# Inferential challenges for machine learning applications in precision medicine and data-driven decision science

Michael R. Kosorok

Department of Biostatistics
Department of Statistics and Operations Research
University of North Carolina at Chapel Hill

January 16, 2018

# Acknowledgments


- Many co-authors helped:
  - Current and former students: Emily Butler, Guanhua Chen, Jingxiang Chen, Yifan Cui, Anna Kahkoska, Daniel Luckett, Yingqi Zhao, Xin Zhou, and Ruoqing Zhu
  - Other colleagues: SM Davis, H Fu, X He, U Khan, EB Laber, Y Liu, DM Maahs, E Mayer-Davis, N Mayer-Hamblett, and D Zeng
- This project was funded in part by US NIH grants P01 CA142538 and UL1 TR001111, and US NSF grant DMS-1407732, among others.

# Outline

Introduction

Outcome weighted learning

Precision medicine in mHealth

Estimation and Optimization of Composite Outcomes

Overall conclusions and future work

# Precision Medicine

- Precision medicine
  - Developing targeted treatments which leverage patient heterogeneity
  - Empirically based, scientifically rigorous, reproducible, and generalizable (i.e., will work with future patients)
  - Philosophically similar to traditional personalized medicine but with greater empirical rigor
- Scientific tools:
  - Biomedical knowledge based on current state of science
  - Data (potentially integrated across many platforms)
  - Knowledge driven vs. data driven approaches
  - Computational, mathematical and statistical tools

# Clinical focus

We want to make the best treatment decisions based on data:

- The single-decision setting:
  - A patient presents with a disease and we need to decide what treatment (or dose) to give from a list of choices
  - We want to make the best decision based on available baseline patient-level feature data (dynamic treatment regime)
- The multi-decision setting:
  - Treat patients for diseases with multiple treatment decision times based on continually accrued patient-level data
  - The best decisions take into account delayed effects
- Real time decision making in mHealth:
  - A large number of decisions need to be made in real time
  - Technical and practical challenges for implementing
- Decision making on social networks and other complex environments

# Statistics and machine learning

- What are the data analytic tasks?
  - Estimate dynamic treatment regimes (DTRs) a.k.a. Individualized treatment rules (ITRs)
  - Inference and prediction for DTRs
  - Etiology?
- What role does statistics play?
  - Estimation
  - Inference: consistency, accuracy (error bounds), confidence regions, efficiency, etc.
- How can machine learning help?
  - Provide a rich set of estimation and prediction tools
  - Perform certain data-drive tasks unusual for statistics: policy learning, reinforcement learning, inverse reinforcement learning, etc.

# Outline

Introduction

Outcome weighted learning

Precision medicine in mHealth

Estimation and Optimization of Composite Outcomes

Overall conclusions and future work

# Single decision setting

- Let $X$ be the vector of patient tailoring variables, $A$ the choice of treatment given, and $R$ the clinical outcome (with larger being better).
- An obvious approach is to first estimate the Q-function

$$Q(x, a) = E\left[R \middle| X = x, A = a\right],$$

through regression of $R$ on $(X, A)$, and invert to obtain

$$\hat{d}(x) = \underset{a}{\operatorname{argmax}}\, \hat{Q}(x, a).$$

- Potential issue: Why estimate all of $\hat{Q}(x, a)$ when focus is on $\hat{d}(x)$?

# Value function and optimal individualized treatment rule

- Let $P$ be the distribution of $(X, A, R)$, with treatments randomized via $\pi(A|X)$, and $P^d$ the distribution of $(X, A, R)$, with treatments chosen according to $d$. The value function of $d$ (Qian & Murphy, 2011) is

$$V(d) = E^d(R) = \int R dP^d = \int R \frac{dP^d}{dP} dP = E\left[\frac{I(A = d)}{\pi(A|X)} R\right].$$

- Optimal Individualized Treatment Rule:

$$d^* \in \underset{d}{\text{argmax}}\, V(d).$$

$$E(R|X, A = 1) > E(R|X, A = -1) \Rightarrow d^*(X) = 1$$
$$E(R|X, A = 1) < E(R|X, A = -1) \Rightarrow d^*(X) = -1$$

# Outcome weighted learning (OWL or O-learning)

## Optimal Individualized Treatment Rule $d^*$

Maximize the value  Minimize the risk

$$E\left[\frac{I(A = d(X))}{\pi(A|X)}R\right] \qquad E\left[\frac{I(A \neq d(X))}{\pi(A|X)}R\right]$$

- For any rule $d$, $d(X) = \text{sign}(f(X))$ for some function $f$.
- Empirical approximation to the risk function:

$$n^{-1}\sum_{i=1}^{n}\frac{R_i}{\pi(A_i|X_i)}I(A_i \neq \text{sign}(f(X_i))).$$

- Computational challenges: non-convexity and discontinuity of 0-1 loss.

