Estimation of Multivariate Treatment Effects in Contaminated Randomized Trials

Zi Ye

Assistant Professor of Statistics
Lehigh University

(Joint work with Solomon W. Harrar, University of Kentucky)

PSI Journal Club, 7th September, 2022

Outline

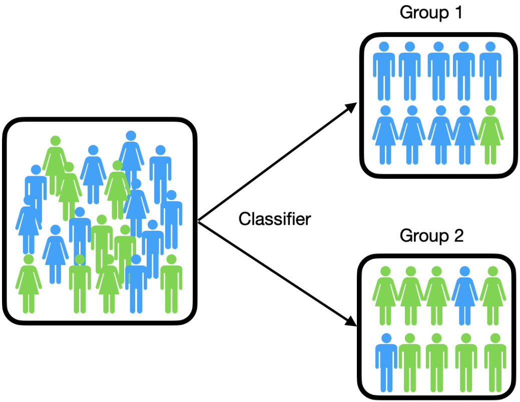
Introduction

Estimation and Hypothesis Testing

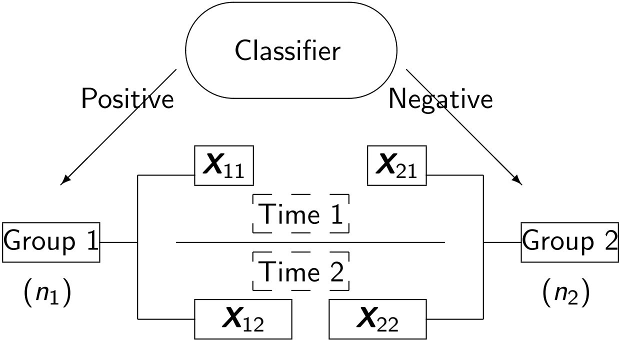
Sample Size Determination

Real Data Analysis

Classification Errors



Pre-Post Design



Probabilities of Classification Errors

- Positive Predictive Value (PPV): is the probability $(1 - \varepsilon)$ that a person with a positive test will have the clinical condition of interest.
- Negative Predictive Value (NPV): is the probability $(1 - \delta)$ that a person with a negative test will be free from the clinical condition of interest.
- PPV (NPV) combine sensitivity (specificity) with prevalence information to provide accuracy of a test result.

Probabilities of Classification Errors

- Problem: Not all classifiers are perfect with 100% PPV and NPV.
    - Ignoring these errors will produce **BIASED** results.
        - The expected outcomes are affected by the contamination:

$$
\begin{aligned}
E(\boldsymbol{X}_{11k}) &= (1-\delta_1)\boldsymbol{\mu}_1 + \delta_1\boldsymbol{\mu}_2 \\
E(\boldsymbol{X}_{21k}) &= \delta_2\boldsymbol{\mu}_1 + (1-\delta_2)\boldsymbol{\mu}_2 \\
E(\boldsymbol{X}_{12k}) &= (1-\delta_1)(\boldsymbol{\mu}_1 + \boldsymbol{\tau}_1) + \delta_1(\boldsymbol{\mu}_2 + \boldsymbol{\tau}_2) \\
E(\boldsymbol{X}_{22k}) &= \delta_2(\boldsymbol{\mu}_1 + \boldsymbol{\tau}_1) + (1-\delta_2)(\boldsymbol{\mu}_2 + \boldsymbol{\tau}_2)
\end{aligned}
$$

    - The sample size and the power calculations will produce overly optimistic results.
    - Under-powered studies can fail to detect a significant effect when, in fact, it is present.

Multivariate Normal Model

- The parameter of interest is still

$$\boldsymbol{\Delta} = \boldsymbol{\tau}_1 - \boldsymbol{\tau}_2.$$

- Assume $0 < \delta_1, \delta_2 < 1/2$.
- Let S be the group membership determined by the classifier,

$$\begin{aligned} f(\mathbf{x}|S = s, \boldsymbol{\theta}) =& \{(1 - \delta_1)\phi(\mathbf{x}|\boldsymbol{\eta}_1, \Sigma) + \delta_1\phi(\mathbf{x}|\boldsymbol{\eta}_2, \Sigma)\}I_{\{1\}}(s) \\ &+ \{\delta_2\phi(\mathbf{x}|\boldsymbol{\eta}_1, \Sigma) + (1 - \delta_2)\phi(\mathbf{x}|\boldsymbol{\eta}_2, \Sigma)\}I_{\{2\}}(s), \end{aligned}$$

where $\boldsymbol{\theta} = (\delta_1, \delta_2, \boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \Sigma)$, $\boldsymbol{\eta}_1 = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_1 + \boldsymbol{\tau}_1)^\top$ and $\boldsymbol{\eta}_2 = (\boldsymbol{\mu}_2, \boldsymbol{\mu}_2 + \boldsymbol{\tau}_2)^\top$.

