An Overview of Dynamic Treatment Regimes

Anastasios (Butch) Tsiatis
Department of Statistics
North Carolina State University

http://www.stat.ncsu.edu/∼tsiatis

Outline

1. Definition of a dynamic treatment regime
2. Casual inference and potential outcomes
3. Assumptions necessary (the no unmeasured confounders or sequential randomization assumption)
4. G-estimation
5. Inverse probability weighting
6. Strategies for modeling
7. Challenges

Dynamic treatment regimes

Treating chronic disease is often complex involving many treatment decisions over time possibly depending on intervening information on the health of the individual. For example, when a physician treats a patient with HIV disease

- Do we start the patient on antiretroviral drugs or wait?
- Which drugs?
- How long do we keep them on drug?
- Do we switch drug?
- Should we use CD4 counts, viral load, health status indicators to aid us in the above decisions?

The goal of a dynamic treatment regime

The overall goal is to give treatment to patients over time that will result in as favorable a clinical outcome as possible. For example, in HIV disease we may be interested ultimately in outcomes such as

- Length of life
- Quality of life
- Long term control of viral load

Current practice

Clinical practice: Treatment of *chronic disease* is an *ongoing process*

- Clinicians *manage* a patient's illness
- Clinicians routinely *adjust*, *change*, *add*, or *discontinue* treatment based on *progress*, *side effects*, *patient burden*, *compliance*, etc.
- Decisions based on *clinical judgment*, *patient preference*, *practice guidelines*
- Goal of "*individualizing*" treatment to the patient to provide the *best care*

Current practice

- In general, we do not study treatment as an ongoing process; rather we focus on each decision point separately (myopic strategy)
- One problem with this: delayed effects
- A treatment early on may give initially a good response but may render subsequent treatments ineffective
- We argue that one should study the entire treatment decision process over time; i.e., the entire treatment regime

How to assess treatment regimes

We need to be able to assess and compare how different treatment decisions affect outcomes of interest. In order to do this we need data on patients which include

- treatments they received over time
- covariate information (baseline characteristics, laboratory measurements, side-effects, etc.) which may affect treatment decisions made over time
- outcomes

Also, in order to make comparisons, we need data where patients receive different treatment combinations that we want to compare. This can come about by

- Controlled intervention studies (sequentially randomized studies)
- Observational data
- Combination

Notation and formal definitions

- For simplicity, we assume that treatment decisions are made at discrete times $t_0, \ldots, t_K$, where t_0 denotes baseline
- The data observed on individuals are time ordered given as $L_0, A_0, L_1, A_1, \ldots, L_K, A_K, Y$, where
  - L_0 denotes baseline covariates measured prior to t_0
  - A_0 denotes treatment assigned at time t_0 among the set of potential treatments $\mathcal{A}_0$
  - L_1 denotes covariate information collected between times t_0 and t_1
  - A_1 denotes treatment assigned at time t_1 among the set of potential treatments $\mathcal{A}_1$
  - So forth for $L_j, A_j,\ j = 2, \ldots, K$
  - Y denotes the outcome of interest, measured after time t_K

Notation and formal definitions

- The overbar notation will be used to denote the history of time-dependent variables

$$\bar{L}_j = (L_0, \ldots, L_j) \text{ and } \bar{A}_j = (A_0, \ldots, A_j), \quad j = 1, \ldots, K$$

- We will often write $\bar{A}_K$ and $\bar{L}_K$ as $\bar{A}$ and $\bar{L}$ for brevity
- **Goal** is to study how treatment decisions $A_0, \ldots, A_K$ affect the response Y
- The point of view taken here is that the treatment assignments $A_0, \ldots, A_K$ are manipulable; i.e., treatment A_j can be chosen by the physician to be any of the choices $a_j \in \mathcal{A}_j$ but the covariates L_j are not manipulable and that the values of L_j may be affected by the treatment choices.