# Using a support vector machine (SVM) approach

## Objective Function: Regularization Framework

$$\min_f \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{R_i}{\pi(A_i|X_i)} \phi(A_i f(X_i)) + \lambda_n \|f\|^2 \right\}. \qquad (1)$$

- $\phi(u) = (1-u)^+$ is the hinge loss surrogate, $\|f\|$ is some norm for $f$, and $\lambda_n$ controls the penalty on $f$.
- A linear decision rule: $f(X) = X^T\beta + \beta_0$, with $\|f\|$ as the Euclidean norm of $\beta$.
- Estimated individualized treatment rule:

$$\hat{d}_n = \text{sign}(\hat{f}_n(X)),$$

where $\hat{f}_n$ is the solution to (1).

# Results for O-Learning

- Can use kernel trick to extend to nonparametric decision rule (e.g., the Gaussian kernel).
- Fisher consistent, consistent, and model robust.
- Risk bounds and convergence rates similar to those observed in SVM literature (Tsybakov, 2004).
- Excellent simulation results and data analysis of Nefazodone-CBASP clinical trial (Keller et al., 2000).
- Promising performance overall (Y.Q. Zhao, et al., 2012).
- An example of a policy learning approach (see also B. Zhang, et al., 2012; Athey and Wager, 2017; others).
- Opens door to a unique application of machine learning techniques to personalized medicine.
- Not semiparametric efficient in finite-dimensional setting.

# O-Learning Extensions

- Multiple decision times (Zhao et al, 2015, *JASA*)
- Location invariance for outcome/utility (Zhou et al, 2017, *JASA*)
- More than two treatment options:
  - Ordinal treatment options (Chen et al, In press, *Biometrics*)
  - Nonordinal treatments (Rashid et al, submitted)
- Censored data (Zhao et al, 2015, *Biometrika*; Cui et al, 2017, *EJS*)
- For observational data, propensity score estimation is needed

# O-Learning and Related Extensions

- Continuous treatment options
  - Chen G, Zeng D, and Kosorok MR (2016). Personalized dose finding using outcome weighted learning (with discussion and rejoinder). *JASA* 111:1509-1547.
  - Consistency and error bounds are difficult, and inference is unclear
- V-learning for (nearly) continuous time and mHealth (Luckett et al, submitted)
- Multiple competing utilities
  - Incorporating patient preferences (Butler et al, In press, *Biometrics*)
  - Inverse reinforcement learning to infer composite utility (Luckett et al, submitted)

# Outline

Introduction

Outcome weighted learning

Precision medicine in mHealth

Estimation and Optimization of Composite Outcomes

Overall conclusions and future work

# Precision medicine in mHealth

Overall research goal:

- Develop estimation techniques (using data collected with mobile devices) for dynamic treatment regimes (which can be implemented as personalized mHealth interventions)

Motivating example: type 1 diabetes

- Understand type 1 diabetes (T1D) and how it is managed (minimizing hypo- and hyperglycemia, controling weight)
- Develop tailored mHealth interventions for T1D management

# The glucose-insulin dynamical system

A day in the life of a T1D patient:

