Model-Robust Inference for Clinical Trials that Improve Precision by Stratified Randomization and Covariate Adjustment

Bingkai Wang[1], Ryoko Susukida[2], Ramin Mojtabai[2], Masoumeh Amin-Esmaeili[2,3], Michael Rosenblum[1]

[1] Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health
[2] Department of Mental Health, Johns Hopkins Bloomberg School of Public Health
[3] Iranian National Center for Addiction Studies (INCAS), Tehran University of Medical Sciences

Open question

In randomized clinical trials, how can we do model-robust inference under stratified randomization?

Open question

In randomized clinical trials, how can we do model-robust inference under stratified randomization?

- Randomized clinical trial (RCT): gold standard for evaluating the efficacy of new treatments.
- Model-robust inference: valid statistical inference even when the assumed model is wrong.
- Stratified randomization: treatment allocation stratified by baseline strata using permuted blocks.

Open question

In randomized clinical trials, how can we do model-robust inference under stratified randomization?

- Randomized clinical trial (RCT): gold standard for evaluating the efficacy of new treatments.
- Model-robust inference: valid statistical inference even when the assumed model is wrong.
- Stratified randomization: treatment allocation stratified by baseline strata using permuted blocks.

Female	1	0	1	0	0	0	1	1	$\cdots$
Male	0	0	1	1	0	1	0	1	$\cdots$

Example

CTN44 (Campbell et al., 2014) is an RCT evaluating internet-delivered treatment for substance abuse.

- **Treatment**: Therapeutic Education System versus Treatment as usual.
- **Outcome**: number of abstinent days (continuous).
- **Baseline variables**: age, sex and urine laboratory result.
- **Stratified randomization**: treatment allocation stratified by abstinence status at baseline (4 strata).

Example

CTN44 (Campbell et al., 2014) is an RCT evaluating internet-delivered treatment for substance abuse.

- **Treatment**: Therapeutic Education System versus Treatment as usual.
- **Outcome**: number of abstinent days (continuous).
- **Baseline variables**: age, sex and urine laboratory result.
- **Stratified randomization**: treatment allocation stratified by abstinence status at baseline (4 strata).

Suppose that ANCOVA is used to estimate the average treatment effect.

- Is it consistent and asymptotically normal?
- What is the asymptotic variance?

Background

- Stratified randomization is different from **simple randomization**, where treatment allocation is determined by independent coin flips and does not involve baseline strata.

Background

- Stratified randomization is different from **simple randomization**, where treatment allocation is determined by independent coin flips and does not involve baseline strata.

- Although commonly used estimators (e.g. ANCOVA) is well studied under simple randomization, little is known about their asymptotics under stratified randomization.

Background

- Stratified randomization is different from **simple randomization**, where treatment allocation is determined by independent coin flips and does not involve baseline strata.

- Although commonly used estimators (e.g. ANCOVA) is well studied under simple randomization, little is known about their asymptotics under stratified randomization.

- The difference between simple randomization and stratified randomization is usually ignored (Kahan and Morris, 2012).

Background

- Stratified randomization is different from **simple randomization**, where treatment allocation is determined by independent coin flips and does not involve baseline strata.
- Although commonly used estimators (e.g. ANCOVA) is well studied under simple randomization, little is known about their asymptotics under stratified randomization.
- The difference between simple randomization and stratified randomization is usually ignored (Kahan and Morris, 2012).
- Stratified randomization is used by 70% of RCT in top medical journals (Lin et al., 2015).

Related work

- Shao et al. (2010); Shao and Yu (2013) proved the validity of the two-sample t-test under the biased-coin design assuming generalized linear model.
- Ma et al. (2015, 2018) assumed a linear model and derived the asymptotic distribution of the test statistic of ATE for the ANCOVA estimator and a class of covariate-adaptive designs.
- **Bugni et al. (2018) established the asymptotic theory of the unadjusted estimator and the ANCOVA estimator (adjusting for strata only) of ATE for a wide range of covariate-adaptive designs.**
- Ye and Shao (2019) derived asymptotics for log-rank and score tests in survival analysis under covariate adaptive randomization.
- Li and Ding (2019) established the asymptotic theory for the ANCOVA estimator under covariate-adaptive randomization in the randomization inference framework.

