GEAR: On Optimal Treatment Decision Making by Auxiliary Data with Application to AIDs Study


Hengrui Cai


Rui Song


Wenbin Lu

Department of Statistics, North Carolina State University, North Carolina, USA

DISS2021

Introduction

Personalized optimal treatment decision making has attracted increasing attention recently.

- Developing an individualized treatment rule for patients to optimize expected clinical outcomes of interest [medicine];
- Offering customized incentives to increase sales and level of engagement [Economics];
- Designing a personalized advertisement recommendation system to raise the click rates [marketing].

Introduction

Personalized optimal treatment decision making has attracted increasing attention recently.

- Developing an individualized treatment rule for patients to optimize expected clinical outcomes of interest [medicine];
- Offering customized incentives to increase sales and level of engagement [Economics];
- Designing a personalized advertisement recommendation system to raise the click rates [marketing].

Introduction

Personalized optimal treatment decision making has attracted increasing attention recently.

- Developing an individualized treatment rule for patients to optimize expected clinical outcomes of interest [medicine];
- Offering customized incentives to increase sales and level of engagement [Economics];
- Designing a personalized advertisement recommendation system to raise the click rates [marketing].

AIDS Clinical Trials Group Protocol 175 (ACTG 175) data

- A randomized trial to examine competitive antiretroviral regimens for HIV-infected subjects.
- Covariates X include four continuous variables and eight categorical variables.
- Treatment A: zidovudine (ZDV) + zalcitabine (ddC) as treatment 0, and ZDV+didanosine (ddI) as treatment 1.
- Outcome of interest Y: the clinical meaningful long-term outcome of interest for the AIDS recovery, such as the mean CD4 count (cells/mm3) at 96 ± 5 weeks.
- Goal: to find the optimal decision rule for the HIV-infected subjects to maximize the AIDS recovery.

AIDS Clinical Trials Group Protocol 175 (ACTG 175) data

- A randomized trial to examine competitive antiretroviral regimens for HIV-infected subjects.
- Covariates X include four continuous variables and eight categorical variables.
- Treatment A: zidovudine (ZDV) + zalcitabine (ddC) as treatment 0, and ZDV+didanosine (ddI) as treatment 1.
- Outcome of interest Y: the clinical meaningful long-term outcome of interest for the AIDS recovery, such as the mean CD4 count (cells/mm3) at 96 ± 5 weeks.
- Goal: to find the optimal decision rule for the HIV-infected subjects to maximize the AIDS recovery.

AIDS Clinical Trials Group Protocol 175 (ACTG 175) data

- A randomized trial to examine competitive antiretroviral regimens for HIV-infected subjects.
- Covariates X include four continuous variables and eight categorical variables.
- Treatment A: zidovudine (ZDV) + zalcitabine (ddC) as treatment 0, and ZDV+didanosine (ddI) as treatment 1.
- Outcome of interest Y: the clinical meaningful long-term outcome of interest for the AIDS recovery, such as the mean CD4 count (cells/mm3) at 96 ± 5 weeks.
- Goal: to find the optimal decision rule for the HIV-infected subjects to maximize the AIDS recovery.

AIDS Clinical Trials Group Protocol 175 (ACTG 175) data

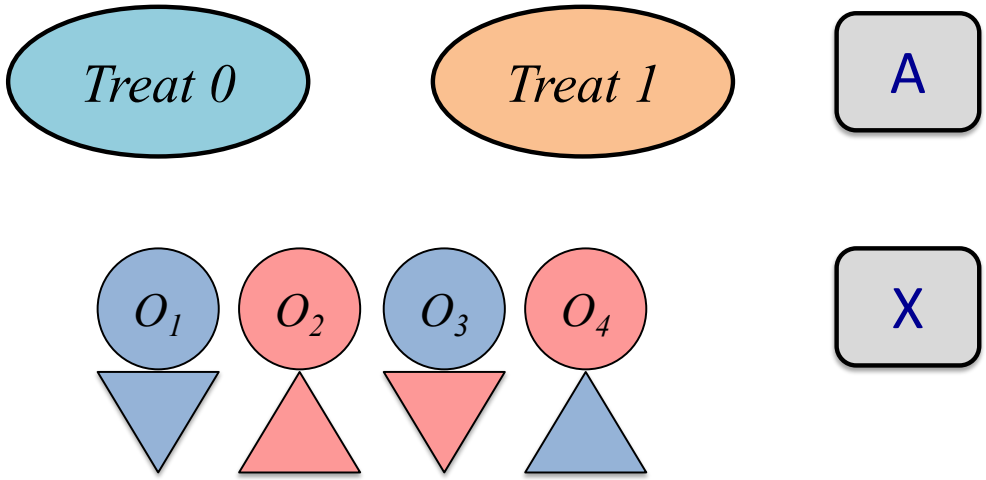
- A randomized trial to examine competitive antiretroviral regimens for HIV-infected subjects.
- Covariates X include four continuous variables and eight categorical variables.
- Treatment A: zidovudine (ZDV) + zalcitabine (ddC) as treatment 0, and ZDV+didanosine (ddI) as treatment 1.
- Outcome of interest Y: the clinical meaningful long-term outcome of interest for the AIDS recovery, such as the mean CD4 count (cells/mm3) at 96 ± 5 weeks.
- Goal: to find the optimal decision rule for the HIV-infected subjects to maximize the AIDS recovery.

AIDS Clinical Trials Group Protocol 175 (ACTG 175) data

- A randomized trial to examine competitive antiretroviral regimens for HIV-infected subjects.
- Covariates X include four continuous variables and eight categorical variables.
- Treatment A: zidovudine (ZDV) + zalcitabine (ddC) as treatment 0, and ZDV+didanosine (ddI) as treatment 1.
- Outcome of interest Y: the clinical meaningful long-term outcome of interest for the AIDS recovery, such as the mean CD4 count (cells/mm3) at 96 ± 5 weeks.
- Goal: to find the optimal decision rule for the HIV-infected subjects to maximize the AIDS recovery.

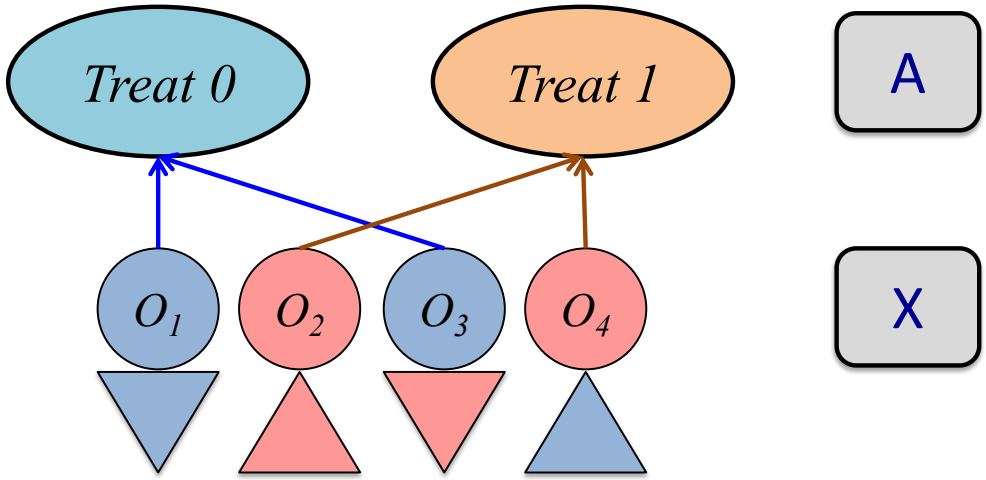
General Setup for Finding the Optimal Decision Rule

Consider a decision making problem to assign individuals with appropriate treatment options.



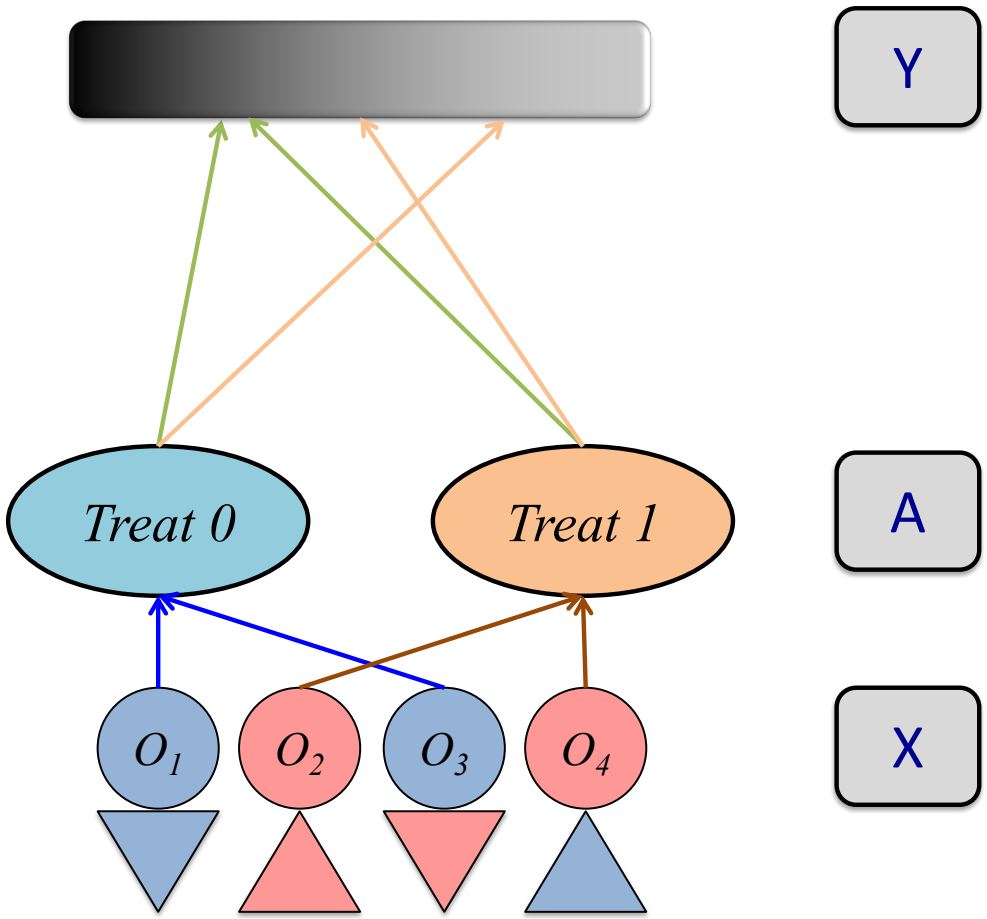
General Setup for Finding the Optimal Decision Rule

Due to individuals' heterogeneity in outcome to different treatment options, there may not exist a unified best decision.