Treatment regime

A **treatment regime** is an algorithm which dictates how each patient in the population is treated possibly based on intervening covariate information. Specifically,

- A treatment regime is a function $g : \bar{\mathcal{L}}_K \to \bar{\mathcal{A}}_K$; where, for every $j = 0, \ldots, K$ and every $\bar{\ell}_j \in \bar{\mathcal{L}}_j$, a realization of $\bar{L}_j$,

$$g(t_j, \bar{\ell}_j) = a_j \in \mathcal{A}_j$$

- For example, if L_j are a set of covariates measured at the time interval preceding t_j (that include CD4 counts) for HIV patients and $a_j = (1, 0)$ denotes whether or not to give antiretroviral therapy at time t_j, then a treatment regime may be to give therapy if a patient's CD4 count is less than 200 and not to give it when it is greater than 200. I.e., $g(t_j, \bar{\ell}_j) = I(\text{CD4}_j \leq 200)$

Treatment regime

- Note that for the treatment regime g every patient in the population is assigned one and only one sequence of treatments at times $t_0, \ldots, t_K$
- If $g(t_j, \bar{\ell}_j) = a_j$ is independent of the time-dependent covariates $\bar{\ell}_j$ for all $j = 0, \ldots, K$, then g is referred to as a **non-dynamic treatment regime**; otherwise, it is a **dynamic treatment regime**.

Goal

- What we would like to know is what would be the probability distribution of the response if we subjected the entire population to treatment regime g for various $g \in \mathcal{G}$. We can use this to deduce the mean response for a given g and compare the mean response for different g.
- How can we estimate the probability distribution for various g?
  - Design a controlled intervention study?
  - Use data from an observational study?

Causal inference and potential outcomes

- Consider the combination of treatment assignments $\bar{a}_K = (a_0, \ldots, a_K) \in \bar{\mathcal{A}}_K$ and denote by $Y^*(\bar{a}_K)$ to be the response of a randomly selected individual in our population if he/she hypothetically received treatment a_0 at time t_0, a_1 at time t_1, . . . , a_K at time t_K.
- Similarly, denote by $L_j^*(\bar{a}_{j-1})$ to be the values of the time dependent covariates between times t_{j-1} and t_j of a randomly selected individual in our population if he/she hypothetically received treatment a_0 at time t_0, . . . , a_{j-1} at time t_{j-1}
- The random variables $Y^*(\bar{a}_K)$ and $L_j^*(\bar{a}_{j-1})$ are referred to as **potential outcomes** or **counterfactuals**

Causal inference and potential outcomes

- The set of all potential outcomes will be denoted by

$$W = \left[\{L_j^*(\bar{a}_{j-1}), j = 0, \ldots, K\}, Y^*(\bar{a}_K) \text{ for all } \bar{a}_K \in \bar{\mathcal{A}}_K \right]$$

- Although not directly observable, we may imagine that there is some population probability distribution of W

Potential outcomes for treatment regime

- Thus for any treatment regime $g \in \mathcal{G}$, we can define the corresponding potential outcomes

$$\{L_0^*(g), \ldots, L_K^*(g), Y^*(g)\}$$

as

$$\begin{aligned}
L_0^*(g) &= L_0 \\
L_1^*(g) &= L_1^*\{g(t_0, L_0)\} \\
L_2^*(g) &= L_2^*\{g(t_0, L_0), g(t_1, \bar{L}_1^*(g))\} \\
&\cdot \\
&\cdot \\
&\cdot \\
L_K^*(g) &= L_K^*\{g(t_0, L_0), \ldots, g(t_{K-1}, \bar{L}_{K-1}^*(g))\} \\
Y^*(g) &= Y^*\{g(t_0, L_0), \ldots, g(t_K, \bar{L}_K^*(g))\}
\end{aligned}$$

Potential outcomes for treatment regime

- Note that $Y^*(g)$ for all $g \in \mathcal{G}$ is a function of W
- Hence, the probability distribution of $Y^*(g)$ is induced from the probability distribution of W
- It is this probability distribution (or aspects of this distribution) for difference treatment regimes g that is of primary inferential interest, e.g., the mean of $Y^*(g)$; $(E\{Y^*(g)\})$

Observed data

- In contrast to the potential outcomes, the observed data from either a controlled intervention study or an observational study can be represented as a sample of data $O_i, i = 1, \ldots, n$ (assumed iid) where $O_i = (L_{0i}, A_{0i}, \ldots, L_{Ki}, A_{Ki}, Y_i)$
- The question then becomes if whether and under what circumstances can we derive the probability distribution of the potential outcome $Y^*(g)$ for $g \in \mathcal{G}$ from the probability distribution of the observed data O.

