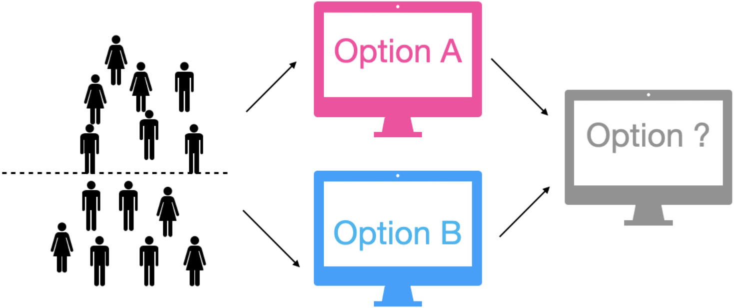
Statistical Designs for Network A/B Tests

Qiong Zhang

School of Mathematical and Statistical Sciences, Clemson University

Feb 23, 2023, Virginia Tech

What is A/B testing?



A/B test is popular in IT companies to compare difference versions of web designs, services, and recommendation algorithms, etc.

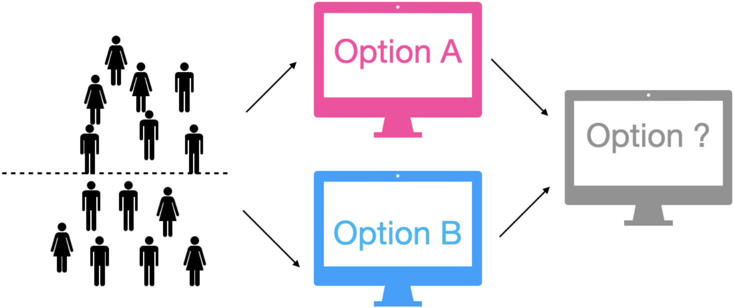
Example: Small Change, Huge Impact

Control: Amazon placed a credit-card offer on the home page in 2004.
Treatment: Move the offer to the shopping cart page, i.e.,

Figure taken from Kohavi et. al., 2014

Impact: The controlled experiment demonstrated that this simple change increased Amazon's annual profit by tens of millions of dollars.

Opportunities in Statistical Designs



A Fundamental Question in Experimental Design

Given n users, how to allocate A and B options to them in order to produce the most accurate estimator of the treatment effect.

Design for A/B Tests

For the i-th user

$$y_i = \alpha + x_i\beta + \delta_i, \quad \text{with} \quad i = 1, \ldots, n$$

with

- Response y_i: a continuous response, e.g., active time, click through rate over a time period, etc.
- Design: $x_i \in \{-1, 1\}$ as the treatment settings.
- α is the intercept and β is the treatment effect.
- δ_i's are iid random variables with zero mean.

Aim: Minimize the variance of the least squared estimator of the treatment effect β:

$$\operatorname{var}\left(\hat{\beta}\right) \propto \frac{1}{n^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \Rightarrow \text{Balanced design}$$

Design for A/B Tests

For the i-th user

$$y_i = \alpha + x_i\beta + \delta_i, \quad \text{with} \quad i = 1, \ldots, n$$

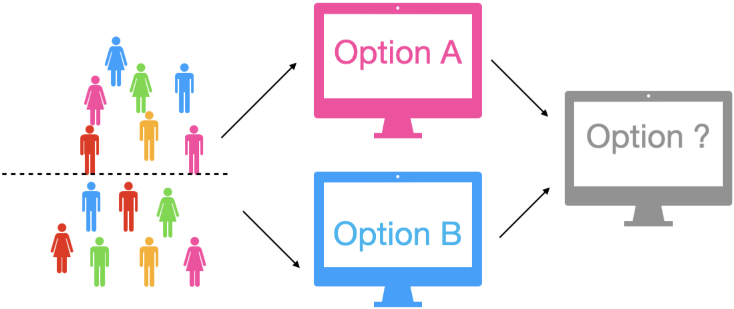
with

- Response y_i: a continuous response, e.g., active time, click through rate over a time period, etc.
- Design: $x_i \in \{-1, 1\}$ as the treatment settings.
- α is the intercept and β is the treatment effect.
- δ_i's are iid random variables with zero mean.

Aim: Minimize the variance of the least squared estimator of the treatment effect β:

$$\text{var}\left(\hat{\beta}\right) \propto \frac{1}{n^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \Rightarrow \text{Balanced design}$$

A/B Tests with User Information



The users' information such as demographics and past online behavior might be available, denoted by $z_{i1}, \ldots, z_{ip}$.

Designs for A/B Tests with Covariates

Incorporate the covariates as additive factors:

$$y_i = \beta x_i + \boldsymbol{z}_i^\top \alpha + \delta_i,$$

with $\boldsymbol{z}_i = (1, z_{i,1}, \ldots, z_{i,p})^\top$.