Our contributions

1. We derive the **model-robust** consistency and asymptotic normality for a wide class of estimators under stratified randomization.
2. We handle various outcome types, repeated-measured outcomes, missing data and covariate adjustment.
3. The asymptotic variance under stratified randomization is **no larger** than under simple randomization.
4. The above results also hold for the biased-coin covariate-adaptive design.

Definition

For participant $i = 1, \ldots, n$, we observe

- Y_i is the outcome variable, which can be continuous, binary or time-to-event,
- A_i is a binary treatment indicator,
- X_i is a vector of baseline variables.

We use the Neyman-Rubin causal model and assume

$$Y_i = A_i Y_i(1) + (1 - A_i) Y_i(0),$$

where $Y_i(1)$ and $Y_i(0)$ are potential outcomes.

Our goal is to estimate a population parameter, for example, the average treatment effect (ATE)

$$\Delta = E[Y_i(1) - Y_i(0)].$$

Definition

Simple randomization

$(A_1, \ldots, A_n)$ are assigned by independent Bernoulli draws with $P(A_i = 1) = \pi, \pi \in (0, 1)$.

Stratified (permuted block) randomization

For each stratum, randomly permuted blocks with fraction π 1's and $1 - \pi$ 0's are used for sequential allocation.

Definition

Simple randomization

$(A_1, \ldots, A_n)$ are assigned by independent Bernoulli draws with $P(A_i = 1) = \pi, \pi \in (0, 1)$.

Stratified (permuted block) randomization

For each stratum, randomly permuted blocks with fraction π 1's and $1 - \pi$ 0's are used for sequential allocation.

Comparison

Under stratified randomization,

1. Within each stratum, exact π fraction of participants are in the treatment group.
2. $(A_1, \ldots, A_n)$ are not longer independent, which leads to the main challenge to derive the asymptotics.

Estimators

M-estimators $\widehat{\Delta}$ (for binary or continuous outcomes)

- the ANCOVA estimator,
- DR-WLS estimator (involving missing outcomes under the missing at random assumption),
- mixed-effects model for repeated measures (MMRM),
- targeted maximum likelihood estimator (TMLE).

Kaplan Meier estimator $\widehat{S}_a(t)$ (for time-to-event outcomes)

- Estimating the survival function $S_a(t) = P(Y_i(a) > t)$ for $t \in [0, \tau]$.

Theorems

Theorem 1

If an M-estimator $\widehat{\Delta}$ has $\sqrt{n}(\widehat{\Delta} - \Delta) \xrightarrow{d} N(0, V_{\mathrm{simple}})$ under simple randomization, then $\sqrt{n}(\widehat{\Delta} - \Delta) \xrightarrow{d} N(0, V_{\mathrm{strat}})$ under stratified randomization with

$$V_{\mathrm{strat}} \leq V_{\mathrm{simple}}.$$

Theorems

Theorem 2

For the Kaplan-Meier estimator $\widehat{S}_a(t)$, if $\left\{\sqrt{n}[\widehat{S}_a(t) - S_a(t)] : t \in [0, \tau]\right\}$ weakly converges to a Gaussian process $\mathcal{GP}(0, V_{simple})$ under simple randomization, then it weakly converges to $\mathcal{GP}(0, V_{strat})$ under stratified randomization with

$$V_{\mathrm{strat}}(t, t) \leq V_{\mathrm{simple}}(t, t).$$

We can use two ways to improve precision/reduce variance of an estimator:

- Stratified randomization.
- Covariate adjustment (FDA, 2019, 2020).

How much variance reduction can be achieved?

Data example 1: variance reduction due to stratified randomization

- Trial: CTN44
- Time-to-event outcome: time to abstinence
- Group: Therapeutic Education System (treatment)
- Estimator: the Kaplan-Meier estimator

Visit	1	2	3	4	5	6
Survival probability	0.58	0.53	0.47	0.40	0.39	0.33
Proportional variance reduction ($1 - V_{strat}/V_{simple}$)	11%	12%	11%	9%	7%	4%

Data example 2: variance reduction due to covariate adjustment

CTN03, CTN30 and CTN44 are RCTs of treatment of substance use disorder using stratified randomization.