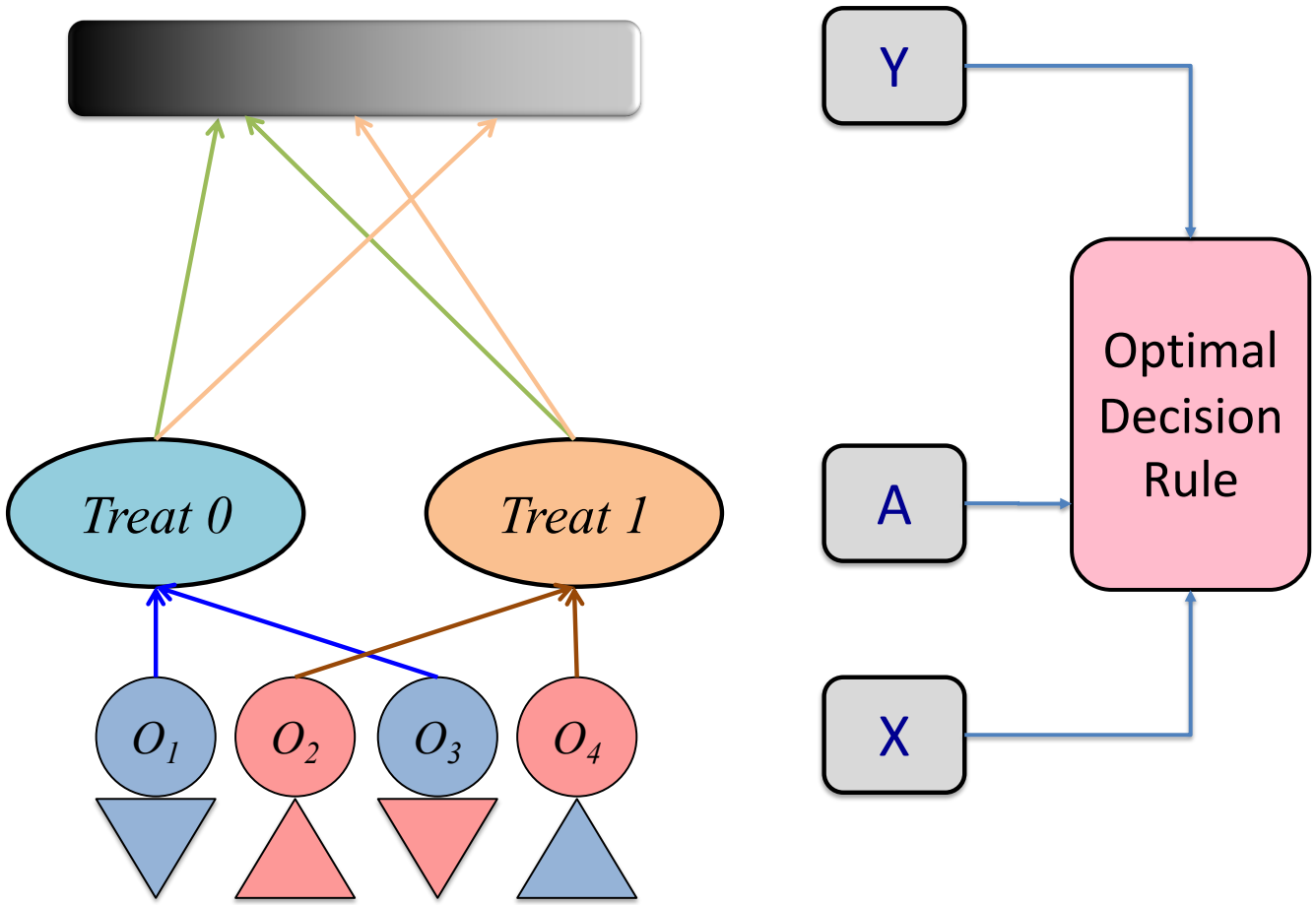
General Setup for Finding the Optimal Decision Rule

The outcome Y can be observed after a treatment is given, in an experimental sample (from either randomized trials or observational studies).



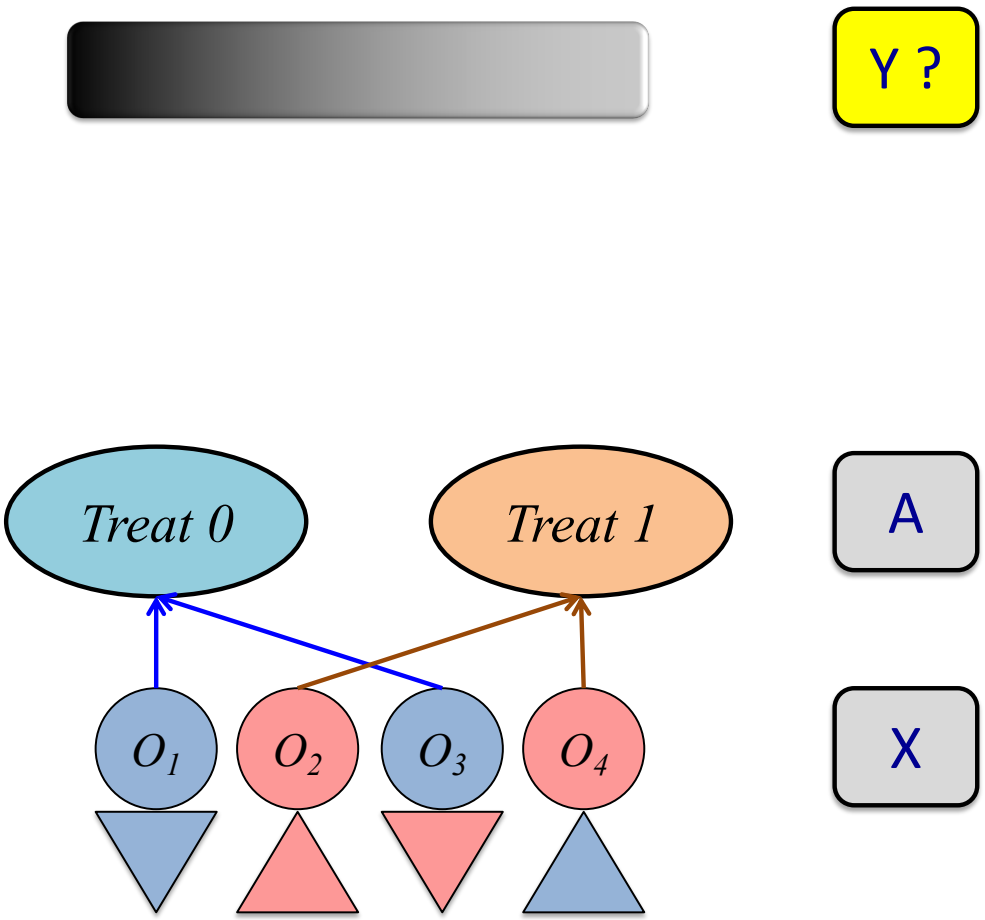
General Setup for Finding the Optimal Decision Rule

The optimal decision rule (ODR) is to assign individuals with the best treatment option according to their covariates.



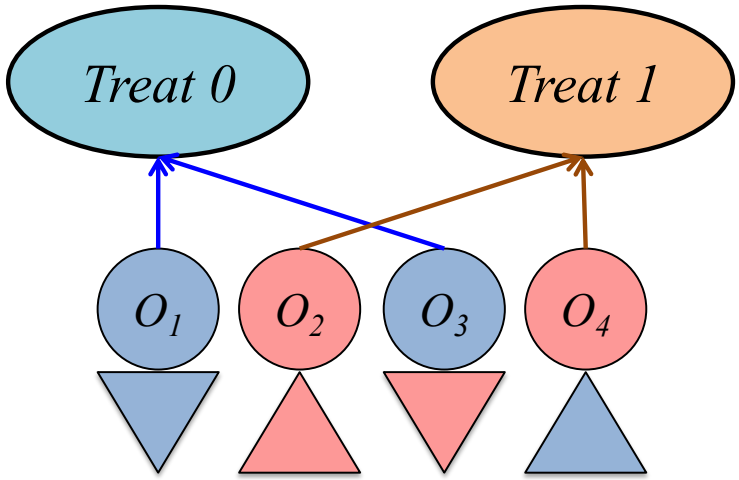
Motivation: Unobserved Long-Term Outcome

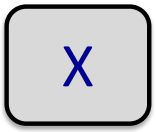
The long-term primary outcome of interest may not be observed due to the limited duration of experiments, such as the AIDS recovery.