Aim: Minimize the variance of the least squared estimator of the treatment effect β:

$$\mathrm{var}\left(\hat{\beta}\right) \propto \frac{1}{n - \boldsymbol{x}^\top Z(Z^\top Z)^{-1} Z^\top \boldsymbol{x}},$$

where $Z = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n)^\top$ and $\boldsymbol{x} = (x_1, \ldots, x_n)^\top$.

Designs for A/B Tests with Covariates

The optimal design $\boldsymbol{x}$ (e.g., Bhat et al, 2020) can be obtained by

$$\begin{aligned} &\min \boldsymbol{x}^\top Z(Z^\top Z)^{-1} Z^\top \boldsymbol{x} \\ &\text{s.t. } \boldsymbol{x} \in \{-1, 1\}^n. \end{aligned}$$

Alternatively,

- The objective $x^\top Z(Z^\top Z)^{-1} Z^\top x$ is the Mahalanobis distance in Morgan and Rubin (2012)
- A rerandomization design can be obtained by utilizing the asymptotic χ^2 distribution of this objective with respect to the randomness of x.

Designs for A/B Tests with Covariates

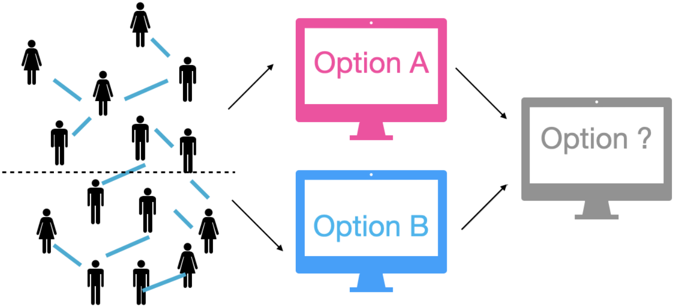
The optimal design $\boldsymbol{x}$ (e.g., Bhat et al, 2020) can be obtained by

$$\min \boldsymbol{x}^\top Z(Z^\top Z)^{-1}Z^\top \boldsymbol{x}$$
$$\text{s.t. } \boldsymbol{x} \in \{-1, 1\}^n.$$

Alternatively,

- The objective $\boldsymbol{x}^\top Z(Z^\top Z)^{-1}Z^\top \boldsymbol{x}$ is the Mahalanobis distance in Morgan and Rubin (2012)
- A rerandomization design can be obtained by utilizing the asymptotic χ^2 distribution of this objective with respect to the randomness of $\boldsymbol{x}$.

A/B Tests with Network



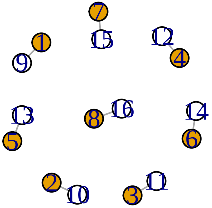
The users are connected in a social network, represented by an $n \times n$ adjacency matrix: $W = \{w_{ij}\}$ with $w_{ii} = 0$ and

$$w_{ij} = \begin{cases} 1, & \text{if user } i \text{ and user } j \text{ are connected} \\ 0, & \text{otherwise,} \end{cases} \quad \text{for} \quad i \neq j.$$

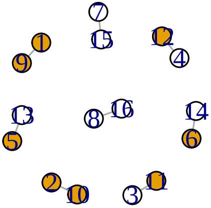
Design for A/B Tests with Network

Question: How should we allocate the users to two treatment groups according to their network connection?

Design Allocation I



Design Allocation II

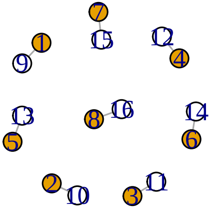


- **Design Allocation I:** $\min_x x^\top W x$ with $\sum x_i = 0$ and $x \in \{-1, 1\}^n$
- **Design Allocation II:** $x^\top W x = 0$ with $\sum x_i = 0$ and $x \in \{-1, 1\}^n$

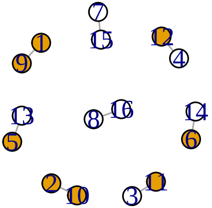
Design for A/B Tests with Network

Question: How should we allocate the users to two treatment groups according to their network connection?

Design Allocation I



Design Allocation II



- **Design Allocation I:** $\min_{\boldsymbol{x}} \boldsymbol{x}^\top W \boldsymbol{x}$ with $\sum x_i = 0$ and $\boldsymbol{x} \in \{-1, 1\}^n$
- **Design Allocation II:** $\boldsymbol{x}^\top W \boldsymbol{x} = 0$ with $\sum x_i = 0$ and $\boldsymbol{x} \in \{-1, 1\}^n$

Design for A/B Tests with Network

Question: How should we allocate the users to two treatment groups according to their network connection?