- CTN03 has binary outcomes and CTN30 and CTN44 have continuous outcomes, all being measures of treatment success.
- Each study has $\sim$5 baseline variables.

Study	Number of Strata	Unadjusted estimator (95% CI)	Adjusted estimator (95% CI)	Proportional variance reduction
CTN03	3	-0.11(-0.21, -0.01)	-0.10(-0.19, -0.02)	**35%**
CTN30	4	0.02(-0.02, 0.05)	0.01(-0.02, 0.04)	**17%**
CTN44	4	-0.09(-0.14, -0.03)	-0.09(-0.14, -0.03)	**2%**

Practical implications

1. When using stratified randomization, doing statistical inference with the correct variance (based on V_{strat} instead of V_{simple}) can avoid being conservative.

2. Adjusting for a set of preplanned baseline variables may lead to substantial variance reduction.

Thank you!

The slides are available at
https://bingkaiwang.com.

The paper is available at
https://arxiv.org/abs/1910.13954.

The R code is available at
https://github.com/BingkaiWang/covariate-adaptive.

Acknowledgement

This project was supported by a research award from Arnold Ventures. The content is solely the responsibility of the authors and does not necessarily represent the official views of Arnold Ventures. The information reported here results from secondary analyses of data from clinical trials conducted by the National Institute on Drug Abuse (NIDA). Specifically, data from NIDA–CTN-0003 (Suboxone (Buprenorphine/Naloxone) Taper: A Comparison of Two Schedules), NIDA-CTN-0030 (Prescription Opioid Addiction Treatment Study) and NIDA-CTN-0044 (Web-delivery of Evidence-Based, Psychosocial Treatment for Substance Use Disorders) were included. NIDA databases and information are available at (https://datashare.nida.nih.gov).

Bugni, F. A., I. A. Canay, and A. M. Shaikh (2018). Inference under covariate-adaptive randomization. *Journal of the American Statistical Association 113*(524), 1784–1796.

Campbell, A. N., E. V. Nunes, A. G. Matthews, M. Stitzer, G. M. Miele, D. Polsky, E. Turrigiano, S. Walters, E. A. McClure, T. L. Kyle, A. Wahle, P. Van Veldhuisen, B. Goldman, D. Babcock, P. Q. Stabile, T. Winhusen, and U. E. Ghitza (2014). Internet-delivered treatment for substance abuse: A multisite randomized controlled trial. *American Journal of Psychiatry 171*(6), 683–690.

FDA (2019). Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biologics with Continuous Outcomes. Draft Guidance for Industry. https://www.fda.gov/media/123801/download.

FDA (2020). COVID-19: Developing Drugs and Biological Products for Treatment or Prevention. *U.S. Food and Drug Administration: CDER*.

Kahan, B. C. and T. P. Morris (2012). Improper analysis of trials randomised using stratified blocks or minimisation. *Statistics in Medicine 31*(4), 328–340.

Li, X. and P. Ding (2019). Rerandomization and regression adjustment. *arXiv https://arxiv.org/abs/1906.11291*.

Lin, Y., M. Zhu, and Z. Su (2015). The pursuit of balance: An overview of covariate-adaptive randomization techniques in clinical trials. *Contemporary Clinical Trials 45*, 21 – 25. 10th Anniversary Special Issue.

Ma, W., F. Hu, and L. Zhang (2015). Testing hypotheses of covariate-adaptive randomized clinical trials. *Journal of the American Statistical Association 110*(510), 669–680.

Ma, W., Y. Qin, Y. Li, and F. Hu (2018). Statistical inference of covariate-adjusted randomized experiments. *arXiv https://arxiv.org/abs/1807.09678*.

Shao, J. and X. Yu (2013). Validity of tests under covariate-adaptive biased coin randomization and generalized linear models. *Biometrics 69*(4), 960–969.

Shao, J., X. Yu, and B. Zhong (2010). A theory for testing hypotheses under covariate-adaptive randomization. *Biometrika 97*(2), 347–360.

Ye, T. and J. Shao (2019). Robust tests for treatment effect in survival analysis under covariate-adaptive randomization. *arXiv https://arxiv.org/abs/1811.07232*.