Motivation: Unobserved Long-Term Outcome

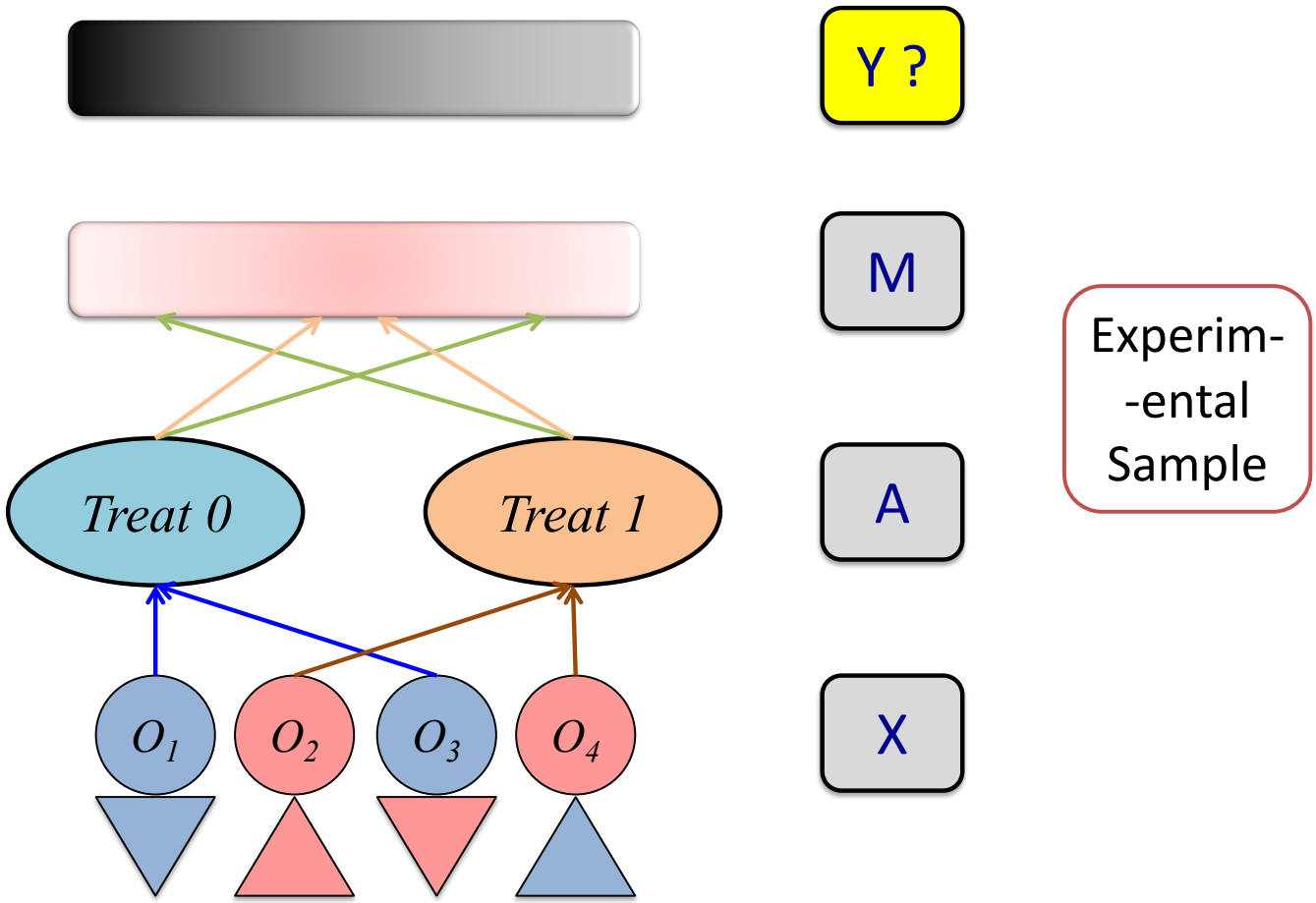
The mean CD4 count (cells/mm3) at 96 $\pm$ 5 weeks is missing for a proportion of patients in ACTG 175 data, due to the limitation of the follow-up.





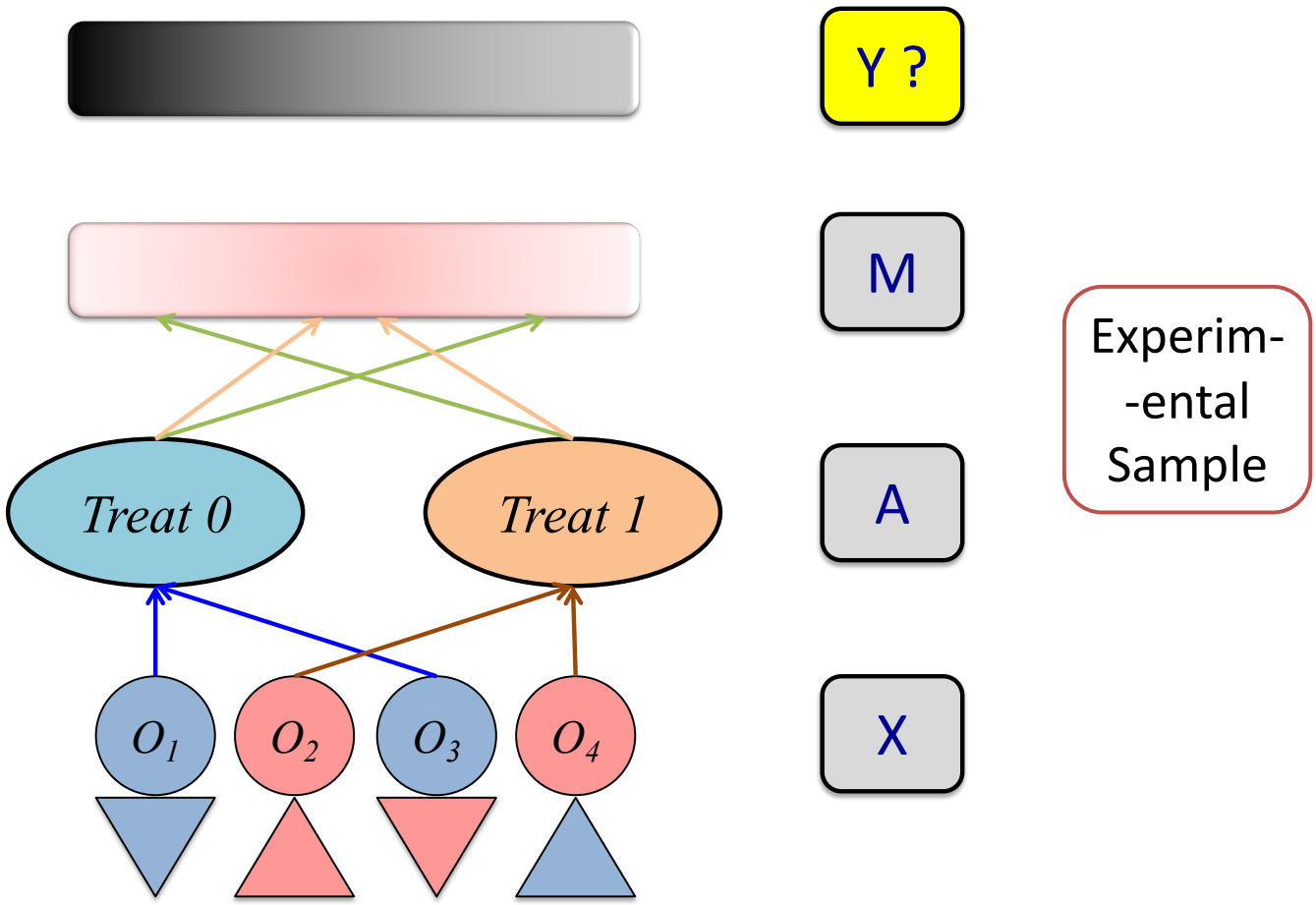
Motivation: Unobserved Long-Term Outcome

Intermediate Outcomes M: the mean CD4 count and the mean CD8 count at 20 $\pm$ 5 weeks.



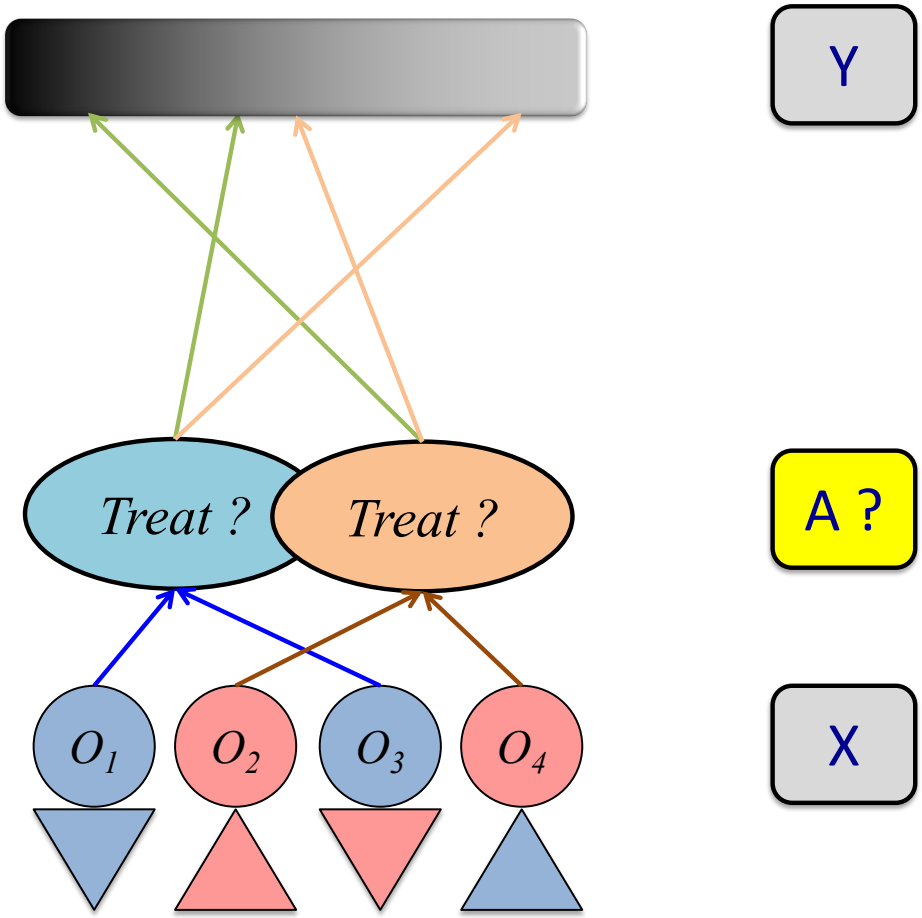
Motivation: Unobserved Long-Term Outcome

How to address the challenge of developing ODR when the long-term outcome cannot be observed in the experimental sample?