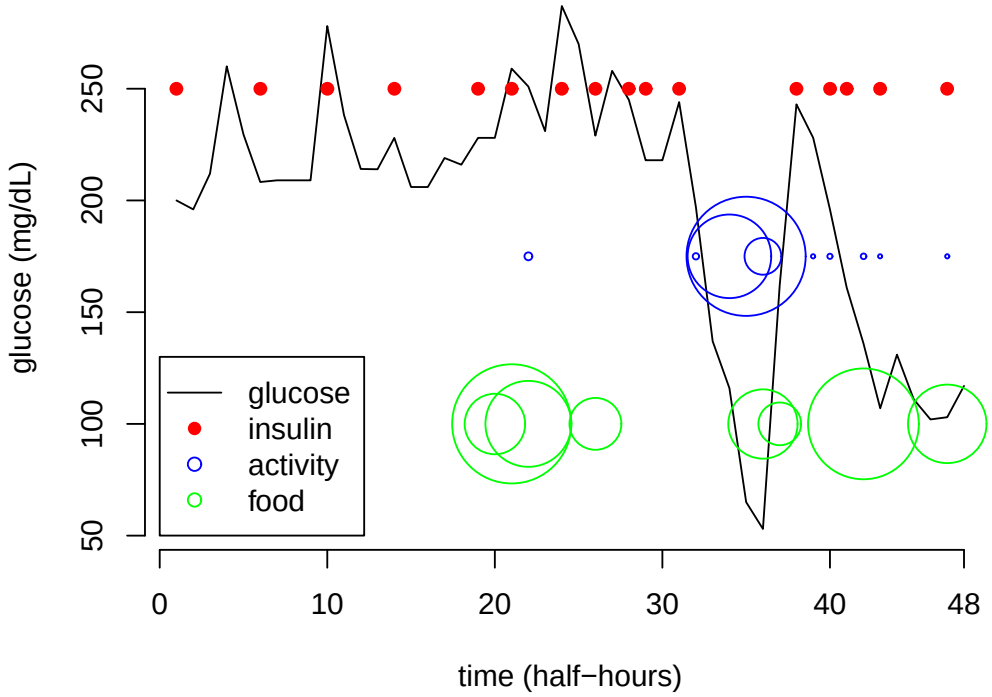


Figure 1: Plot of glucose, insulin, physical activity, and food intake.

# Mobile technology in T1D care

Mobile devices can be used to administer treatment and assist with data collection in an outpatient setting, including

- Continuous glucose monitoring
- Accelerometers to track physical activity
- Insulin pumps to administer and log injections automatically

These technologies can be incorporated using mobile phones.

# Research goals

Methodological goals:

- Estimate dynamic treatment regimes for use in mobile health
- Infinite time horizon, minimal modeling assumptions
- Observational data with minute-by-minute observations
- Online estimation to facilitate real-time decision making

Clinical goals:

- Provide patients information on the best actions to stabilize glucose
- Recommendations that are dynamic and personalized to the patient

# Conceptual framework

- We use a Markov decision process (MDP) context
- One potential approach is to use infinite horizon Q-learning (models state-value as a function of action assuming all future actions are optimal):
  - Ertefaie A (2014). Constructing dynamic treatment regimes in infinite-horizon settings. *arXiv preprint arXiv:1406.0764*.
- We developed V-learning which uses a policy learning approach (models state-value as a function of policy):
  - Luckett DJ, Laber EB, Kahkoska AR, Maahs DM, Mayer-Davis E, Kosorok MR (2016). Estimating dynamic treatment regimes in mobile health using V-learning. *arXiv preprint arXiv:1611.03531*.

# Markov decision processes (MDP's)

Assume the data consist of $n$ i.i.d. trajectories $(\mathbf{S}^1, A^1, \mathbf{S}^2, \ldots, \mathbf{S}^T, A^T, \mathbf{S}^{T+1})$ where $\mathbf{S}^t \in \mathbb{R}^p$, $A^t \in \mathcal{A}$, and there exists a known utility function $U^t = u(\mathbf{S}^{t+1}, A^t, \mathbf{S}^t)$.

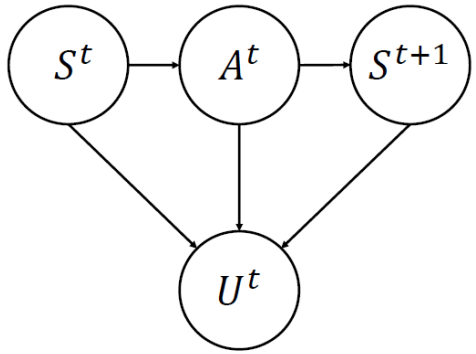


Figure 2: Graphical depiction of a Markov decision process.