Moment-Based Method (ϵ and δ are Known)

- If ϵ and δ are known,

$$C\mathrm{E}(\overline{\mathbf{Y}}_D - \overline{\mathbf{Y}}_H) = (1-\epsilon-\delta)(\boldsymbol{\tau}_D - \boldsymbol{\tau}_H) = (1-\epsilon-\delta)\boldsymbol{\Delta}.$$

- Thus, an unbiased estimator of $\boldsymbol{\Delta} = \boldsymbol{\tau}_D - \boldsymbol{\tau}_H$ is

$$\widetilde{\boldsymbol{\Delta}} = \frac{1}{1-\epsilon-\delta} C(\overline{\mathbf{Y}}_D - \overline{\mathbf{Y}}_H).$$

The variance of $\widetilde{\boldsymbol{\Delta}}$ is

$$\begin{aligned} Var(\widetilde{\boldsymbol{\Delta}}) = & \frac{1}{(1-\epsilon-\delta)^2} C\Sigma C^\top \left\{ \frac{1}{n_D} + \frac{1}{n_H} \right\} \\ & + \left\{ \frac{\epsilon(1-\epsilon)}{n_D} + \frac{\delta(1-\delta)}{n_H} \right\} \boldsymbol{\Delta}\boldsymbol{\Delta}^\top. \end{aligned}$$

EM-Based Method (ϵ and δ are Unknown)

- Let Z_{ij} be the true group of the jth subject classified in the ith group.
    - Z_{ij} is missing information.
- Via EM algorithm, the MLE of $\boldsymbol{\theta}$ is

$$\widehat{\boldsymbol{\theta}} = \left(\widehat{\delta}_1, \widehat{\delta}_2, \widehat{\boldsymbol{\eta}}_1, \widehat{\boldsymbol{\eta}}_2, \widehat{\Sigma}\right)$$

- Estimate $\boldsymbol{\Delta}$ by $\widehat{\boldsymbol{\Delta}} = C(\widehat{\boldsymbol{\eta}}_1 - \widehat{\boldsymbol{\eta}}_2)$.
- Apply bootstrap to estimate the covariance matrix of $\widehat{\boldsymbol{\Delta}}$, denoted by S_B.

Hybrid Method

- Hybrid estimator of $\mathbf{\Delta}$ is

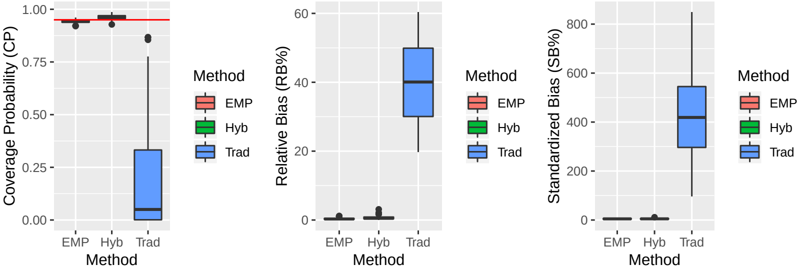
$$\widetilde{\mathbf{\Delta}} = (1 - \widehat{\delta}_1 - \widehat{\delta}_2)^{-1} C \left(\overline{\mathbf{X}}_{1\cdot} - \overline{\mathbf{X}}_{2\cdot}\right)$$

and its variance can be estimated by

$$\widehat{Var(\widetilde{\mathbf{\Delta}})} = \left(1 - \widehat{\delta}_1 - \widehat{\delta}_2\right)^{-2} CSC^{\top}.$$

where $\widehat{\delta}_1$ and $\widehat{\delta}_2$ are EM estimators.

Overall Comparison



(a) Comparison of CP

(b) Comparison of RB%

(c) Comparison of SB%

Figure: Boxplots of CP, RB%, and SB% for all methods.Trad is for the traditional method; Hyb is for the hybrid method; EMP is for the MLE via EM algorithm.

Sample Size Determination

- Hypothesis Test: $H_0 : \mathbf{\Delta} = \mathbf{\Delta}_0$ vs. $H_1 : \mathbf{\Delta} = \mathbf{\Delta}_1$ (s.t. $\mathbf{\Delta}_1 \neq \mathbf{\Delta}_0$).
- For the nominal test size α and power $1 - \beta$, the required sample size of $n = n_D + n_H$, where $n_D/n_H = \pi$ and $0 < \pi < \infty$ can be derived based on test statistic

  $$\widetilde{T} = [C(\overline{\mathbf{Y}}_D - \overline{\mathbf{Y}}_H) - \psi \mathbf{\Delta}_0]^\top \left(CSC^\top \right)^{-1} [C(\overline{\mathbf{Y}}_D - \overline{\mathbf{Y}}_H) - \psi \mathbf{\Delta}_0], \tag{1}$$

  where $C = (-I_p, I_p)$, $S = \frac{1}{n_D}(S_D + \pi S_H)$ and S_D and S_H are the sample covariances of $\mathbf{Y}_{D_i}$ and $\mathbf{Y}_{Hi}$, respectively, and $\psi = 1 - \epsilon - \delta$.