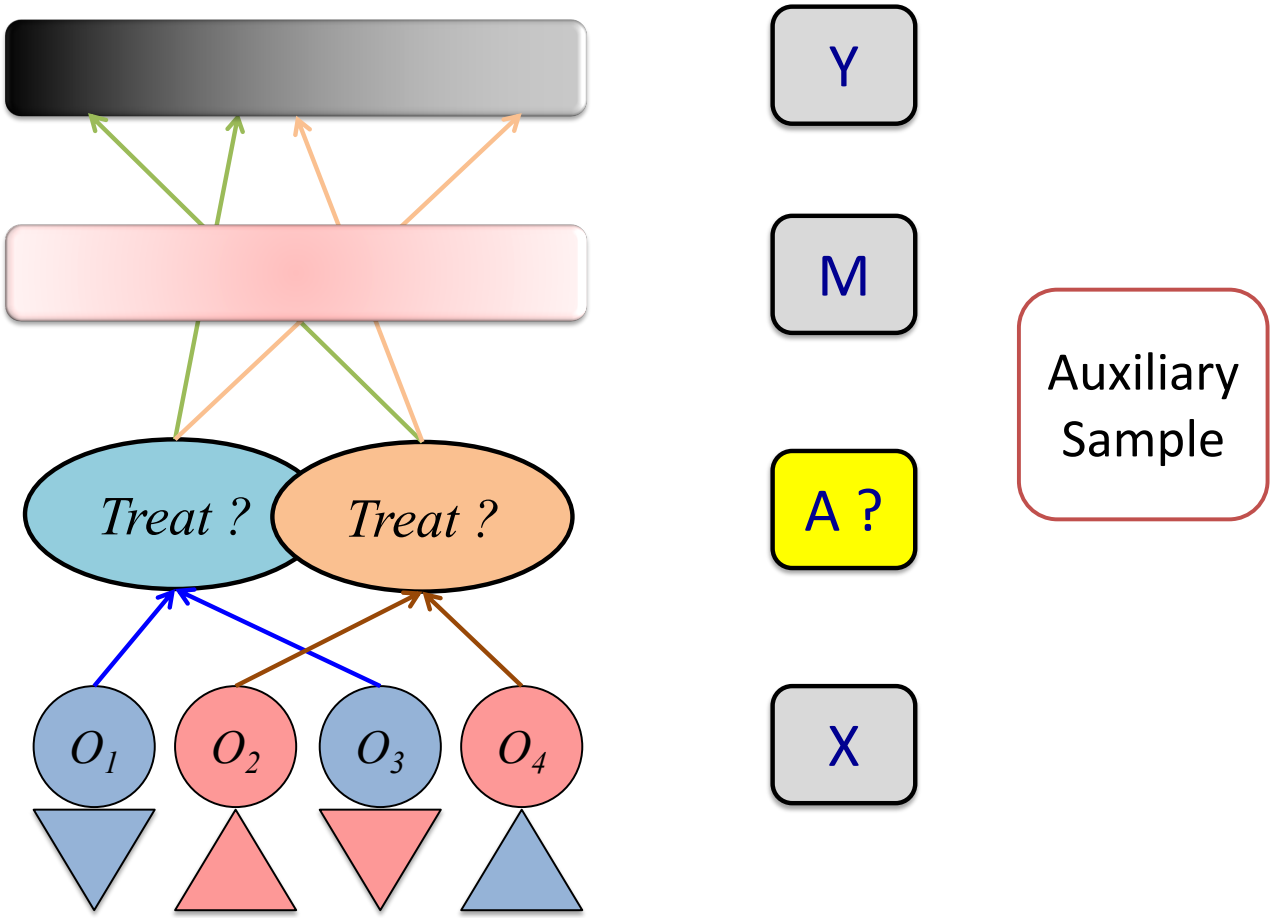
Motivation: Auxiliary Data Source

Auxiliary data (such as electronic medical records or administrative records) contains rich information, except for treatment information.



Motivation: Auxiliary Data Source

We make use of the auxiliary sample to facilitate the estimation of ODR in the experimental sample.



Overview

- Statistical framework to estimate the heterogeneous treatment effect for the long-term outcome that cannot be observed in an experiment.
- Proposing an **auGmented inverse propensity weighted Experimental and Auxiliary sample-based decision Rule**, named GEAR, by augmented inverse propensity weighted (AIPW) estimator.
- Deriving the consistency and asymptotic distribution of the AIPW estimator under the proposed GEAR. A confidence interval (CI) for the estimated expected outcome is provided.
- Simulation studies are conducted to demonstrate its empirical validity with an AIDS application.

Overview

- Statistical framework to estimate the heterogeneous treatment effect for the long-term outcome that cannot be observed in an experiment.
- Proposing an **auGmented inverse propensity weighted Experimental and Auxiliary sample-based decision Rule**, named GEAR, by augmented inverse propensity weighted (AIPW) estimator.
- Deriving the consistency and asymptotic distribution of the AIPW estimator under the proposed GEAR. A confidence interval (CI) for the estimated expected outcome is provided.
- Simulation studies are conducted to demonstrate its empirical validity with an AIDS application.

Overview

- Statistical framework to estimate the heterogeneous treatment effect for the long-term outcome that cannot be observed in an experiment.
- Proposing an **auGmented inverse propensity weighted Experimental and Auxiliary sample-based decision Rule**, named GEAR, by augmented inverse propensity weighted (AIPW) estimator.
- Deriving the consistency and asymptotic distribution of the AIPW estimator under the proposed GEAR. A confidence interval (CI) for the estimated expected outcome is provided.
- Simulation studies are conducted to demonstrate its empirical validity with an AIDS application.

Overview

- Statistical framework to estimate the heterogeneous treatment effect for the long-term outcome that cannot be observed in an experiment.
- Proposing an **auGmented inverse propensity weighted Experimental and Auxiliary sample-based decision Rule**, named GEAR, by augmented inverse propensity weighted (AIPW) estimator.
- Deriving the consistency and asymptotic distribution of the AIPW estimator under the proposed GEAR. A confidence interval (CI) for the estimated expected outcome is provided.
- Simulation studies are conducted to demonstrate its empirical validity with an AIDS application.

Statistical Framework

E: an experimental sample of interest with sample size N_E.

- X_E: r-dimensional individual's baseline covariates.
- $A_E \in \{0, 1\}$: the treatment an individual receives.
- M_E: s-dimensional intermediate outcomes.
- Dataset $\{E_i = (X_{E,i}, A_{E,i}, M_{E,i}), i = 1, \ldots, N_E\}$.
- Y_E: long-term outcome of interest cannot be observed.

U: an auxiliary sample with sample size N_U.

- X_U: r-dimensional individual's baseline covariates.
- M_E: s-dimensional intermediate outcomes.
- Y_E: long-term outcome of interest.
- Dataset $\{U_i = (X_{U,i}, M_{U,i}, Y_{U,i}), i = 1, \ldots, N_U\}$.
- However, treatment information is not available.

$R = \{E, U\}$ to indicate the missingness of each sample.

Statistical Framework

E: an experimental sample of interest with sample size N_E.

- X_E: r-dimensional individual's baseline covariates.
- $A_E \in \{0, 1\}$: the treatment an individual receives.
- M_E: s-dimensional intermediate outcomes.
- Dataset $\{E_i = (X_{E,i}, A_{E,i}, M_{E,i}), i = 1, \ldots, N_E\}$.
- Y_E: long-term outcome of interest cannot be observed.

U: an auxiliary sample with sample size N_U.

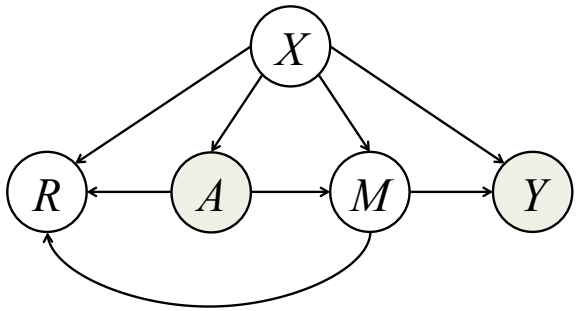
- X_U: r-dimensional individual's baseline covariates.
- M_E: s-dimensional intermediate outcomes.
- Y_E: long-term outcome of interest.
- Dataset $\{U_i = (X_{U,i}, M_{U,i}, Y_{U,i}), i = 1, \ldots, N_U\}$.
- However, treatment information is not available.

$R = \{E, U\}$ to indicate the missingness of each sample.

Assumptions

Define the potential outcomes $Y_E^*(0)$ and $Y_E^*(1)$ as the long-term outcome that would be observed after treatment 0 or 1, respectively.
Let the propensity score function as $\pi(x) = Pr_E(A_{E,i} = 1 | X_{E,i} = x)$.

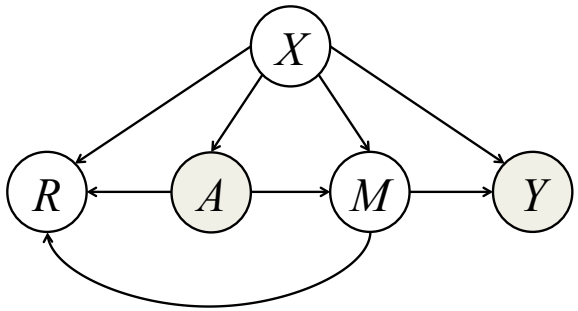
- (A1). Stable Unit Treatment Value Assumption (SUTVA): $Y_E = A_E Y_E^\star(1) + (1 - A_E) Y_E^\star(0)$.
- (A2). No Unmeasured Confounders Assumption: $\{Y_E^*(0), Y_E^*(1)\} \perp\!\!\!\perp A_E \mid X_E$.
- (A3). Positivity: $0 < \pi(x) < 1$ for all $x \in \mathbb{X}_E$.
- (A4). Comparability Assumption: $Y \perp\!\!\!\perp R \mid X, M$.
- (A5). Surrogacy Assumption: $Y_E \perp\!\!\!\perp A_E \mid X_E, M_E$.