Treatment regimes:

- Let $\mathcal{B}(\mathcal{A})$ be the space of distributions on $\mathcal{A}$
- A policy, $\pi$, is a function $\pi : \mathrm{dom}\, \mathbf{S}^t \to \mathcal{B}(\mathcal{A})$
- $\pi(a^t; \mathbf{s}^t)$ gives the probability of selecting $a^t \in \mathcal{A}$ when in state $\mathbf{S}^t = \mathbf{s}^t$

# The state-value function

- The state-value function is

$$V(\pi, \mathbf{s}^t) = \mathbb{E}\left\{ \sum_{k \geq 0} \gamma^k U^{*(t+k)}(\pi) \middle| \mathbf{S}^t = \mathbf{s}^t \right\}$$

  for a discount factor $\gamma \in (0, 1)$

- For a distribution, $\mathcal{R}$, define the value of $\pi$, $V_{\mathcal{R}}(\pi) = \int V(\pi, \mathbf{s}) \mathrm{d}\mathcal{R}(\mathbf{s})$

- For a class of regimes, $\Pi$, the optimal regime, $\pi_{\mathcal{R}}^{\mathrm{opt}} \in \Pi$, satisfies

$$V_{\mathcal{R}}(\pi_{\mathcal{R}}^{\mathrm{opt}}) \geq V_{\mathcal{R}}(\pi)$$

  for all $\pi \in \Pi$

# An estimating equation for $V(\pi, \mathbf{s})$

Let $\mu^t(a^t; \mathbf{s}^t) = \Pr(A^t = a^t | \mathbf{S}^t = \mathbf{s}^t)$ for each $t \geq 1$.

## Lemma

*Assume strong ignorability, consistency, and positivity. Let $\pi$ denote an arbitrary regime and $\gamma \in (0, 1)$ a discount factor. Then, provided interchange of the sum and integration is justified, the state-value function of $\pi$ at $\mathbf{s}^t$ is*

$$V(\pi, \mathbf{s}^t) = \sum_{k \geq 0} \mathbb{E}\left[\gamma^k U^{t+k} \left\{\prod_{v=0}^{k} \frac{\pi(A^{v+t}; \mathbf{S}^{v+t})}{\mu^{v+t}(A^{v+t}; \mathbf{S}^{v+t})}\right\} \middle| \mathbf{S}^t = \mathbf{s}^t\right].$$

This result will form the basis of an estimating equation for $V(\pi, \mathbf{s})$.

# An estimating equation for $V(\pi, \mathbf{s})$ (continued)

From Lemma 3.1, it follows that

$$0 = \mathbb{E}\left[\frac{\pi(A^t; \mathbf{S}^t)}{\mu^t(A^t; \mathbf{S}^t)} \left\{U^t + \gamma V(\pi, \mathbf{S}^{t+1}) - V(\pi, \mathbf{S}^t)\right\} \psi(\mathbf{S}^t)\right],$$

for any function $\psi$ (an importance-weighted version of the Bellman equation). An estimating equation for $V(\pi, \mathbf{s})$ is

$$\Lambda_n(\pi, \theta^\pi) = \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T_i} \frac{\pi(A_i^t; \mathbf{S}_i^t)}{\mu^t(A_i^t; \mathbf{S}_i^t)} \left\{U_i^t + \gamma V(\pi, \mathbf{S}_i^{t+1}; \theta^\pi) - V(\pi, \mathbf{S}_i^t; \theta^\pi)\right\} \nabla_{\theta^\pi} V(\pi, \mathbf{S}_i^t; \theta^\pi),$$

where $V(\pi, \mathbf{S}; \theta^\pi)$ is a parametric model for the state-value function.

# V-learning

Given an estimate $\widehat{\theta}_n^{\pi}$, an estimate of the value of $\pi$ under $\mathcal{R}$ is $\widehat{V}_{n,\mathcal{R}}(\pi) = \int V\left(\pi, \mathbf{s}; \widehat{\theta}_n^{\pi}\right) \mathrm{d}\mathcal{R}(\mathbf{s})$ and an estimate of the optimal policy is $\widehat{\pi}_n = \arg\max_{\pi \in \Pi} \widehat{V}_{n,\mathcal{R}}(\pi)$. Start with an initial policy, $\pi$, and repeat until convergence:

1. Estimate $\widehat{\theta}_n^{\pi}$
2. Evaluate $\widehat{V}_{n,\mathcal{R}}(\pi) = \int V\left(\pi, \mathbf{s}; \widehat{\theta}_n^{\pi}\right) \mathrm{d}\mathcal{R}(\mathbf{s})$
3. Take a step to maximize $\widehat{V}_{n,\mathcal{R}}(\pi)$ over a class of policies

# Summary for V-Learning

- Features of V-learning include
  - Flexibility in choosing a model for $V(\pi, \mathbf{s}; \theta^{\pi})$
  - Online estimation, randomized decision rules
  - Flexibility in specifying reference distribution
  - Parametric value estimates
- A tailored treatment regime delivered through mobile devices may help to reduce hypo- and hyperglycemia in T1D patients