Assumptions

- The first assumption is a **consistency assumption** where we assume that

$$L_j = \sum_{\bar{a}_{j-1} \in \bar{\mathcal{A}}_{j-1}} L_j^*(\bar{a}_{j-1}) I(\bar{A}_{j-1} = \bar{a}_{j-1}), j = 1, \ldots, K$$

$$Y = \sum_{\bar{a}_K \in \bar{\mathcal{A}}_K} Y^*(\bar{a}) I(\bar{A} = \bar{a})$$

- That is, the response that is observed corresponds to the potential response (outcome) for the observed treatment actually received. (No interference among patients)

Assumptions

- The second and critical assumption is referred to as the

No unmeasured confounders assumption

or the

sequential randomization assumption (SRA)

$$\{W \perp\!\!\!\perp A_j | (\bar{A}_{j-1}, \bar{L}_j) \text{ for all } j = 0, \ldots, K\}$$

- SRA assumes that, conditional on the observed time-dependent covariate history and time-dependent treatment history up to time t_j, the treatment assignment at time t_j, A_j, is made independently of the potential outcomes of the individual

Sequential randomization assumption

- Suppose we conducted the following experiment: At time t_j we randomized patients to treatment $a_j \in \mathcal{A}_j$ with prespecified probabilities

$$P(A_j = a_j | \bar{L}_j = \bar{\ell}_j, \bar{A}_{j-1} = \bar{a}_{j-1}) = \pi_j(a_j | \bar{\ell}_j, \bar{a}_{j-1}), j = 0, \ldots, K$$

that may depend on prior covariate-treatment history, where

$$\sum_{a_j \in \mathcal{A}_j} \pi_j(a_j | \bar{\ell}_j, \bar{a}_{j-1}) = 1 \text{ for all } \bar{\ell}_j \text{ and } \bar{a}_{j-1}$$

- In a controlled intervention study such an experiment can actually be carried out and the SRA assumption automatically holds and the π_j's are known.

- In an observational study this is an assumption that must be evaluated on subject matter grounds and the π_j's must be estimated

Finding the probability distribution of $Y^*(g)$

- There are two general approaches for showing how the probability distribution of the potential response for treatment regime g; i.e., $P\{Y^*(g) = y\}$ can be derived from the probability distribution of the observed data O. These are
  - G-computation algorithm
  - Inverse probability weighting

G-computation algorithm

- The joint density of $\{L_0, L_1^*(g), \ldots, L_K^*(g), Y^*(g)\}$ can be obtained as

$$
\begin{aligned}
p_{L_0,L_1^*(g),\ldots,L_K^*(g),Y^*(g)}&(\ell_0,\ldots,\ell_K,y) \qquad (1)\\
&= p_{L_0}(\ell_0)\\
&\times p_{L_1|L_0,A_0}\{\ell_1|\ell_0,\bar{g}_0(\ell_0)\}\\
&\times p_{L_2|\bar{L}_1,\bar{A}_1}\{\ell_2|\bar{\ell}_1,\bar{g}_1(\bar{\ell}_1)\}\\
&\times \ldots \times p_{L_K|\bar{L}_{K-1},\bar{A}_{K-1}}\{\ell_K|\bar{\ell}_{K-1},\bar{g}_{K-1}(\bar{\ell}_{K-1})\}\\
&\times p_{Y|\bar{L}_K,\bar{A}_K}\{y|\bar{\ell}_K,\bar{g}_K(\bar{\ell}_K)\}
\end{aligned}
$$

- The probability distribution of $Y^*(g)$ can then be obtained by integrating out $\ell_0,\ldots,\ell_K$