Assumptions

Define the potential outcomes $Y_E^*(0)$ and $Y_E^*(1)$ as the long-term outcome that would be observed after treatment 0 or 1, respectively.
Let the propensity score function as $\pi(x) = Pr_E(A_{E,i} = 1 | X_{E,i} = x)$.

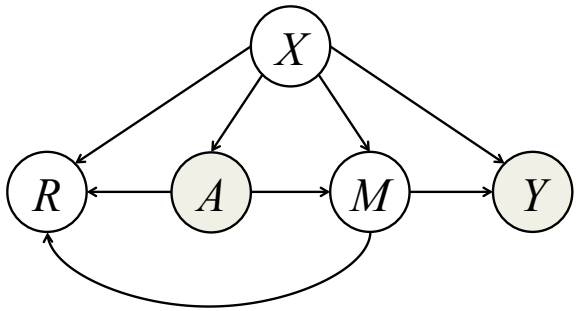
- (A1). Stable Unit Treatment Value Assumption (SUTVA): $Y_E = A_E Y_E^\star(1) + (1 - A_E) Y_E^\star(0)$.
- (A2). No Unmeasured Confounders Assumption: $\{Y_E^*(0), Y_E^*(1)\} \perp\!\!\!\perp A_E \mid X_E$.
- (A3). Positivity: $0 < \pi(x) < 1$ for all $x \in \mathbb{X}_E$.
- (A4). Comparability Assumption: $Y \perp\!\!\!\perp R \mid X, M$.
- (A5). Surrogacy Assumption: $Y_E \perp\!\!\!\perp A_E \mid X_E, M_E$.



Assumptions

Define the potential outcomes $Y_E^*(0)$ and $Y_E^*(1)$ as the long-term outcome that would be observed after treatment 0 or 1, respectively.
Let the propensity score function as $\pi(x) = Pr_E(A_{E,i} = 1 | X_{E,i} = x)$.

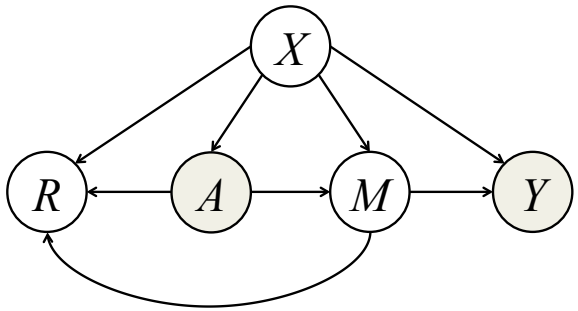
- (A1). Stable Unit Treatment Value Assumption (SUTVA): $Y_E = A_E Y_E^\star(1) + (1 - A_E) Y_E^\star(0)$.
- (A2). No Unmeasured Confounders Assumption: $\{Y_E^*(0), Y_E^*(1)\} \perp\!\!\!\perp A_E \mid X_E$.
- (A3). Positivity: $0 < \pi(x) < 1$ for all $x \in \mathbb{X}_E$.
- (A4). Comparability Assumption: $Y \perp\!\!\!\perp R \mid X, M$.
- (A5). Surrogacy Assumption: $Y_E \perp\!\!\!\perp A_E \mid X_E, M_E$.



Assumptions

Define the potential outcomes $Y_E^*(0)$ and $Y_E^*(1)$ as the long-term outcome that would be observed after treatment 0 or 1, respectively.
Let the propensity score function as $\pi(x) = Pr_E(A_{E,i} = 1 | X_{E,i} = x)$.

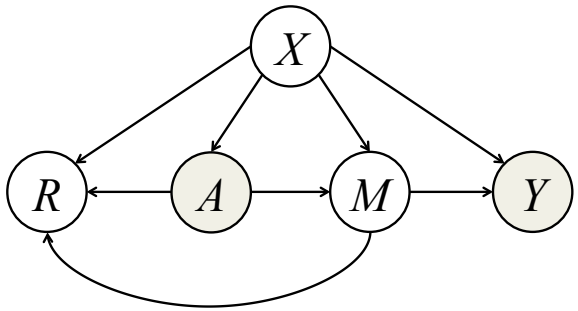
- (A1). Stable Unit Treatment Value Assumption (SUTVA): $Y_E = A_E Y_E^\star(1) + (1 - A_E) Y_E^\star(0)$.
- (A2). No Unmeasured Confounders Assumption: $\{Y_E^*(0), Y_E^*(1)\} \perp\!\!\!\perp A_E \mid X_E$.
- (A3). Positivity: $0 < \pi(x) < 1$ for all $x \in \mathbb{X}_E$.
- (A4). Comparability Assumption: $Y \perp\!\!\!\perp R \mid X, M$.
- (A5). Surrogacy Assumption: $Y_E \perp\!\!\!\perp A_E \mid X_E, M_E$.



Assumptions

Define the potential outcomes $Y_E^*(0)$ and $Y_E^*(1)$ as the long-term outcome that would be observed after treatment 0 or 1, respectively.
Let the propensity score function as $\pi(x) = Pr_E(A_{E,i} = 1 | X_{E,i} = x)$.

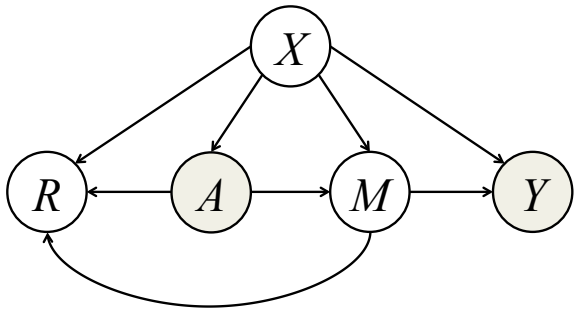
- (A1). Stable Unit Treatment Value Assumption (SUTVA): $Y_E = A_E Y_E^\star(1) + (1 - A_E) Y_E^\star(0)$.
- (A2). No Unmeasured Confounders Assumption: $\{Y_E^*(0), Y_E^*(1)\} \perp\!\!\!\perp A_E \mid X_E$.
- (A3). Positivity: $0 < \pi(x) < 1$ for all $x \in \mathbb{X}_E$.
- (A4). Comparability Assumption: $Y \perp\!\!\!\perp R \mid X, M$.
- (A5). Surrogacy Assumption: $Y_E \perp\!\!\!\perp A_E \mid X_E, M_E$.



Assumptions

Define the potential outcomes $Y_E^*(0)$ and $Y_E^*(1)$ as the long-term outcome that would be observed after treatment 0 or 1, respectively.
Let the propensity score function as $\pi(x) = Pr_E(A_{E,i} = 1 | X_{E,i} = x)$.

- (A1). Stable Unit Treatment Value Assumption (SUTVA): $Y_E = A_E Y_E^\star(1) + (1 - A_E) Y_E^\star(0)$.
- (A2). No Unmeasured Confounders Assumption: $\{Y_E^*(0), Y_E^*(1)\} \perp\!\!\!\perp A_E \mid X_E$.
- (A3). Positivity: $0 < \pi(x) < 1$ for all $x \in \mathbb{X}_E$.
- (A4). Comparability Assumption: $Y \perp\!\!\!\perp R \mid X, M$.
- (A5). Surrogacy Assumption: $Y_E \perp\!\!\!\perp A_E \mid X_E, M_E$.



Statistical Framework

- Decision Rule $d(\cdot)$ is a deterministic function that maps $\mathbb{X}_E$ to $\{0, 1\}$.
- Value function under $d(\cdot)$ is $V(d) = \mathrm{E}\{Y^*(d)\}$, where $Y_E^*(d) = Y_E^*(0)\{1 - d(X_E)\} + Y_E^*(1)d(X_E)$ is the potential outcome under $d(\cdot)$ that would be observed if an individual had received a treatment according to $d(\cdot)$.
- Optimal Decision Rule (ODR) is to maximize the value function over the experimental sample among a class of decision rules of interest as $d^{opt}(\cdot) = \arg\min_{d(\cdot)} V(d)$.

Table 1: The data structure of the experimental sample and the auxiliary sample.

Sample	X	A	M	Y	ODR for Y	ODR for M
Experimental	✓	✓	✓	×	×	✓
Auxiliary	✓	×	✓	✓	×	×

Statistical Framework

- Decision Rule $d(\cdot)$ is a deterministic function that maps $\mathbb{X}_E$ to $\{0, 1\}$.
- Value function under $d(\cdot)$ is $V(d) = \mathrm{E}\{Y^*(d)\}$, where $Y_E^*(d) = Y_E^*(0)\{1 - d(X_E)\} + Y_E^*(1)d(X_E)$ is the potential outcome under $d(\cdot)$ that would be observed if an individual had received a treatment according to $d(\cdot)$.
- Optimal Decision Rule (ODR) is to maximize the value function over the experimental sample among a class of decision rules of interest as $d^{opt}(\cdot) = \arg\min_{d(\cdot)} V(d)$.

Table 1: The data structure of the experimental sample and the auxiliary sample.

Sample	X	A	M	Y	ODR for Y	ODR for M
Experimental	✓	✓	✓	×	×	✓
Auxiliary	✓	×	✓	✓	×	×

Statistical Framework

- Decision Rule $d(\cdot)$ is a deterministic function that maps $\mathbb{X}_E$ to $\{0, 1\}$.
- Value function under $d(\cdot)$ is $V(d) = \mathrm{E}\{Y^*(d)\}$, where $Y_E^*(d) = Y_E^*(0)\{1 - d(X_E)\} + Y_E^*(1)d(X_E)$ is the potential outcome under $d(\cdot)$ that would be observed if an individual had received a treatment according to $d(\cdot)$.
- Optimal Decision Rule (ODR) is to maximize the value function over the experimental sample among a class of decision rules of interest as $d^{opt}(\cdot) = \arg\min_{d(\cdot)} V(d)$.