Design Allocation I **Design Allocation II**

- **Design Allocation I:** $\min_{\boldsymbol{x}} \boldsymbol{x}^\top W \boldsymbol{x}$ with $\sum x_i = 0$ and $\boldsymbol{x} \in \{-1, 1\}^n$
- **Design Allocation II:** $\boldsymbol{x}^\top W \boldsymbol{x} = 0$ with $\sum x_i = 0$ and $\boldsymbol{x} \in \{-1, 1\}^n$

Assumptions on the Network Dependence

Consider the model

$$y_i = \alpha + \beta x_i + \delta_i,$$

where δ_is are not iid random variables.

Instead, we model the network dependence based on

$$\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)^\top.$$

Popular Assumptions:

- Network Correlated Responses, e.g., conditional autoregressive model (CAR)
- Network Interference

Assumptions on the Network Dependence

Consider the model

$$y_i = \alpha + \beta x_i + \delta_i,$$

where δ_is are not iid random variables.

Instead, we model the network dependence based on

$$\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)^\top.$$

Popular Assumptions:

- Network Correlated Responses, e.g., conditional autoregressive model (CAR)
- Network Interference

Network Correlated Responses with CAR Model

Consider the model

$$y_i = \alpha + \beta x_i + \delta_i,$$

where

$$\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)^\top \sim \mathcal{MVN}_n(0, \sigma^2 (D - \rho W)^{-1}),$$

where $D = \mathrm{diag}(d_1, \ldots, d_n)$ with $d_i = \sum_{j=1}^n w_{ij} > 0$ and $0 < \rho < 1$ is the correlation parameter.

The optimal design that minimizes the variance of estimated β (Pokhiko et al, 2019) requires that

$$\min_x \boldsymbol{x}^\top W \boldsymbol{x},$$

and

$$\sum d_i x_i = 0 \quad \text{and} \quad \sum x_i = 0$$

Network Correlated Responses with CAR Model

Consider the model

$$y_i = \alpha + \beta x_i + \delta_i,$$

where

$$\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)^\top \sim \mathcal{MVN}_n(0, \sigma^2 (D - \rho W)^{-1}),$$

where $D = \mathrm{diag}(d_1, \ldots, d_n)$ with $d_i = \sum_{j=1}^n w_{ij} > 0$ and $0 < \rho < 1$ is the correlation parameter.

The optimal design that minimizes the variance of estimated β (Pokhiko et al, 2019) requires that

$$\min_{\boldsymbol{x}} \boldsymbol{x}^\top W \boldsymbol{x},$$

and

$$\sum d_i x_i = 0 \quad \text{and} \quad \sum x_i = 0$$

Network Interference

Consider the model

$$y_i = \alpha + \beta x_i + \delta_i,$$

where

$$\delta_i = \sum_{j=1}^{n} w_{ij}\tau + \left(\sum_{j=1}^{n} w_{ij}x_j\right)\gamma + \varepsilon_i,$$

with ε_i be iid random variables with mean zero.

The optimal design that minimizes the variance of estimated β (e.g., Zhang, 2023+) requires that

$$\boldsymbol{x}^\top W\boldsymbol{x} = 0, \quad \sum x_i = 0, \quad \sum d_i x_i = 0$$

where $d_i = \sum_{j=1}^{n} w_{ij}$.

Network Interference

Consider the model

$$y_i = \alpha + \beta x_i + \delta_i,$$

where

$$\delta_i = \sum_{j=1}^{n} w_{ij}\tau + \left(\sum_{j=1}^{n} w_{ij}x_j\right)\gamma + \varepsilon_i,$$

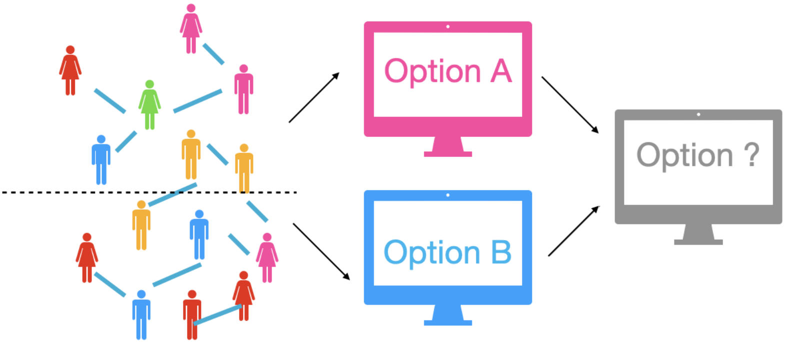
with ε_i be iid random variables with mean zero.

The optimal design that minimizes the variance of estimated β (e.g., Zhang, 2023+) requires that

$$\boldsymbol{x}^\top W\boldsymbol{x} = 0, \quad \sum x_i = 0, \quad \sum d_i x_i = 0$$

where $d_i = \sum_{j=1}^{n} w_{ij}$.

