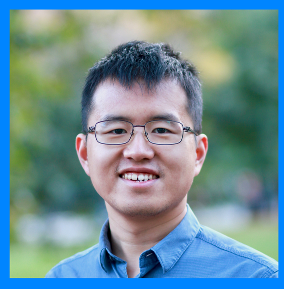
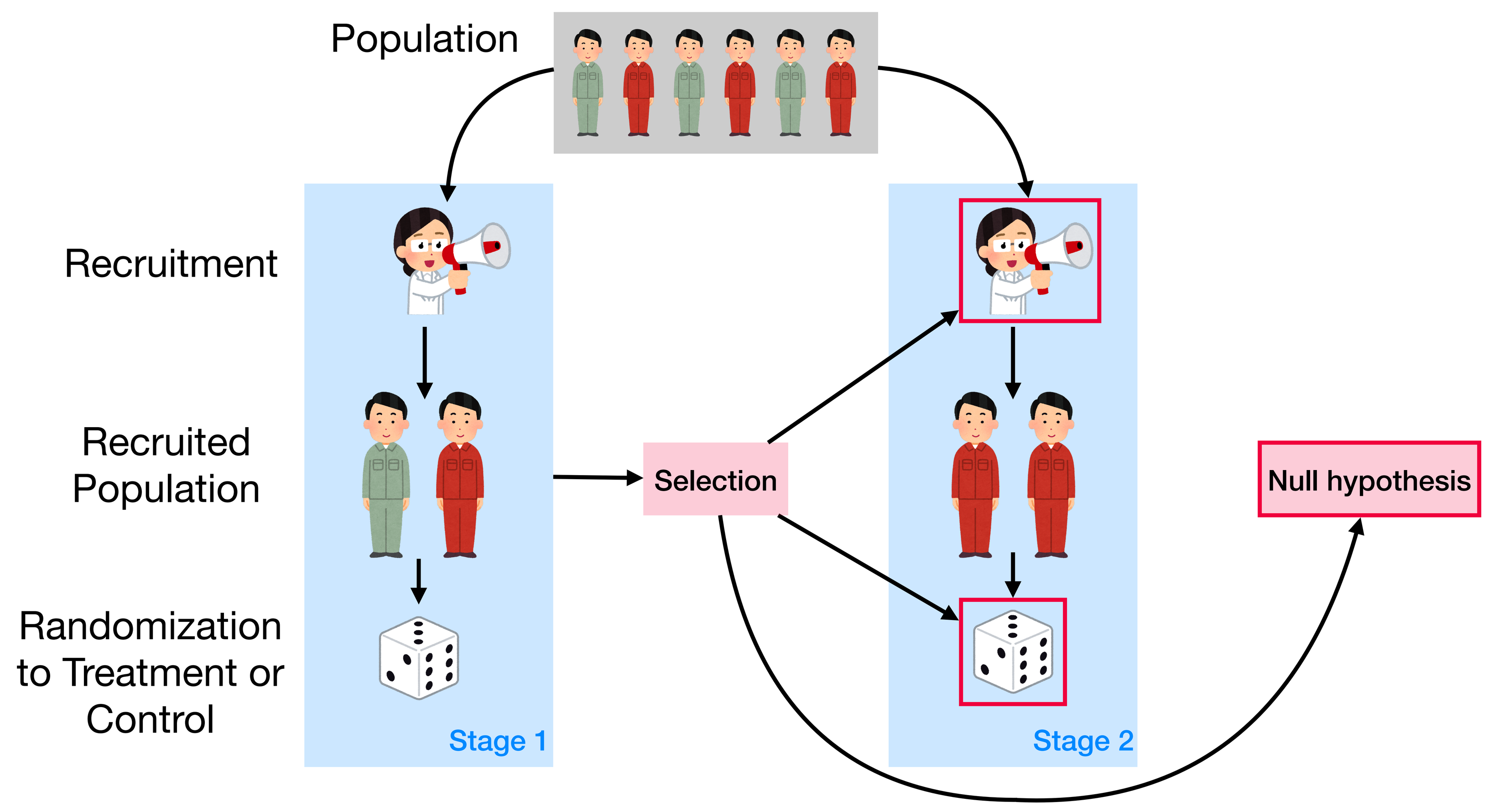
Selective Randomization Inference for Adaptive Studies

Tobias Freidling[1] Qingyuan Zhao[1] Zijun Gao[2]