Table 1: The data structure of the experimental sample and the auxiliary sample.

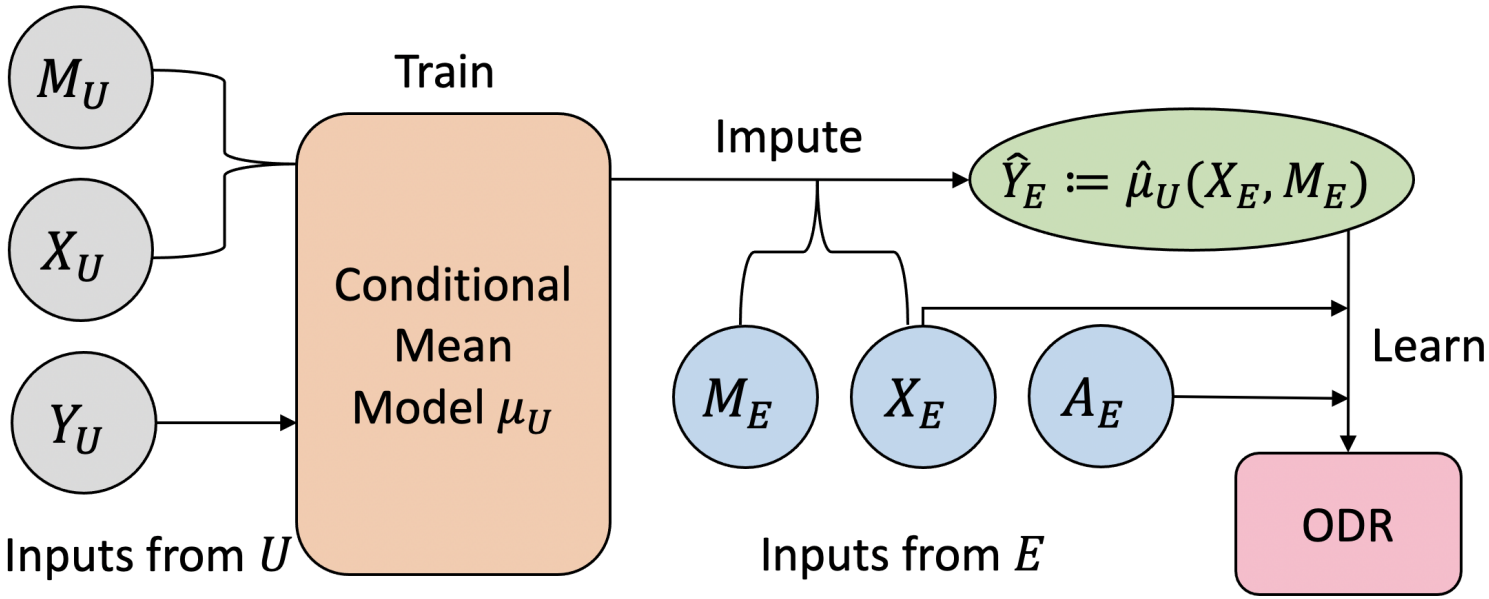
Sample	X	A	M	Y	ODR for Y	ODR for M
Experimental	✓	✓	✓	×	×	✓
Auxiliary	✓	×	✓	✓	×	×

Architecture of the Proposed GEAR

Denote $\mu_E(m,x) \equiv \mathsf{E}[Y_E|M_E = m, X_E = x]$, and $\mu_U(m,x) \equiv \mathsf{E}[Y_U|M_U = m, X_U = x]$.

Corollary (Equal Conditional Mean)

Under (A1) - (A4), $\mu_E(m,x) = \mu_U(m,x)$, for any x and m.



AIPW Estimator for Long-Term Outcome

Suppose the decision rule $d(\cdot)$ relies on a model parameter β, denoted as $d(\cdot) \equiv d(\cdot;\beta)$. We use a shorthand to write $V(d)$ as $V(\beta)$.

Theorem 1

Under (A1)-(A5), given $d(\cdot;\beta)$, we have

$$\begin{aligned} & E_{Y_E|X_E}\{Y_E|A_E = d(X_E;\beta), X_E\} \\ = \quad & E_{M_E|X_E}\{\mu_U(M_E, X_E)|A_E = d(X_E;\beta), X_E\}. \end{aligned}$$

AIPW Estimator

$$\widehat{V}_{AIP}(\beta) = \frac{1}{N_E}\sum_{i=1}^{N_E}\left[\widehat{\nu}_i + \frac{\mathbb{I}\{A_{E,i} = d(X_{E,i};\beta)\}\{\widehat{\mu}_U(M_{E,i}, X_{E,i}) - \widehat{\nu}_i\}}{A_{E,i}\widehat{\pi}(X_{E,i}) + (1 - A_{E,i})\{1 - \widehat{\pi}(X_{E,i})\}}\right],$$

where $\widehat{\nu}_i \equiv \widehat{\mathsf{E}}\{\widehat{\mu}_U(M_{E,i}, X_{E,i})|A_{E,i} = d(X_{E,i};\beta), X_{E,i}\}$ is the estimator for $\nu_i \equiv \mathsf{E}\{\mu_U(M_{E,i}, X_{E,i})|A_{E,i} = d(X_{E,i};\beta), X_{E,i}\}$ as the augmented term.

GEAR and its Implementation

GEAR

Define $\widehat{\beta}^G = \arg\max_\beta \widehat{V}_{AIP}(\beta)$. The GEAR is $d(X; \widehat{\beta}^G)$ with the corresponding estimated value function as $\widehat{V}_{AIP}(\widehat{\beta}^G)$.

- Class of decision rules: generalized linear rules, fixed depth decision trees, and threshold rules.
- Estimation models: the propensity score function π and the conditional mean μ_U can be estimated through any parametric or nonparametric model.
- Estimation of the augmented term:
  1). model $\mu_U(m, x)$ through $\{X_U, M_U, Y_U\}$ as $\widehat{\mu}_U(m, x)$;
  2). plug $\{M_E, X_E\}$ into $\widehat{\mu}_U(m, x)$ and get $\widehat{\mu}_U(M_E, X_E)$ as the conditional mean outcome to impute the missing Y_E;
  3). fit $\widehat{\mu}_U(M_E, X_E)$ on $\{A_E, X_E\}$ to get $\widehat{\nu}_i$.

GEAR and its Implementation

GEAR

Define $\widehat{\beta}^G = \arg\max_\beta \widehat{V}_{AIP}(\beta)$. The GEAR is $d(X; \widehat{\beta}^G)$ with the corresponding estimated value function as $\widehat{V}_{AIP}(\widehat{\beta}^G)$.

- Class of decision rules: generalized linear rules, fixed depth decision trees, and threshold rules.
- Estimation models: the propensity score function π and the conditional mean μ_U can be estimated through any parametric or nonparametric model.
- Estimation of the augmented term:
  1). model $\mu_U(m, x)$ through $\{X_U, M_U, Y_U\}$ as $\widehat{\mu}_U(m, x)$;
  2). plug $\{M_E, X_E\}$ into $\widehat{\mu}_U(m, x)$ and get $\widehat{\mu}_U(M_E, X_E)$ as the conditional mean outcome to impute the missing Y_E;
  3). fit $\widehat{\mu}_U(M_E, X_E)$ on $\{A_E, X_E\}$ to get $\widehat{\nu}_i$.

GEAR and its Implementation

GEAR

Define $\widehat{\beta}^G = \arg\max_\beta \widehat{V}_{AIP}(\beta)$. The GEAR is $d(X; \widehat{\beta}^G)$ with the corresponding estimated value function as $\widehat{V}_{AIP}(\widehat{\beta}^G)$.

- Class of decision rules: generalized linear rules, fixed depth decision trees, and threshold rules.
- Estimation models: the propensity score function π and the conditional mean μ_U can be estimated through any parametric or nonparametric model.
- Estimation of the augmented term:
  1). model $\mu_U(m, x)$ through $\{X_U, M_U, Y_U\}$ as $\widehat{\mu}_U(m, x)$;
  2). plug $\{M_E, X_E\}$ into $\widehat{\mu}_U(m, x)$ and get $\widehat{\mu}_U(M_E, X_E)$ as the conditional mean outcome to impute the missing Y_E;
  3). fit $\widehat{\mu}_U(M_E, X_E)$ on $\{A_E, X_E\}$ to get $\widehat{\nu}_i$.

Technical Assumptions

- (A6). The density of covariates $f_X(x)$ is bounded away from 0 and ∞ and is twice continuously differentiable with bounded derivatives.
- (A7). Both π and μ_U are smooth bounded functions, with their first derivatives exist and bounded.
- (A8). Model for μ_U is correctly specified.
- (A9). Denote $t = \sqrt{\frac{N_E}{N_U}}$ and assume $0 < t < +\infty$.
- (A10). The true value function $V(\beta)$ is twice continuously differentiable at a neighborhood of true.
- (A11). Either the model of the propensity score or the model of the augmented term is correctly specified.

Theoretical Properties

Theorem 2 (Consistency)

Under (A1)-(A9) and (A11),

$$\widehat{V}_{AIP}(\beta) = V(\beta) + o_p(1), \quad \forall \beta.$$

Remark: When the model for $\mu_U(m, x)$ is correctly specified, our AIPW estimator is doubly robust given either the model of the propensity score or the model of the augmented term is correctly specified.

Theoretical Properties

Theorem 3 (Asymptotic Distribution)

Under (A1)-(A11),

$$\sqrt{N_E}\{\widehat{V}_{AIP}(\widehat{\beta}^G) - V(\beta_0)\} \xrightarrow{\mathcal{D}} N(0, \sigma^2_{AIP}),$$

where $\sigma^2_{AIP} = t\sigma^2_U + \sigma^2_E$, $\sigma^2_U = \mathsf{E}[\{\xi_i^{(U)}\}^2]$, and $\sigma^2_E = \mathsf{E}[\{\xi_i^{(E)}\}^2]$. Here, $\xi_i^{(E)}$ and $\xi_i^{(U)}$ are the I.I.D. terms in the experimental sample and auxiliary sample, respectively.