# Asymptotic Inference

- We obtain uniform asymptotic normality for key parameters and predictions
- Main technical tools:
  - Donsker theorem for $\beta$-mixing stationary processes based on bracketing entropy (Dedecker and Louhichi, 2002)
  - New bracketing entropy preservation results for products of function classes
- Issue: Need Donsker theorems for non-stationary processes for certain types of online V-learning

# Outline

Introduction

Outcome weighted learning

Precision medicine in mHealth

Estimation and Optimization of Composite Outcomes

Overall conclusions and future work

# Precision medicine revisited

- Patients can exhibit significant heterogeneity in response to treatment
- Outcomes can be improved by tailoring treatment to individuals
- Standard components:
  - An outcome to optimize
  - A set of treatment options
  - A set of tailoring variables
- The goal is to estimate a decision rule for treatment to optimize the outcome in a population
- How do we handle the case where there are multiple outcomes?

# Motivating example: bipolar disorder

- The Systematic Treatment Enhancement Program for Bipolar Disorder Standard Care Pathway (STEP-BD SCP)
- Characterized by episodes of depression and mania
- Anti-depressants can be used to treat depressive episodes
- Anti-depressants may induce manic episodes
- An example of precision medicine: determine which patients will benefit from anti-depressants
- Clinical decision making needs to balance the trade-off between depression and mania

# Notation

- $\mathbf{X} \in \mathcal{X} \subseteq \mathbb{R}^p$ are tailoring variables
- $A \in \{-1, 1\}$ is treatment
- $Y$ and $Z$ are two real-valued outcomes with higher values preferable
- $Q_Y(\mathbf{x}, a) = \mathbb{E}\{Y|\mathbf{X} = \mathbf{x}, A = a\}$ is the mean of $Y$ given $\mathbf{X}$ and $A$, with $R_Y(\mathbf{x}) = Q_Y(\mathbf{x}, 1) - Q_Y(\mathbf{x}, -1)$
- $d_Y^{\text{opt}}(\mathbf{x}) = \arg\max_{a \in \{-1,1\}} Q_Y(\mathbf{x}, a) = \text{sign}(R_Y(\mathbf{x}))$ is the decision to maximize $Y$
- $Q_Z$, $R_Z$ and $d_Z^{\text{opt}}$ are defined similarly

# Utility functions

- If both $Y$ and $Z$ are relevant, neither $d_Y^{\text{opt}}$ nor $d_Z^{\text{opt}}$ may be acceptable
- Define the composite outcome $U = u(Y, Z)$ for a utility function, $u$
- Define $Q_U(\mathbf{x}, a) = \mathbb{E}\{U|\mathbf{X} = \mathbf{x}, A = a\}$, $R_U(\mathbf{x}) = Q_U(\mathbf{x}, 1) - Q_U(\mathbf{x}, -1)$, and

$$d_U^{\text{opt}}(\mathbf{x}) = \arg\max_{a \in \{-1,1\}} Q_U(\mathbf{x}, a) = \text{sign}(R_U(\mathbf{x}))$$

- Assume $u(Y, Z; \omega) = \omega Y + (1 - \omega)Z$; we will refer to $Q_\omega$, $R_\omega$, and $d_\omega^{\text{opt}}$
- For a broad class of utility functions, $d_U^{\text{opt}}$ is equivalent to $d_\omega^{\text{opt}}$ for some $\omega \in [0, 1]$ (Butler et al., 2017)

# Pseudo-likelihood estimation of utility functions

- Assume there exists a true utility function defined by $\omega_0$ such that observed decisions were made with the intent of maximizing $U = u(Y, Z; \omega_0)$
- Assume that

$$\Pr\left\{A = d^{\text{opt}}_{\omega_0}(\mathbf{X})\right\} = \text{expit}(\mathbf{X}^\intercal \beta_0)$$

  for some $\beta_0 \in \mathbb{R}^p$
- The likelihood for $(\omega, \beta)$ is

$$\mathcal{L}_n(\omega, \beta) \propto \prod_{i=1}^{n} \frac{\exp\left[\mathbf{X}_i^\intercal \beta 1\left\{A_i = d^{\text{opt}}_{\omega}(\mathbf{X}_i)\right\}\right]}{1 + \exp\left(\mathbf{X}_i^\intercal \beta\right)},$$

  which can be used to estimate the true utility function and the probability that any patient would be treated optimally in standard practice