Focus in This Talk (Zhang and Kang, 2022)



- Covariates and network connection information are given.
- Assume that the users' responses are more likely to be similar:
  - if their covariates share common features
  - if they are connected in the social network, i.e., network-correlated responses

Optimal Design with Covariates under CAR Model

Recall

$$y_i = \beta x_i + \boldsymbol{z}_i^\top \alpha + \delta_i,$$

with

$$\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)^\top \sim \mathcal{MVN}_n(0, \sigma^2 (D - \rho W)^{-1}).$$

Aim: minimize the variance of the estimated treatment effect β under CAR model assumption is

$$\mathrm{var}(\hat{\beta}) \propto \frac{1}{\boldsymbol{x}^\top K \boldsymbol{x}}$$

where K is an $n \times n$ matrix depending on ρ

$$K = (D - \rho W) - (D - \rho W) Z \left[Z^\top (D - \rho W) Z \right]^{-1} Z^\top (D - \rho W),$$

Optimal Design with Covariates under CAR Model

Recall

$$y_i = \beta x_i + \boldsymbol{z}_i^\top \alpha + \delta_i,$$

with

$$\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)^\top \sim \mathcal{MVN}_n(0, \sigma^2 (D - \rho W)^{-1}).$$

Aim: minimize the variance of the estimated treatment effect β under CAR model assumption is

$$\mathrm{var}(\hat{\beta}) \propto \frac{1}{\boldsymbol{x}^\top K \boldsymbol{x}}$$

where K is an $n \times n$ matrix depending on ρ

$$K = (D - \rho W) - (D - \rho W) Z \left[Z^\top (D - \rho W) Z \right]^{-1} Z^\top (D - \rho W),$$

Optimal Design with Covariates under CAR Model

Denote

$$T(\boldsymbol{x}, \rho) = \boldsymbol{x}^\top K \boldsymbol{x}$$

If ρ is given, the optimal design is obtained by

$$\max T(\boldsymbol{x}, \rho),$$

$$\text{s.t. } -1 \leq \sum_{i=1}^{n} x_i \leq 1, \text{ and } \boldsymbol{x} \in \{-1, 1\}^n,$$

Two Issues

- The correlation parameter ρ is unknown.
- K is a positive definite matrix, thus maximizing the objective $\boldsymbol{x}^\top K \boldsymbol{x}$ is a challenging problem to solve compared with minimizing $\boldsymbol{x}^\top K \boldsymbol{x}$.

Optimal Design with Covariates under CAR Model

Denote

$$T(\boldsymbol{x}, \rho) = \boldsymbol{x}^\top K \boldsymbol{x}$$

If ρ is given, the optimal design is obtained by

$$\max \; T(\boldsymbol{x}, \rho),$$

$$\text{s.t. } -1 \leq \sum_{i=1}^{n} x_i \leq 1, \text{ and } \boldsymbol{x} \in \{-1, 1\}^n,$$

Two Issues

- The correlation parameter ρ is unknown.
- K is a positive definite matrix, thus maximizing the objective $\boldsymbol{x}^\top K \boldsymbol{x}$ is a challenging problem to solve compared with minimizing $\boldsymbol{x}^\top K \boldsymbol{x}$.

Locally Optimal Design

Theorem

For $\rho \in (0,1)$, and any given design $\boldsymbol{x}$, the design criterion $T(\boldsymbol{x}, \rho)$ is a concave function with respect to ρ.

Corollary

Given a prior distribution of ρ, $p(\rho)$, for $\rho \in (0,1)$, an upper bound for $\mathbb{E}\left[T(\boldsymbol{x}, \rho)\right]$ is $\mathbb{E}\left[T(\boldsymbol{x}, \rho)\right] \leq T(\boldsymbol{x}, \rho_0)$, where $\rho_0 := \mathbb{E}(\rho)$ is the population mean of ρ based on $p(\rho)$.

We define the locally optimal design by solving the original problem with a plug-in value ρ_0:

$$\max T(\boldsymbol{x}, \rho_0)$$

$$\text{s.t. } -1 \leq \sum_{i=1}^{n} x_i \leq 1, \text{ and } \boldsymbol{x} \in \{-1, 1\}^n.$$

Locally Optimal Design

Theorem

For $\rho \in (0, 1)$, and any given design $\boldsymbol{x}$, the design criterion $T(\boldsymbol{x}, \rho)$ is a concave function with respect to ρ.

Corollary

Given a prior distribution of ρ, $p(\rho)$, for $\rho \in (0, 1)$, an upper bound for $\mathbb{E}\left[T(\boldsymbol{x}, \rho)\right]$ is $\mathbb{E}\left[T(\boldsymbol{x}, \rho)\right] \leq T(\boldsymbol{x}, \rho_0)$, where $\rho_0 := \mathbb{E}(\rho)$ is the population mean of ρ based on $p(\rho)$.