Remark: The asymptotic variance of the AIPW estimator has an additive form that consists of the estimation error from each sample. Proportion of these two estimation variances is controlled by the sample ratio.

Confidence Interval for the Optimal Value

- The variance σ_E^2 and σ_U^2 can be consistently estimated by $\widehat{\sigma}_E^2 = \frac{1}{N_E}\sum_{i=1}^{N_E}\{\widehat{\xi}_i^{(E)}\}^2$ and $\widehat{\sigma}_U^2 = \frac{1}{N_U}\sum_{i=1}^{N_U}\{\widehat{\xi}_i^{(U)}\}^2$.
- Estimate σ_{AIP} through $\widehat{\sigma_{AIP}} \equiv \sqrt{t\widehat{\sigma}_U^2 + \widehat{\sigma}_E^2}$.

A two-sided $1-\alpha$ confidence interval (CI) for $V(\beta_0)$ under GAER

$$\left[\widehat{V}_{AIP}(\widehat{\beta}^G) - \frac{z_{\alpha/2}\widehat{\sigma_{AIP}}}{\sqrt{N_E}}, \quad \widehat{V}_{AIP}(\widehat{\beta}^G) + \frac{z_{\alpha/2}\widehat{\sigma_{AIP}}}{\sqrt{N_E}}\right],$$

where $z_{\alpha/2}$ denotes the upper $\alpha/2-$th quantile of a standard normal distribution.

Simulation Studies

Data generated from

$$X^{(1)}, X^{(2)}, \cdots, X^{(r)} \overset{iid}{\sim} Uniform[-1,1], \quad A \overset{iid}{\sim} Bernoulli(0.5),$$
$$M = H^M(X) + AC^M(X) + \epsilon^M, \quad Y = H^Y(X) + C^Y(X,M) + \epsilon^Y,$$

where ϵ^M and ϵ^Y are random errors following $N(0, 0.5)$.

$$\textbf{S1}: \begin{cases} H^M(X) = \begin{bmatrix} X^{(3)} \\ X^{(1)} \end{bmatrix}, C^M(X) = \begin{bmatrix} 4\{X^{(1)} - X^{(2)}\} \\ 4\{X^{(4)} - X^{(3)}\} \end{bmatrix}, \\ H^Y(X) = -1 + X^{(2)} + X^{(4)}, C^Y(X,M) = M^{(1)} + M^{(2)}. \end{cases}$$

$$\textbf{S2}: \begin{cases} H^M(X) = \begin{bmatrix} \{X^{(1)}\}^2 X^{(3)} + \sin\{X^{(4)}\} \\ \{X^{(1)}\}^3 - \{X^{(2)} - X^{(4)}\}^2 \end{bmatrix}, \\ C^M(X) = \begin{bmatrix} 4\{X^{(1)} - X^{(2)}\} \\ 4\{X^{(4)} - X^{(3)}\} \end{bmatrix}, \\ H^Y(X) = -1 + X^{(2)} + X^{(4)}, C^Y(X,M) = M^{(1)} + M^{(2)}. \end{cases}$$

Simulation Studies

Table 2: Empirical results under the GEAR for Scenario 1 and 2.

	Scenario 1			Scenario 2		
$N_E =$	200	400	800	200	400	800
$V(\beta_0)$		0.87			0.20	
$\widehat{V}_{AIP}(\widehat{\beta}^G)$	0.89	0.89	0.88	0.24	0.24	0.22
$SE\{\widehat{V}_{AIP}\}$	0.02	0.01	0.01	0.02	0.01	0.01
$\mathrm{E}\{\widehat{\sigma}_{AIP}\}$	0.02	0.01	0.01	0.02	0.01	0.01
$V(\widehat{\beta}^G)$	0.85	0.86	0.86	0.18	0.18	0.19
CP (%)	94.6	94.8	94.8	95.0	94.4	94.8
RCD (%)	95.9	96.6	97.3	95.0	95.8	96.7
$\|\|\widehat{\beta}^G - \beta_0\|\|_2$	0.12	0.09	0.07	0.14	0.11	0.09

Sensitivity Studies when Surrogacy Assumption is Violated

Scenario 3:

$$\begin{cases} H^M(X) = \begin{bmatrix} 0 \\ X^{(1)} \end{bmatrix}, C^M(X) = \begin{bmatrix} -0.5 + 0.4X^{(1)} - 0.6X^{(2)} \\ 0.5 + 0.6X^{(1)} - 0.4X^{(2)} \end{bmatrix}, \\ H^Y(X) = X^{(2)}, \quad C^Y(X, M) = M^{(1)} + M^{(2)}. \end{cases}$$

We use the following $M^{(1)}_{par}$ as one contaminated intermediate outcome we collected instead of the original $M^{(1)}$,

$$M^{(1)}_{par} = M^{(1)} + A(1-l)\{-0.5 + 0.4X^{(1)}\},$$

where the parameter l chosen from $\{0, 0.2, 0.4, 0.6, 0.8, 1\}$ reflects the uncollected information related to the long-term outcome.

Sensitivity Studies when Surrogacy Assumption is Violated

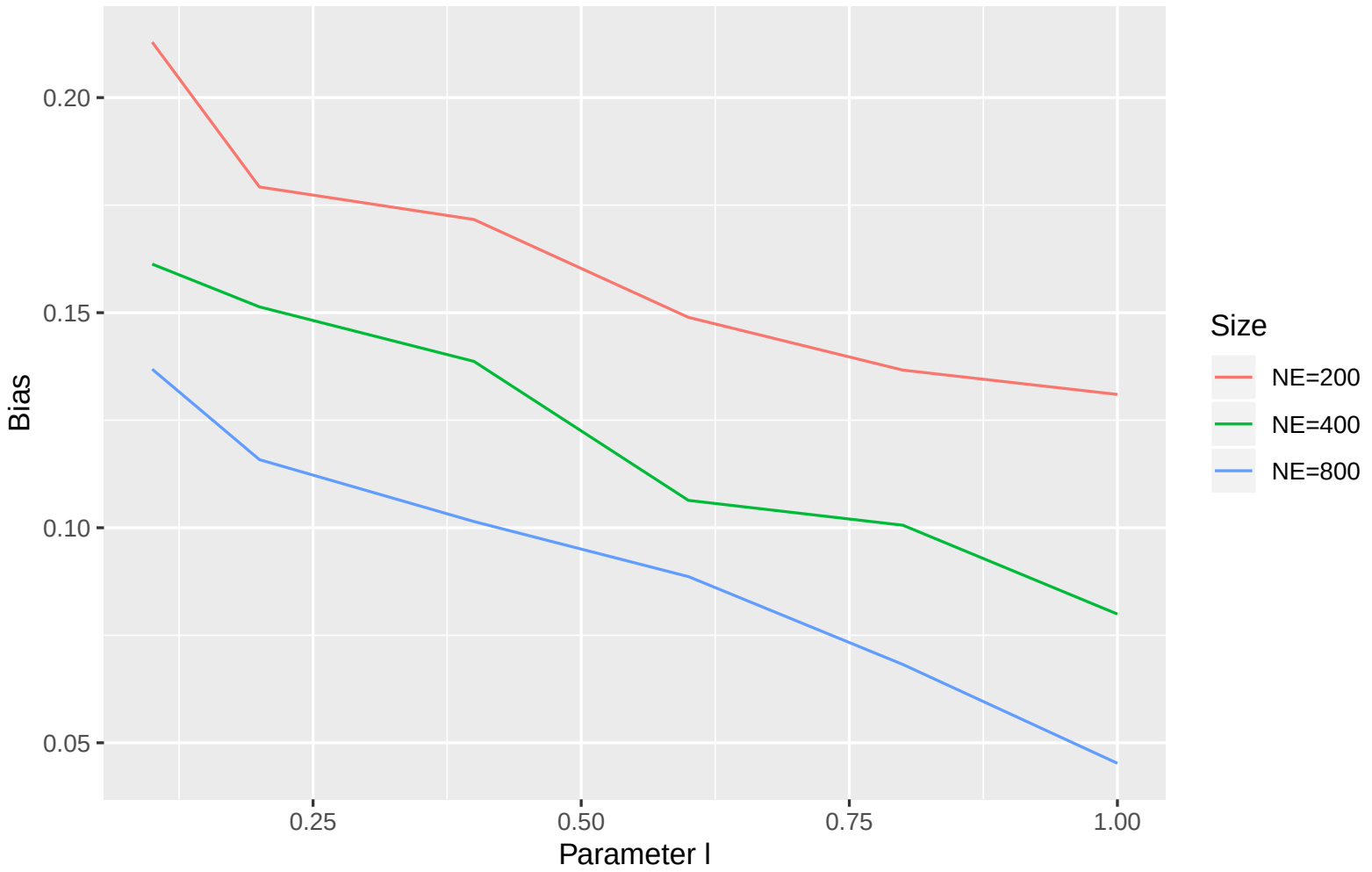


Figure 1: The trend of the bias of $V(\widehat{\beta}^G)$ under the GEAR over the parameter l.