# Pseudo-likelihood estimation (continued)

- The likelihood for $(\omega, \beta)$ depends on the unknown function $d_\omega^{\text{opt}}$
- Let $\widehat{Q}_{Y,n}$ and $\widehat{Q}_{Z,n}$ be estimators for $Q_Y$ and $Q_Z$, etc.
- For any $\omega \in [0,1]$, let

$$\widehat{Q}_{\omega,n}(\mathbf{x}, a) = \omega \widehat{Q}_{Y,n}(\mathbf{x}, a) + (1-\omega)\widehat{Q}_{Z,n}(\mathbf{x}, a),$$

$$\widehat{R}_{\omega,n}(\mathbf{x}) = \omega \widehat{R}_{Y,n}(\mathbf{x}) + (1-\omega)\widehat{R}_{Z,n}(\mathbf{x}), \text{ and}$$

$$\widehat{d}_{\omega,n}(\mathbf{x}) = \arg\max_{a \in \{-1,1\}} \widehat{Q}_{\omega,n}(\mathbf{x}, a) = \text{sign}(\widehat{R}_{\omega,n}(\mathbf{x}))$$

- We can replace $d_\omega^{\text{opt}}$ with $\widehat{d}_{\omega,n}$ to obtain the pseudo-likelihood

$$\widehat{\mathcal{L}}_n(\omega, \beta) \propto \prod_{i=1}^{n} \frac{\exp\left[\mathbf{X}_i^\intercal \beta 1\left\{A_i = \widehat{d}_{\omega,n}(\mathbf{X}_i)\right\}\right]}{1 + \exp\left(\mathbf{X}_i^\intercal \beta\right)}$$

# Patient-specific utility functions

- Let $\theta \in \mathbb{R}^d$ and assume

$$u(Y, Z; \mathbf{X}, \theta) = m(\mathbf{X}; \theta) Y + \{1 - m(\mathbf{X}; \theta)\} Z,$$

  where $m \mapsto (0, 1)$ is continuously differentiable in $\theta$

- Define $\widehat{d}_{\theta,n}$ analogously to $\widehat{d}_{\omega,n}$, etc.

- The pseudo-likelihood is

$$\widehat{\mathcal{L}}_n(\theta, \beta) \propto \prod_{i=1}^{n} \frac{\exp\left[\mathbf{X}_i^\intercal \beta 1\left\{A_i = \widehat{d}_{\theta,n}(\mathbf{X}_i)\right\}\right]}{1 + \exp\left(\mathbf{X}_i^\intercal \beta\right)}$$

# Asymptotic Inference

We need some basic assumptions and definitions, including:

- $\sqrt{n}\left[\widehat{R}_{Y,n}(\mathbf{x}) - R_Y(\mathbf{x})\right] = \phi_Y^T(\mathbf{x}) n^{-1/2} \sum_{i=1}^n \psi_{iY} + o_P(1)$, where $o_P(1)$ is uniform over $\mathbf{x}$, the i.i.d. influence functions $\psi_{iY} \in \mathbb{R}^{q_1}$, and $\phi_Y$ are basis functions
- $\sqrt{n}\left[\widehat{R}_{Z,n}(\mathbf{x}) - R_Z(\mathbf{x})\right] = \phi_Z^T(\mathbf{x}) n^{-1/2} \sum_{i=1}^n \psi_{iZ} + o_P(1)$, similarly, and $\psi_{iZ} \in \mathbb{R}^{q_2}$
- $P_\beta(\mathbf{x}) = \text{expit}(\mathbf{x}^T\beta)$, $\psi_{iA} = \mathbf{X}_i(A_i - P_{\beta_0}(\mathbf{X}_i))$ and $I_0 = P\left[\mathbf{X}\mathbf{X}^T P_{\beta_0}(\mathbf{X})(1 - P_{\beta_0}(\mathbf{X}))\right]$, where $P$ is the expectation over $\mathbf{X}$
- $\widehat{D}_{\theta,n}(\mathbf{x}) = m(\mathbf{x};\theta)\widehat{R}_{Y,n}(\mathbf{x}) + (1 - m(\mathbf{x};\theta))\widehat{R}_{Z,n}(\mathbf{x})$
- $D_\theta(\mathbf{x}) = m(\mathbf{x};\theta)R_Y(\mathbf{x}) + (1 - m(\mathbf{x};\theta))R_Z(\mathbf{x})$
- The density of $D_{\theta_0}(\mathbf{X})$ at zero is $0 < f_0 < \infty$.