[1]University of Cambridge [2]University of Southern California



Analysing Adaptive Studies



Adaptive Studies

- Characteristics: Recruitment, treatment assignment and null hypothesis can depend on data from previous stages
- Benefits: reacting to external circumstances, more ethical treatment allocation, saving time and money [1]

Data Analysis

- Difficulty: data informs design and null hypothesis → risk of double dipping
- Existing methods: design-specific, strong assumptions
- Our approach: randomization inference → no modelling assumptions or i.i.d. data needed

Selective Randomization P-value

Insight: only use randomness of Z as its distribution is known

Testing the null hypothesis

$$H_0: \quad Y_i(1) - Y_i(0) = 0 \quad \text{for all } i \in R \text{ (or a subset)}$$

with the statistic T. ($Z^* \overset{D}{=} Z$ and $Z^* \perp\!\!\!\perp Z \mid W$)

- Usual randomization p-value[2]: invalid due to double dipping

$$P^*(T(Z^*, W) \leq T(Z, W) \mid Z, W)$$

- Data splitting[3] / 2nd stage randomization p-value: loses power

$$P^*(T(Z^*, W) \leq T(Z, W) \mid Z, W, Z_1^* = Z_1)$$

- **Selective randomization p-value:** valid & more powerful [4, 5]

$$p(Z) := P^*(T(Z^*, W) \leq T(Z, W) \mid Z, W, S(Z_1^*) = S(Z_1))$$

Inference and Computation

Inference for a homogeneous treatment effect $\beta = Y_i(1) - Y_i(0)$, where $i \in R$ (or a subset):

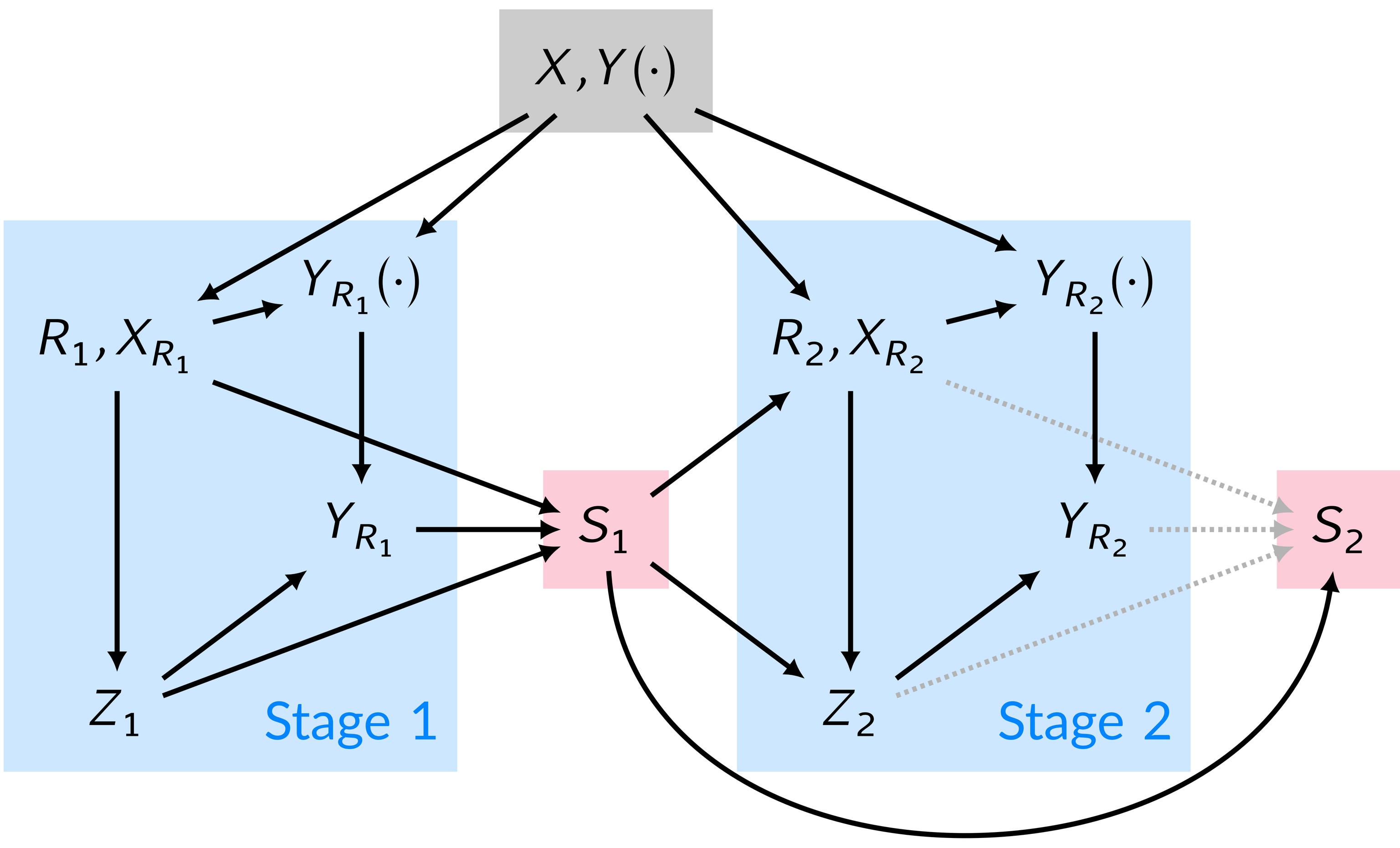
- $(1-\alpha)$ confidence interval: inversion of tests $\{\beta \colon p_\beta(Z) \geq \alpha\}$
- Estimation: $\hat{\beta} = \beta$ such that $p_\beta(Z) = 0.5$