Sensitivity Studies when Surrogacy Assumption is Violated

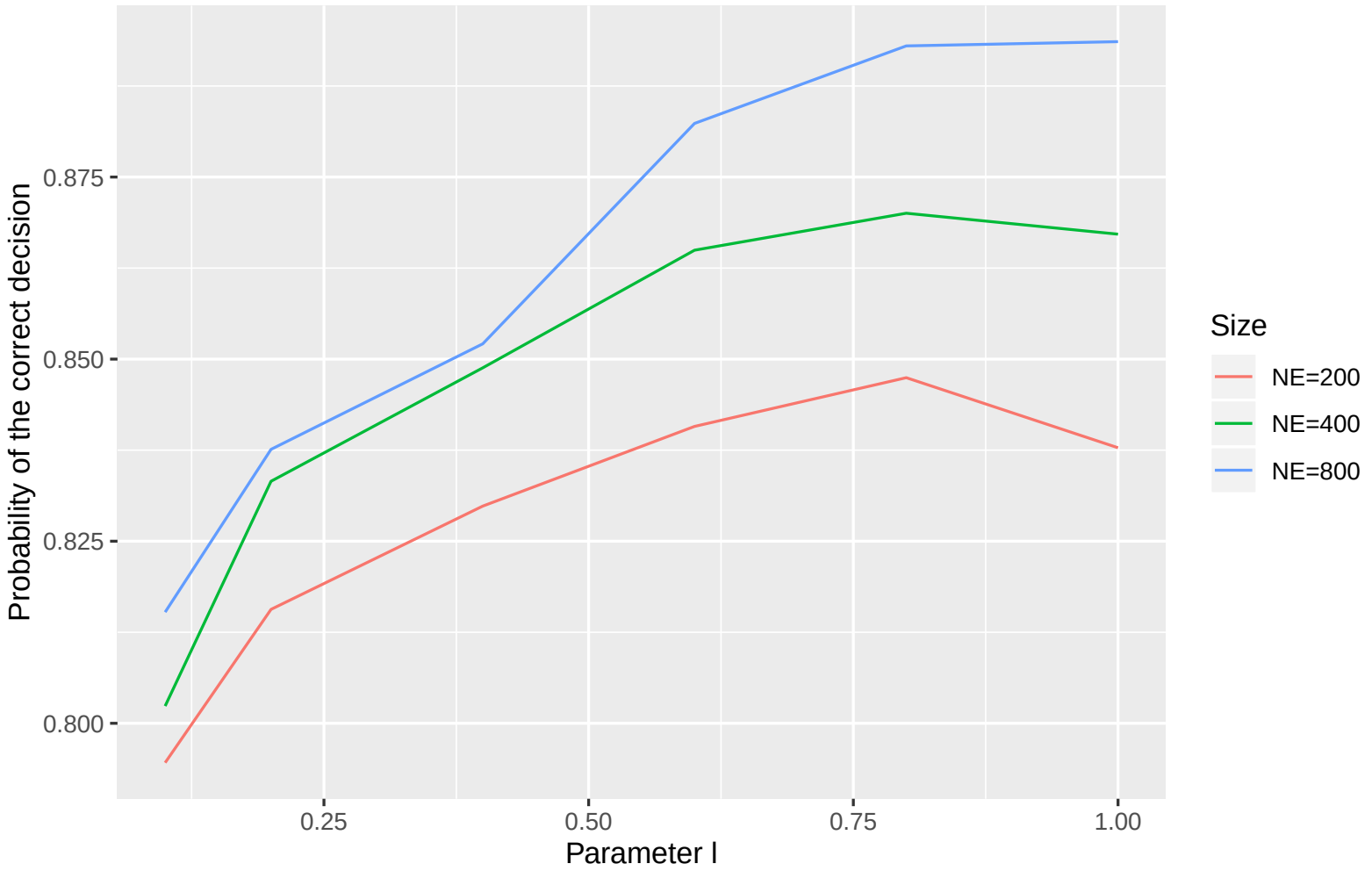


Figure 2: The trend of the average rate of the correct decision made by the GEAR over the parameter l.

Real Data Analysis: ACTG 175 Dataset

Experimental sample E:

- There are $N_E = 376$ HIV-infected subjects randomized to two competitive antiretroviral regimens.
  1). 187 patients were randomized to zidovudine (ZDV) + zalcitabine (ddC), denoted as 'ZDV+ddC' as treatment 0;
  2). 189 patients were randomized to ZDV+didanosine (ddI), denoted as 'ZDV+ddI' as treatment 1.
- $r = 12$ Covariates X:
  1). 4 continuous variables: age (years), weight (kg), CD4 count (cells/mm3) at baseline, and CD8 count (cells/mm3) at baseline;
  2). 8 categorical variables: hemophilia, homosexual activity, history of intravenous drug use, Karnofsky score (scale of 0-100), race (0=white, 1=non-white), gender (0=female), antiretroviral history (0=naive, 1=experienced), and symptomatic status (0=asymptomatic).
- $s = 2$ Intermediate Outcomes M: CD4 count at 20 ± 5 weeks, and CD8 count at 20 ± 5 weeks.

Real Data Analysis: ACTG 175 Dataset

Experimental sample E:

- There are $N_E = 376$ HIV-infected subjects randomized to two competitive antiretroviral regimens.
  1). 187 patients were randomized to zidovudine (ZDV) + zalcitabine (ddC), denoted as 'ZDV+ddC' as treatment 0;
  2). 189 patients were randomized to ZDV+didanosine (ddI), denoted as 'ZDV+ddI' as treatment 1.
- $r = 12$ Covariates X:
  1). 4 continuous variables: age (years), weight (kg), CD4 count (cells/mm3) at baseline, and CD8 count (cells/mm3) at baseline;
  2). 8 categorical variables: hemophilia, homosexual activity, history of intravenous drug use, Karnofsky score (scale of 0-100), race (0=white, 1=non-white), gender (0=female), antiretroviral history (0=naive, 1=experienced), and symptomatic status (0=asymptomatic).
- $s = 2$ Intermediate Outcomes M: CD4 count at 20 ± 5 weeks, and CD8 count at 20 ± 5 weeks.

Real Data Analysis: ACTG 175 Dataset

Experimental sample E:

- There are $N_E = 376$ HIV-infected subjects randomized to two competitive antiretroviral regimens.
  1). 187 patients were randomized to zidovudine (ZDV) + zalcitabine (ddC), denoted as 'ZDV+ddC' as treatment 0;
  2). 189 patients were randomized to ZDV+didanosine (ddI), denoted as 'ZDV+ddI' as treatment 1.
- $r = 12$ Covariates X:
  1). 4 continuous variables: age (years), weight (kg), CD4 count (cells/mm3) at baseline, and CD8 count (cells/mm3) at baseline;
  2). 8 categorical variables: hemophilia, homosexual activity, history of intravenous drug use, Karnofsky score (scale of 0-100), race (0=white, 1=non-white), gender (0=female), antiretroviral history (0=naive, 1=experienced), and symptomatic status (0=asymptomatic).
- $s = 2$ Intermediate Outcomes M: CD4 count at 20 $\pm$ 5 weeks, and CD8 count at 20 $\pm$ 5 weeks.

Real Data Analysis: ACTG 175 Dataset

Auxiliary sample U:

- Include $N_U = 1342$ HIV-infected subjects.
- Same $r = 12$ Covariates X.
- Same $s = 2$ Intermediate Outcomes M.
- Long-Term Outcome of Interest Y: the mean CD4 count (cells/mm3) at 96 ± 5 weeks.
- It can be shown in the auxiliary data that intermediate outcomes are highly related to the long-term outcome via a linear regression of Y_U on $\{X_U, M_U\}$.

Real Data Analysis: ACTG 175 Dataset

Auxiliary sample U:

- Include $N_U = 1342$ HIV-infected subjects.
- Same $r = 12$ Covariates X.
- Same $s = 2$ Intermediate Outcomes M.
- Long-Term Outcome of Interest Y: the mean CD4 count (cells/mm3) at 96 ± 5 weeks.
- It can be shown in the auxiliary data that intermediate outcomes are highly related to the long-term outcome via a linear regression of Y_U on $\{X_U, M_U\}$.

Real Data Analysis: ACTG 175 Dataset

Auxiliary sample U:

- Include $N_U = 1342$ HIV-infected subjects.
- Same $r = 12$ Covariates X.
- Same $s = 2$ Intermediate Outcomes M.
- Long-Term Outcome of Interest Y: the mean CD4 count (cells/mm3) at 96 ± 5 weeks.
- It can be shown in the auxiliary data that intermediate outcomes are highly related to the long-term outcome via a linear regression of Y_U on $\{X_U, M_U\}$.

Real Data Analysis: ACTG 175 Dataset

Auxiliary sample U:

- Include $N_U = 1342$ HIV-infected subjects.
- Same $r = 12$ Covariates X.
- Same $s = 2$ Intermediate Outcomes M.
- Long-Term Outcome of Interest Y: the mean CD4 count (cells/mm3) at 96 ± 5 weeks.
- It can be shown in the auxiliary data that intermediate outcomes are highly related to the long-term outcome via a linear regression of Y_U on $\{X_U, M_U\}$.

Real Data Analysis: ACTG 175 Dataset

Auxiliary sample U:

- Include $N_U = 1342$ HIV-infected subjects.
- Same $r = 12$ Covariates X.
- Same $s = 2$ Intermediate Outcomes M.
- Long-Term Outcome of Interest Y: the mean CD4 count (cells/mm3) at 96 ± 5 weeks.
- It can be shown in the auxiliary data that intermediate outcomes are highly related to the long-term outcome via a linear regression of Y_U on $\{X_U, M_U\}$.

Implementation and Results

To apply the GEAR, we model the conditional mean of the long-term outcome $\mu_U(m, x)$ and the augmented term v_i in the auxiliary data via

- 'Linear': the linear regression;
- 'B-spline': the tensor-product B-splines.

Table 3: Comparison results for ACTG 175 data.

	Linear	B-spline
$\widehat{V}_{AIP}(0)$	327.8	325.5
$\widehat{V}_{AIP}(1)$	334.0	329.0
$\widehat{V}_{AIP}(\widehat{\beta}^G)$ [SD]	351.4 [10.2]	346.4 [9.6]
95% CI for $\widehat{V}_{AIP}(\widehat{\beta}^G)$	(331.4, 371.3)	(327.7, 365.1)
Assign to 'ZDV+ddC'	185	184
Assign to 'ZDV+ddI'	191	192

Thank You!