G-computation algorithm

- Notice that the joint distribution of the observed data $O = (L_0, A_0, \ldots, L_K, A_K, Y)$ can be expressed as

$$
\begin{aligned}
p_O(o) \quad &= \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (2) \\
& \quad p_{L_0}(\ell_0) \\
&\times \quad p_{A_0|L_0}(a_0|\ell_0) \\
&\times \quad p_{L_1|L_0,A_0}(\ell_1|\ell_0, a_0) \\
&\times \quad \ldots \times p_{A_K|\bar{L}_K,\bar{A}_{K-1}}(a_K|\bar{\ell}_K, \bar{a}_{K-1}) \\
&\times \quad p_{Y|\bar{L}_K,\bar{A}_K}(y|\bar{\ell}_K, \bar{a}_K)
\end{aligned}
$$

- The G-computation algorithm uses only the terms in blue

Inverse probability weighting

- If everyone in a study were assigned treatment according to a treatment regime g then with a sample of data we could estimate $P\{Y^*(g) = y\}$ by simply taking the empirical average; i.e., $n^{-1} \sum_{i=1}^{n} I\{Y_i^*(g) = y\}$.
- However only a subset of individuals have treatment assignment consistent with the treatment regime g. The others were assigned other treatments.
- The idea is to take a weighted average of the response among individuals consistent with receiving treatment regime g weighted by the inverse of the probability that they received such a regime.
- This inverse probability weighting thus allows the individual who received treatment according to the regime g to not only represent themselves but also similar individuals who did not receive this treatment

Inverse probability weighting

- We remind that

$$I\{\bar{A}_K = \bar{g}_K(\bar{L}_K)\} = \prod_{j=0}^{K} I\{A_j = g(t_j, \bar{L}_j\}$$

and that

$$P(A_j = a_j | \bar{L}_j, \bar{A}_{j-1}) = \pi_j(a_j | \bar{L}_j, \bar{A}_{j-1})$$

- In which case we use

$$\frac{I\{\bar{A}_K = \bar{g}_K(\bar{L}_K)\} I(Y = y)}{\prod_{j=0}^{K} \pi_j\{g(t_j, \bar{L}_j) | \bar{L}_j, \bar{g}_{j-1}(\bar{L}_{j-1})\}} \tag{3}$$

- An estimator for $P\{Y^*(g) = y\}$ is given by

$$n^{-1} \sum_{i=1}^{n} \frac{I\{\bar{A}_{K,i} = \bar{g}_K(\bar{L}_{K,i})\} I(Y_i = y)}{\prod_{j=0}^{K} \pi_j\{g(t_j, \bar{L}_{j,i}) | \bar{L}_{j,i}, \bar{g}_{j-1}(\bar{L}_{j-1,i})\}}$$

Complexity of treatment regimes

- Let the number of different treatments at time t_j be k_j for $j = 0, \ldots, K$
- The number of non-dynamic treatment regimes can be as large as $k_0 \times \ldots \times k_K$.
- Let the number of levels of the covariate L_j be u_j for $j = 0, \ldots, K$
- The number of treatment regimes is

$$k_0^{(u_0)} \times \ldots \times k_j^{(u_0 \times \ldots \times u_j)} \times \ldots \times k_K^{(u_0 \times \ldots \times u_K)}$$

- The total number of observed covariate-treatment combinations is

$$(k_0 \times \ldots \times k_K) \times (u_0 \times \ldots \times u_K)$$

- This makes modeling the relationship of response as a function of dynamic treatment regimes with easily interpretable parameters a difficult task

Structural nested models

- The structural nested model (SNM) models the effect that "a blip of treatment a_j would have on response as compared to $a_j = 0$ on a group of individuals who at time t_j are prognostically similar; (i.e., have the same treatment-covariate history) for all $j = 0, \ldots, K$
- For a structural nested mean model we model

$$\gamma(\bar{\ell}_j, \bar{a}_j, \psi) = E\{Y^*(\bar{a}_j, 0) - Y^*(\bar{a}_{j-1}, 0) | \bar{L}_j = \bar{\ell}_j, \bar{A}_j = \bar{a}_j\}$$

where $\gamma(\bar{\ell}_j, \bar{a}_{j-1}, a_j = 0, \psi) = 0$ for all $\bar{\ell}_j, \bar{a}_{j-1}, j = 0, \ldots, K$ and all ψ and $\gamma(\bar{\ell}_j, \bar{a}_j, \psi) = 0$ when $\psi = 0$