# Asymptotic Inference

Assumptions and definitions, continued:

- Assume

$$\Sigma_0 = E\begin{pmatrix}\psi_{1Y}\\ \psi_{1Z}\\ \psi_{1A}\end{pmatrix}^{\otimes 2} = \begin{pmatrix}\Sigma_{YY} & \Sigma_{YZ} & \Sigma_{YA}\\ \Sigma_{YZ}^T & \Sigma_{ZZ} & \Sigma_{ZA}\\ \Sigma_{YA}^T & \Sigma_{ZA}^T & \Sigma_{AA}\end{pmatrix}$$

  is positive definite (note that $\Sigma_{AA} = I_0$)
- Let $a_Y(\mathbf{x}) = m(\mathbf{x};\theta_0)R_Y(\mathbf{x})\phi_Y(\mathbf{x})$, $a_Z(\mathbf{x}) = (1 - m(\mathbf{x};\theta_0))R_Z(\mathbf{x})\phi_Z(\mathbf{x})$, $b(\mathbf{x}) = (R_Y(\mathbf{x}) - R_Z(\mathbf{x}))\dot{m}_{\theta_0}(\mathbf{x})$, and $c(\mathbf{x}) = \mathbf{x}\,(2P_{\beta_0}(\mathbf{x}) - 1)$, where $\dot{m}_\theta = \partial m/(\partial\theta)$
- For any $z_Y \in \mathbb{R}^{q_1}$, $z_Z \in \mathbb{R}^{q_2}$, $u \in \mathbb{R}^d$, define the function $(z_Y, z_Z, u) \mapsto k_0(z_y, z_Z, u) =$

$$P\left[c(\mathbf{X})\left|a_Y(\mathbf{X})^T z_Y + a_Z(\mathbf{X})^T z_Z + b(\mathbf{X})^T u\right| \,|D_{\theta_0}(\mathbf{X}) = 0\right] f_0$$

# Asymptotic Inference

**Theorem**

*Under regularity conditions, the pseudo-likelihood maximizers $\hat{\beta}_n$ and $\hat{\theta}_n$ satisfy*

$$\sqrt{n}\begin{pmatrix} \hat{\beta}_n - \beta_0 \\ \hat{\theta}_n - \theta_0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} I_0^{-1}\left[Z_A - k_0(Z_Y, Z_Z, U)\right] \\ U \end{pmatrix} = \begin{pmatrix} B \\ U \end{pmatrix},$$

*where* $U = \text{argmin}_u\, \beta_0^T k_0(Z_Y, Z_Z, u)$, *and*

$$\begin{pmatrix} Z_Y \\ Z_Z \\ Z_A \end{pmatrix} \sim N(0, \Sigma_0).$$

*A certain semiparametric bootstrap is also consistent in probability.*

# Asymptotic Inference

Main technical tools:

- The Argmax theorem
- The following for the bootstrap:

**Theorem**

*Let $H$ be compact with respect to a metric $d$ and $\mathcal{F} \subset C[H]$ be compact with respect to $\|\cdot\|_H$. For each $f \in \mathcal{F}$, let $u(f) = \text{argmax}_{u \in H} f(u)$, where we arbitrarily choose a value if nonunique. Suppose also that there exists an $\mathcal{F}_1 \subset \mathcal{F}$ such that each $f \in \mathcal{F}_1$ has a unique maximum. Then*

$$\lim_{\delta \downarrow 0} \sup_{f \in \mathcal{F}_1} \sup_{g \in \mathcal{F}: \|f-g\|_H < \delta} d(u(f), u(g)) = 0.$$

# Outline

Introduction

Outcome weighted learning

Precision medicine in mHealth

Estimation and Optimization of Composite Outcomes

Overall conclusions and future work

# Overall Conclusions and Future Work

- This is an exciting time for precision medicine at the confluence of machine learning and statistics.
- There are numerous open questions.
- Inference can be challenging and nonstandard.
- Consistency, or zero order inference, is often an important first step.
- This work is part of the emergence of a new (or renewed) discipline focused on data driven decision making and precision medicine and has many connections in many quantitative and nonquantitative disciplines.