We define the locally optimal design by solving the original problem with a plug-in value ρ_0:

$$\max T(\boldsymbol{x}, \rho_0)$$

$$\text{s.t. } -1 \leq \sum_{i=1}^{n} x_i \leq 1, \text{ and } \boldsymbol{x} \in \{-1, 1\}^n.$$

Reformulation

Notice that

$$T(\boldsymbol{x},\rho)=\boldsymbol{x}^\top K\boldsymbol{x}=\sum_{ij} w_{ij}-\rho T_1(\boldsymbol{x})-T_2(\boldsymbol{x},\rho)$$

with

$$T_1(\boldsymbol{x})=\boldsymbol{x}^\top W\boldsymbol{x}$$
$$T_2(\boldsymbol{x},\rho)=\boldsymbol{x}^\top(D-\rho W)Z\left[Z^\top(D-\rho W)Z\right]^{-1}Z^\top(D-\rho W)\boldsymbol{x}$$

where $(D-\rho W)Z\left[Z^\top(D-\rho W)Z\right]^{-1}Z^\top(D-\rho W)$ is a positive definite matrix.

Reformulation

Given the value of $\rho = \rho_0$, the original optimization problem can be reformulated by

$$\begin{aligned} &\min\ T_2(\boldsymbol{x}, \rho_0) \\ &\text{s.t. } T_1(\boldsymbol{x}) \le q, \\ &\quad -1 \le \sum_{i=1}^{n} x_i \le 1, \\ &\quad \boldsymbol{x} \in \{-1, 1\}^n, \end{aligned}$$

where

$$T_1(\boldsymbol{x}) = \boldsymbol{x}^\top W \boldsymbol{x}$$

$$T_2(\boldsymbol{x}, \rho) = \boldsymbol{x}^\top (D - \rho W) Z \left[Z^\top (D - \rho W) Z \right]^{-1} Z^\top (D - \rho W) \boldsymbol{x}$$

Robustness with respect to ρ_0

Consider that $x_1, \ldots, x_n$ in $\boldsymbol{x}$ are independent and identically distributed random variables from the discrete distribution with

$$\Pr(x_i = 1) = \Pr(x_i = -1) = 0.5.$$

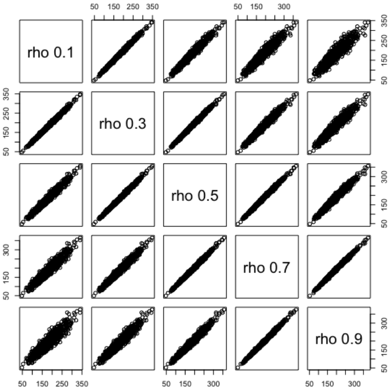
For any two symmetric $n \times n$ fixed matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, we have that

$$\operatorname{cor}(\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}^\top \boldsymbol{B}\boldsymbol{x}) = \frac{\sum_{i<j} a_{ij} b_{ij}}{\sqrt{\sum_{i<j} a_{ij}^2}\sqrt{\sum_{i<j} b_{ij}^2}},$$

where a_{ij} and b_{ij} are the (i,j)-th entries of matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ respectively.

Robustness with respect to ρ_0

Draw 1000 random samples of $\boldsymbol{x}$ and compute the correlation parameters between $T_2(\boldsymbol{x}, \rho)$ and $T_2(\boldsymbol{x}, \rho')$



For 100 users each with 20 covariates, the correlation values are all above 0.9. *We specify that* $\rho_0 = 0.5$.

Upper Bound of $T_1(\boldsymbol{x})$

Consider that $x_1, \ldots, x_n$ in $\boldsymbol{x}$ are independent and identically distributed random variables from the discrete distribution with

$$\Pr(x_i = 1) = \Pr(x_i = -1) = 0.5.$$

Under minor assumptions to the sparsity of the adjacency matrix, as $n \to \infty$,

$$\frac{T_1(\boldsymbol{x})}{\mathrm{sd}(T_1(\boldsymbol{x}))} \xrightarrow{d} N(0,1),$$

where $\xrightarrow{d}$ represents convergence in distribution.

The reformulated constraint can be changed to

$$T_1(\boldsymbol{x}) \leq \mathrm{sd}(T_1(\boldsymbol{x}))q_\alpha,$$

where q_α is the α-th quantile of $N(0,1)$ and $\mathrm{sd}(T_1(\boldsymbol{x})) = \sqrt{2\sum_{ij} w_{ij}}$

Upper Bound of $T_1(\boldsymbol{x})$

Consider that $x_1, \ldots, x_n$ in $\boldsymbol{x}$ are independent and identically distributed random variables from the discrete distribution with

$$\Pr(x_i = 1) = \Pr(x_i = -1) = 0.5.$$

Under minor assumptions to the sparsity of the adjacency matrix, as $n \to \infty$,

$$\frac{T_1(\boldsymbol{x})}{\mathrm{sd}(T_1(\boldsymbol{x}))} \xrightarrow{d} N(0,1),$$

where $\xrightarrow{d}$ represents convergence in distribution.