Computation of p-value via Monte Carlo approximation

$$\frac{\sum_{j=1}^m \mathbf{1}\{T(z_j^*, W) \leq T(Z, W)\} \cdot P^*(Z^* = z_j^* \mid W)}{\sum_{j=1}^m P^*(Z^* = z_j^* \mid W)},$$

where sample $(z_j^*)_{j=1}^m$ is generated via rejection sampling or MCMC

DAG and Notation



- Covariates X and potential outcomes $Y(\cdot)$ of population
- Recruitment: R_1, R_2
- Treatment assignment: Z_1, Z_2
- Observed outcomes: $Y_{R_i} = Y_{R_i}(Z_i)$
- Selective choice: S_1, S_2
- Short-hand: $W = (R, X_R, Y_R(\cdot))$

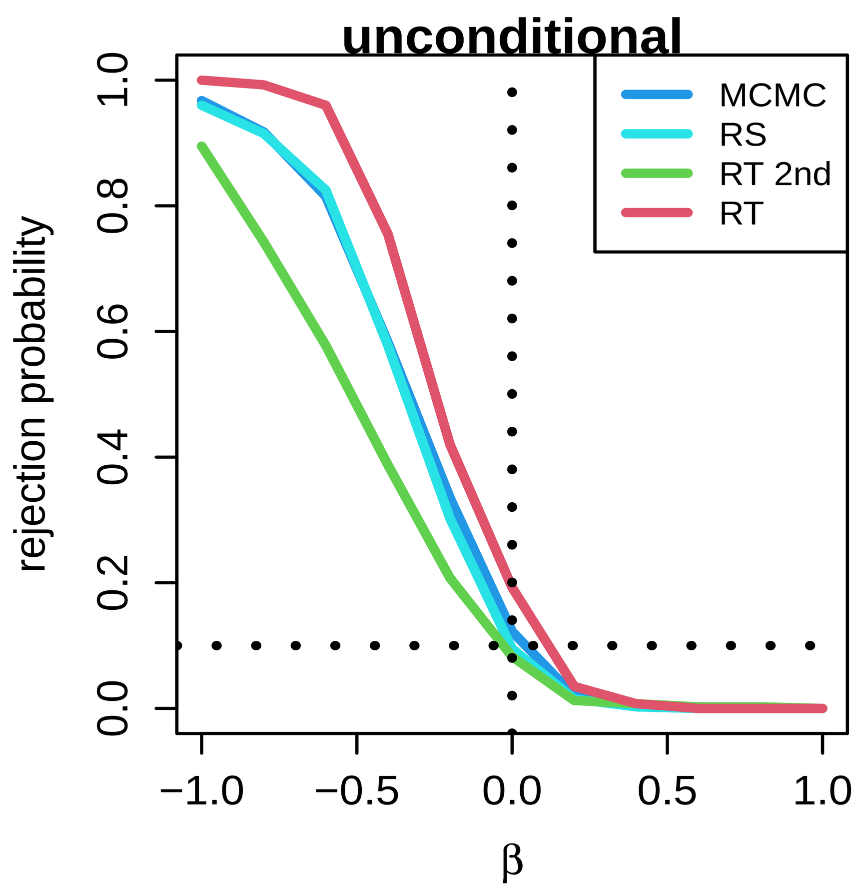
Simulation Study

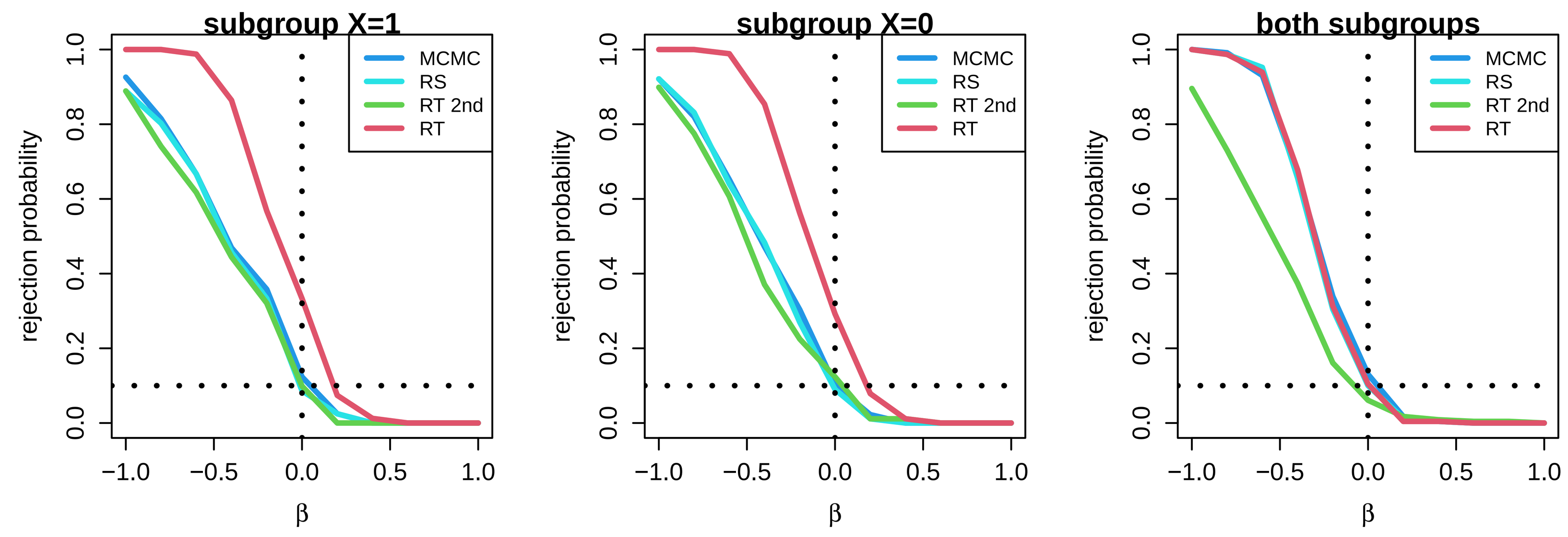
- 2 stages, 2 treatments $Z_i \in \{0, 1\}$, 2 groups $X_i \in \{0, 1\}$
- Potential outcomes: $Y_i(0) = Y_i(1) \sim N(0, 1)$ i.i.d.
- First stage: 50 patients
- Δ = standardized difference in SATEs between groups
- Selection variable and recruitment in second stage:

$$S = \begin{cases} 0, & \Delta < \Phi^{-1}(0.2), & \text{recruit 25 from group } X_i = 0, \\ 1, & \Delta > \Phi^{-1}(0.8), & \text{recruit 25 from group } X_i = 1, \\ 2, & \text{otherwise}, & \text{recruit 13/12.} \end{cases}$$

Power analysis:

- Type-I error control overall and in subgroups
- More powerful than data splitting
- Similar approximations for rejection sampling and MCMC





References

[1] Philip Pallmann, Alun W. Bedding, Babak Choodari-Oskooei, and collab. Adaptive designs in clinical trials: why use them, and how to run and report them. *BMC Medicine*, 16(1):29, 2018.

[2] R. A. Fisher. *The design of experiments*. Oliver & Boyd, Edinburgh, 1935.

[3] D. R. Cox. A note on data-splitting for the evaluation of significance levels. *Biometrika*, 62(2):441–444, 1975.

[4] Yao Zhang and Qingyuan Zhao. What is a Randomization Test? *Journal of the American Statistical Association*, 0(0):1–15, 2023.

[5] William Fithian, Dennis Sun, and Jonathan Taylor. Optimal Inference After Model Selection. *arXiv: 1410.2597*, 2017.