- E.g., $\gamma(\bar{\ell}_j, \bar{a}_j, \psi) = \psi_1 a_j + \psi_2 a_j a_{j-1} + \psi_3 a_j w_j(\bar{\ell}_j)$

Marginal structural models

- One builds models only for non-dynamic treatment regimes

$$E\{Y^*(\bar{a}_K)\} = \mu(\bar{a}_K, \psi)$$

  e.g., $E\{Y^*(\bar{a}_K)\} = \psi_0 + \psi_1 cum(\bar{a}_K)$

- Estimation of ψ uses inverse probability weighting
- Useful for comparing non-dynamic treatment regimes
- Not very useful if interest is in dynamic treatment regimes

Optimal dynamic treatment regime

- Here the goal is to find the regime g that maximizes $E\{Y^*(g)\}$
- The optimal regime is found through dynamic programming and backward induction
- Murphy and Robins have proposed methods for estimating optimal dynamic treatment regimes

Q-learning

- Using backwards induction we first start at the last time point and build a model for

$$E(Y|\bar{L}_K, \bar{A}_K) = \mu_K(\bar{L}_K, \bar{A}_K, \psi_K),$$

where ψ_K is estimated using the data $(Y_i, \bar{L}_{Ki}, \bar{A}_{Ki}); i = 1, \ldots, n$.

- Define

$$g_K^{\text{opt}}(\bar{L}_K, \bar{A}_{K-1}) = \text{arg max}_{a_K} \mu_K(\bar{L}_K, \bar{A}_{K-1}, a_K, \hat{\psi}_K)$$

and

$$\mu_K^{\text{opt}}(\bar{L}_K, \bar{A}_{K-1}, \hat{\psi}_K) = \text{max}_{a_K} \mu_K(\bar{L}_K, \bar{A}_{K-1}, a_K, \hat{\psi}_K)$$

Q-learning

- We next develop a model for

$$E\{\mu_K^{\text{opt}}(\bar{L}_K, \bar{A}_{K-1}) | \bar{L}_{K-1}, \bar{A}_{K-1}\} = \mu_{K-1}(\bar{L}_{K-1}, \bar{A}_{K-1}, \psi_{K-1}),$$

where ψ_{K-1} is estimated using data

$$\left(\mu_K^{\text{opt}}(\bar{L}_{K,i}, \bar{A}_{K-1,i}, \hat{\psi}_K), \bar{L}_{K-1,i}, \bar{A}_{K-1,i}\right); i = 1, \ldots, n.$$

- Define

$$g_{K-1}^{\text{opt}}(\bar{L}_{K-1}, \bar{A}_{K-2}) = \text{arg max}_{a_{K-1}} \mu_{K-1}(\bar{L}_{K-1}, \bar{A}_{K-2}, a_{K-1}, \hat{\psi}_{K-1})$$

and

$$\mu_{K-1}^{\text{opt}}(\bar{L}_{K-1}, \bar{A}_{K-2}, \hat{\psi}_{K-1}) = \text{max}_{a_{K-1}} \mu_{K-1}(\bar{L}_{K-1}, \bar{A}_{K-2}, a_{K-1}, \hat{\psi}_{K-1})$$

- Continue iteratively until first time point ...

Challenges

Although the conceptual theory for dynamic treatment regimes has been laid out there remain many challenges

- Even if the probability distributions were known, dynamic programming is very difficult. For our problems, we need to combine dynamic programming together with statistical modeling for high-dimensional problems (We hope to learn from other disciplines)

Challenges

- For sequentially randomized controlled intervention studies
  - How many time points can sequential randomization be feasibly implemented?
  - What sample sizes are necessary?
- For observational studies
  - How can we ensure that sufficient information has been collected to make the no unmeasured confounders (SRA) assumption tenable?
  - Sensitivity analysis?