The reformulated constraint can be changed to

$$T_1(\boldsymbol{x}) \leq \mathrm{sd}(T_1(\boldsymbol{x}))q_\alpha,$$

where q_α is the α-th quantile of $N(0,1)$ and $\mathrm{sd}(T_1(\boldsymbol{x})) = \sqrt{2\sum_{ij} w_{ij}}$

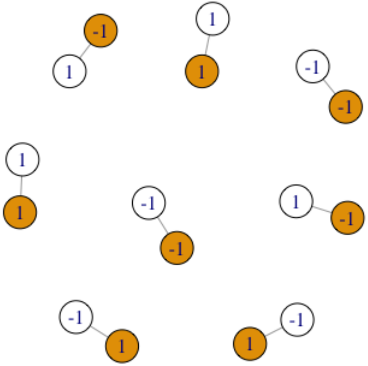
A Hybrid Solution Approach

Set $\rho_0 = 0.5$, the optimal design is obtained by

$$\begin{aligned} \min T_2(\boldsymbol{x}, \rho_0) = \boldsymbol{x}^\top (D - \rho_0 W) Z \left[Z^\top (D - \rho_0 W) Z \right]^{-1} Z^\top (D - \rho_0 W) \boldsymbol{x} \\ \text{s.t. } T_1(\boldsymbol{x}) \leq \mathrm{sd}(T_1(\boldsymbol{x})) q_\alpha, \\ -1 \leq \sum_{i=1}^{n} x_i \leq 1, \\ \boldsymbol{x} \in \{-1, 1\}^n. \end{aligned}$$

Illustration

- Design Allocation $\boldsymbol{x}$: color
- 1-d Covariate z: 1 or -1
- $T_1(\boldsymbol{x}) = \boldsymbol{x}^\top W \boldsymbol{x}$: network balance
- $T_2(\boldsymbol{x}, \rho)$: network adjusted covariates balance



Numerical Study: Evaluation

Consider a random balanced design $\boldsymbol{x}$ with

$$\Pr(x_i = 1) = \Pr(x_i = -1) = 0.5 \quad \text{and} \quad -1 \le \sum_{i=1}^{n} x_i \le 1.$$

The expected precision of the random balanced design is

$$\mathbb{E}_{\boldsymbol{x}} \left(\sigma^{-2} \boldsymbol{x}^\top K \boldsymbol{x} \right) = \sigma^{-2} \mathrm{tr}(K\boldsymbol{C}),$$

where $\boldsymbol{C}$ is an $n \times n$ matrix with all of the diagonal entries equal to 1 and all of the off-diagonal entries equal to a fixed constant $-1/n$ or $-1/(n-1)$.

A given design x_0 is evaluated by Percentage of Improvement in Precision

$$\frac{\sigma^{-2} \boldsymbol{x}_0^\top K \boldsymbol{x}_0 - \mathbb{E}_{\boldsymbol{x}} \left(\sigma^{-2} \boldsymbol{x}^\top K \boldsymbol{x} \right)}{\sigma^{-2} \boldsymbol{x}_0^\top K \boldsymbol{x}_0} = 1 - \frac{\mathrm{tr}(K\boldsymbol{C})}{\boldsymbol{x}_0^\top K \boldsymbol{x}_0}.$$

Numerical Study: Evaluation

Consider a random balanced design $\boldsymbol{x}$ with

$$\Pr(x_i = 1) = \Pr(x_i = -1) = 0.5 \quad \text{and} \quad -1 \leq \sum_{i=1}^{n} x_i \leq 1.$$

The expected precision of the random balanced design is

$$\mathbb{E}_{\boldsymbol{x}} \left(\sigma^{-2} \boldsymbol{x}^\top K \boldsymbol{x} \right) = \sigma^{-2} \mathrm{tr}(K\boldsymbol{C}),$$

where $\boldsymbol{C}$ is an $n \times n$ matrix with all of the diagonal entries equal to 1 and all of the off-diagonal entries equal to a fixed constant $-1/n$ or $-1/(n-1)$.

