\text{fs}}\right) = \frac{1}{(N/2)^2}\sum_{j=1}^{N/2}\left(s_c^2(j) + s_t^2(j)\right)$$

  where $s_c^2(j) = \frac{1}{N_\text{c}(j)-1}\sum_{i:G_i=j,W_i=0}\left(Y_i^{\text{obs}} - \bar{Y}_\text{c}^{\text{obs}}\right)^2$ and $s_t^2(j) = \frac{1}{N_\text{t}(j)-1}\sum_{i:G_i=j,W_i=1}\left(Y_i^{\text{obs}} - \bar{Y}_\text{t}^{\text{obs}}\right)^2$.

- $s_c^2(j)$ and $s_t^2(j)$ can not be used.

Solution

- Treatment effect is constant and **across pairs**

$$\tau_{\text{pair}}(j) = \tau_{\text{S}} \text{ for all } j$$

- The sampling variance can be written

$$\mathbb{V}_W\left(\hat{\tau}^{\text{fs}}\right) = \frac{1}{(N/2)^2} \sum_{j=1}^{N/2} \left(S_c^2(j) + S_t^2(j) - \frac{S_{ct}^2(j)}{2} \right) = \frac{4}{N} \cdot S^2$$

where $S^2 := S^2(j) = S_t^2(j) = S_c^2(j)$ for all j.

- It can be estimated by

$$\hat{\mathbb{V}}^{\text{pair}}\left(\hat{\tau}^{\text{fs}}\right) = \frac{4}{N \cdot (N-2)} \cdot \sum_{j=1}^{N/2} \left(\hat{\tau}^{\text{pair}}(j) - \hat{\tau}^{\text{fs}}\right)^2 .$$

Thm 10.1

- The fact from the thm 10.1

$$\mathbb{E}\left[\hat{\mathbb{V}}^{\textsf{pair}}\left(\hat{\tau}^{\mathrm{fs}}\right)\right] = \mathbb{V}_W\left(\hat{\tau}^{\mathrm{fs}}\right) + \frac{4}{N \cdot (N-2)} \cdot \sum_{j=1}^{N/2} \left(\tau_{\mathrm{pair}}(j) - \tau\right)^2$$

- Heterogenity in the treatment effects → upwardly biased and C.I. will be conservative.

10.6 Regression-Based Analysis of Pairwise Randomized Experiments

- Primary difference
  - Outcome of regression ($\tau^{\text{pair}}(j) = Y_{j,t} - Y_{j,c}$)
  - Consider pair as the unit of analysis.
- $\tau^{\text{pair}}(j) = \tau_{\text{sp}} + \varepsilon_j$. ($\tau_{\text{sp}} = \mathbb{E}_{\text{sp}}[\tau_{\text{pair}}(j)]$)
- How to include additional covariates
  - $\Delta_{X,j} = \left(W_{j,A} \cdot (X_{j,A} - X_{j,B}) + (1 - W_{j,A}) \cdot (X_{j,B} - X_{j,A})\right)$
  - $\bar{X}_j = (X_{j,A} + X_{j,B})/2$
- The general version of regression model

$$\tau^{\text{pair}}(j) = \tau + \beta \cdot \Delta_{X,j} + \gamma \cdot \left(\bar{X}_j - \bar{X}\right) + \varepsilon_j$$

10.7 Model-Based Analysis of Pairwise Randomized Experiments

- Model-based imputation approach
  - Little different from that for the case of stratified randomized experiments.
  - One unit (trt/crtl) $\rightarrow$ Not flexible to specify the joint distributions of the potential outcomes
- Some structure on the variance within pairs, such as a hierarchical structure

Hierarchical Structure

- Independent priors for μ, σ_μ^2, σ_c^2, σ_t^2, γ, and β
- Conditional on pair indicators, covariates, and parameters.

$$\begin{pmatrix} Y_i(0) \\ Y_i(1) \end{pmatrix} \mid G_i = j, X_i = x, \mu(1), \ldots, \mu(N/2), \gamma, \beta, \sigma_c^2, \sigma_t^2$$

$$\sim \mathcal{N}\left(\begin{pmatrix} \mu(j) + x \cdot \beta \\ \mu(j) + \gamma + x \cdot \beta \end{pmatrix}, \begin{pmatrix} \sigma_c^2 & 0 \\ 0 & \sigma_t^2 \end{pmatrix} \right).$$

- Specification for the pair-specific means

$$\begin{pmatrix} \mu(1) \\ \vdots \\ \mu(N/2) \end{pmatrix} \mid \mathrm{G}, \mathrm{X}, \mathrm{W}, \gamma, \beta, \sigma_c^2, \sigma_t^2, \mu \sim \mathcal{N}\left(\begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix}, \begin{pmatrix} \sigma_\mu^2 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \sigma_\mu^2 \end{pmatrix} \right)$$

10.8 Conclusion

- The Fisher exact p-value is conceptually not affected
- The Neyman repeated sampling variance
  - A homogeneity of treatment effect.
- In the regression analyses
  - Regression outcome and pair as the unit of analysis.
- Model-based analyses are modified
  - A hierarchical structure

The End