F Approximation

- We have the approximation

$$\widetilde{T} \approx pF \sim pF_{p,f}(n_D \psi^2 (\mathbf{\Delta}_1 - \mathbf{\Delta}_0)^\top \Phi^{-1} (\mathbf{\Delta}_1 - \mathbf{\Delta}_0))$$

  Therefore, to find n_D we need to solve the equation

$$P(T > pF_{p,f_0}(1-\alpha)|H_1) \approx P(F > F_{p,f_0}(1-\alpha)) = 1 - \beta,$$

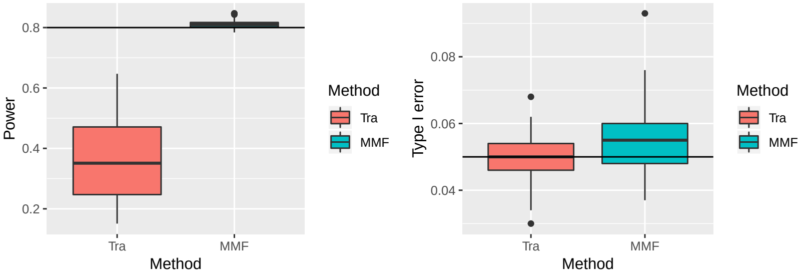
Sample Size

- Due to misclassification rates, the sample size required are larger than the traditional methods.

		ϵ		
δ	Method	0.1	0.2	0.3
0.1	Trad	17	17	17
	F	26	34	47
0.2	Trad	17	17	17
	F	34	48	70
0.3	Trad	17	17	17
	F	47	70	111

Table: Sample size required for parametric setting 1 when $p = 2$. Trad, Traditional method; F, F Approximation

Overall Comparisons



(a) Comparison of power among different methods

(b) Comparison of Type I error among different methods

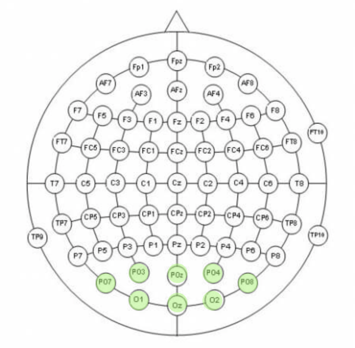
Figure: Boxplots of power and Type I error for all methods. Tra, traditional test that ignores group classification errors; MMF, moment-based test; The sample sizes for MMF are calculated based on F Approximation.

EEG Data: Data Background

- This data was collected to examine Electroencephalograph (EEG) correlates of genetic predisposition to alcohol use disorder.
- 122 subjects: 77 alcohol use disorder (1) and 45 not having alcohol use disorder (2).
- Their baseline brain activities were recorded using EEG.
- After the baseline assessment, a visual stimuli was presented and the brain activities were measured again.

EEG Data: Data Background

- We focus on the activity recorded on EEG electrodes placed at the O1, Oz, O2, PO7, PO3, POz, PO4, and PO8.
- These channels corresponds to the occipital lobe and parietal lobe of the brain.
- They are responsible for visual processing and spatial relationships.



EEG Data: Results and Conclusions

Method	O1	O2	Oz	PO3	PO4	PO7	PO8	POz	p-value
Trad	1.209	0.865	0.205	0.692	0.849	0.797	0.968	0.442	0.660
Hyb	1.514	1.084	0.256	0.867	1.063	0.998	1.212	0.553	0.750
EMP	2.143	1.840	0.637	1.451	1.549	1.368	1.778	0.926	<0.001

Table: Differences in pre and post brain activity ($\mathbf{\Delta}$)

• Trad is traditional estimator that ignores the misclassification probability in diagnostic test;

• Hyb is the hybrid estimator that combines maximum likelihood estimator of ϵ and δ with the moment-based estimator of $\mathbf{\Delta}$;

• EMP is the maximum likelihood estimator of all the parameters via EM algorithm.

Conclusion and Summary

- We should check the diagnostic device's accuracy before applying the traditional method.
- Traditional methods that ignore misclassification errors lead to unacceptably-large bias in estimating treatment effects. We may fail to detect significant differences in treatment effects.
- Sample sizes required from traditional methods are overly optimistic.
- The EM-based methods provide more accurate estimators for misclassification error rates and treatment effects.
- The hybrid method is easy to use and fast to compute.