A given design $\boldsymbol{x}_0$ is evaluated by Percentage of Improvement in Precision

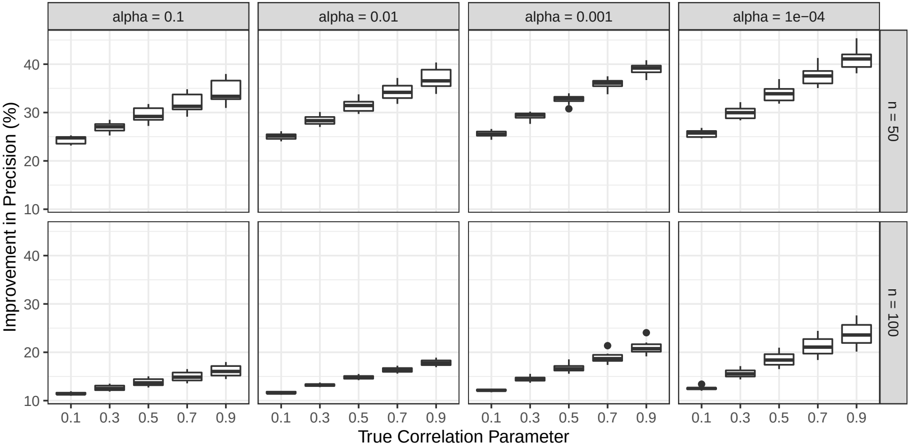
$$\frac{\sigma^{-2} \boldsymbol{x}_0^\top K \boldsymbol{x}_0 - \mathbb{E}_{\boldsymbol{x}} \left(\sigma^{-2} \boldsymbol{x}^\top K \boldsymbol{x} \right)}{\sigma^{-2} \boldsymbol{x}_0^\top K \boldsymbol{x}_0} = 1 - \frac{\mathrm{tr}(K\boldsymbol{C})}{\boldsymbol{x}_0^\top K \boldsymbol{x}_0}.$$

Numerical Study: Set Up

- Generate random adjacency matrix with a constant network density (i.e., $\Pr(w_{ij} = 1)$ for $i \neq j$)
- Generate covariates vectors with balanced entries $\{-1, 1\}$
- Optimization problems can be solved in Gurobi with time limit setup to be 500 sec.
- The correlation value ρ_0 in the optimization is fixed as 0.5.

Numerical Study: Different α

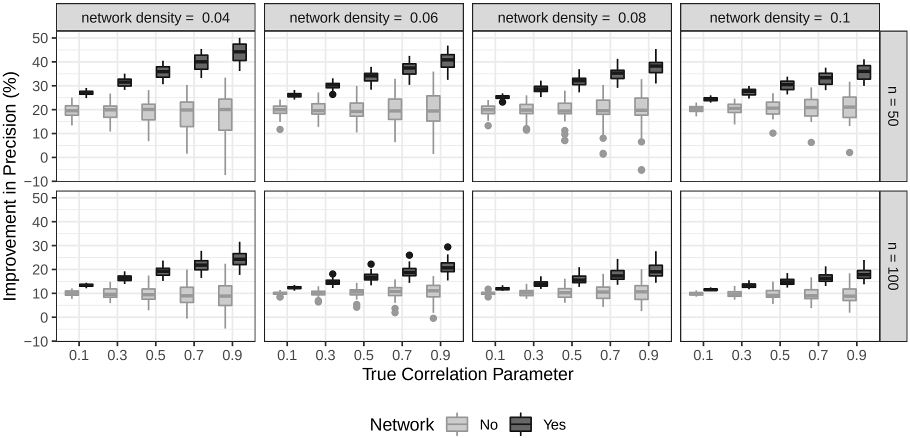
Recall the constraint: $T_1(\boldsymbol{x}) \leq \mathrm{sd}(T_1(\boldsymbol{x}))q_\alpha$.



We fix $\alpha = 0.001$ in the rest of the numerical study.

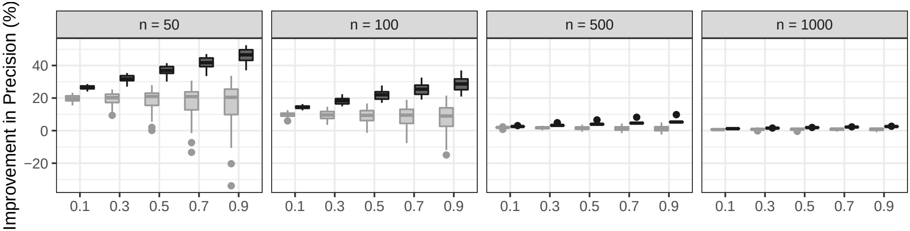
Numerical Study: Network Density

Recall that network density $= \Pr(w_{ij} = 1)$ for $i \neq j$ approx $\sum_{ij} w_{ij}/n^2$.



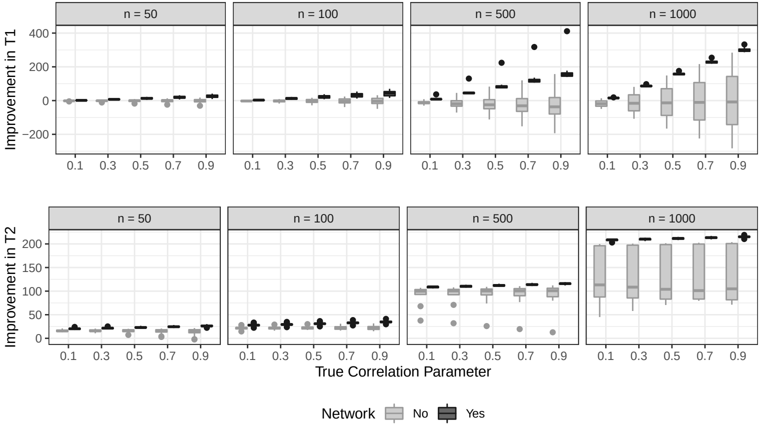
Numerical Study: Size of Network

Fix network density $= 0.02$.



Recall that $\boldsymbol{x}^\top K\boldsymbol{x} = \sum_{ij} w_{ij} - \rho T_1(\boldsymbol{x}) - T_2(\boldsymbol{x}, \rho)$

Numerical Study: Size of Network



Recall that $\boldsymbol{x}^\top K \boldsymbol{x} = \sum_{ij} w_{ij} - \rho T_1(\boldsymbol{x}) - T_2(\boldsymbol{x}, \rho)$

Case Study

Data

- Consider a public available dataset from a music streaming service with $\approx$50k users.
- Friendship network between users
- Covariates: the users' preference (recorded by 1 or 0) to 84 distinct music genres.

Potential Application:

- This type of dataset is typically relevant to the experiments of assessing different algorithms for recommendation systems.
- Such type of experiments requires to split users into two groups to allocate two versions of a recommendation system.
- A numerical outcome will be computed for each user, such as total time spending on recommended products or the click through rate to recommended products over a time period.

Case Study

Data

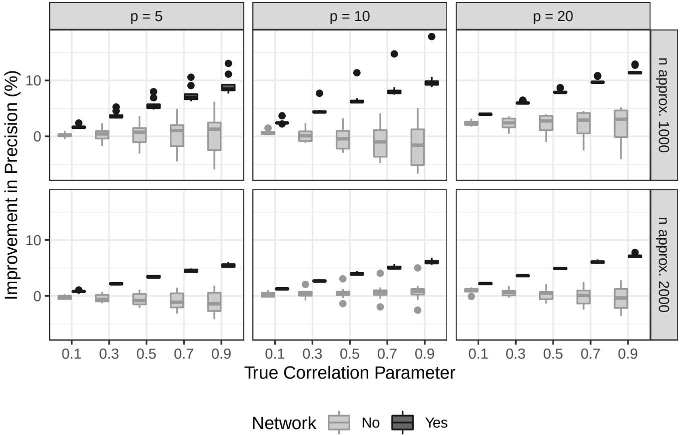
- Consider a public available dataset from a music streaming service with $\approx$50k users.
- Friendship network between users
- Covariates: the users' preference (recorded by 1 or 0) to 84 distinct music genres.

Potential Application:

- This type of dataset is typically relevant to the experiments of assessing different algorithms for recommendation systems.
- Such type of experiments requires to split users into two groups to allocate two versions of a recommendation system.
- A numerical outcome will be computed for each user, such as total time spending on recommended products or the click through rate to recommended products over a time period.

Improvement to Precision

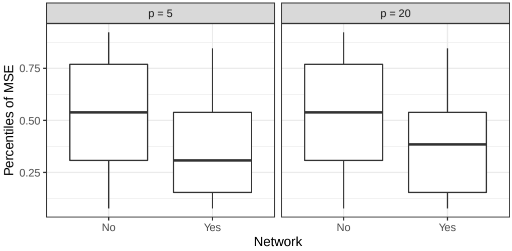
After taking a subset of users and subset of covariates



Note: the network density of each subset is between 0.001-0.002.

Pseudo Experiment

- Generate responses based on CAR model with network dependence $\sigma^2(D - PWP)^{-1}$ with $P = diag(\sqrt{\rho_1}, \ldots, \sqrt{\rho_n})$ and $\rho_1, \ldots, \rho_n$ randomly generated from $U(0,1)$ for $n = 1000$.
- Compute the mean squared errors of the estimated treatment effect over 100 replication.
- Compute the quantiles of MSE of the optimal designs among 10 random designs for each replication



Summary

- Discussed general design problems for network A/B tests
- Introduced a locally optimal design approach for A/B testing with network and covariates information.
- The proposed design is robust to unknown parameters.
- For the model with different network dependence structure, robustness of the proposed design can be investigated before data collection.

My Works

- Pokhiko, V., Zhang, Q., Kang, L., Mays, D. (2019), D-optimal design for network A/B testing. Journal of Statistical Theory and Practice, 13(4), 1-23.
- Zhang, Q. and Kang, L. (2022), Locally Optimal Design for A/B Tests in the Presence of Covariates and Network Dependence. Technometrics. 64(3): 358-369.
- Zhang, Q. (2023+), On the Asymptotics of Graph Cut for Network A/B Tests. To be released.

Thank You!