Section V: Extension

- Survival outcomes with censoring
- Multicategory Treatment
- Observational Study
- Dynamic Treatment Regimes

Survival outcomes with censoring

Survival Outcomes with Censoring

- Interested in time-to-event outcome.
- Observe independently and identically distributed training data $(X_i, A_i, D_i, \Omega_i), i = 1, \ldots, n$.

  X: baseline variables, $X \in \mathbb{R}^p$,
  A: binary treatment options, $A \in \{0, 1\}$,
  D: observed event time.
  Ω: censoring indicator $\Omega_i = I(T_i \leq C_i)$.

- $D = \min(T, C)$: T survival time, C censoring time.
- Randomized study with known randomization probability of the treatment.

Survival Outcomes with Censoring

- Two possible objectives
  - Maximize expected survival time
  - Maximize the probability of surviving beyond a landmark time.

Cox Regression

- Classic approach: Cox proportional hazard regression
- Cox regression: cox proportional hazards model with treatment-by-covariate interactions
- Estimates for $d^*(x)$ can be derived based on the predicted outcomes.

Cox (*JRSSB*, 1972)

Regression Modeling Approach

- Inverse probability of censoring weighted (IPW) Q-learning: solve for

$$\underset{\theta}{\operatorname{argmin}} \sum_{i=1}^{n} \{D_i - \Phi(X_i, A_i)\theta\}^2 \frac{\Omega_i}{\widehat{S}_C(D_i|A_i, X_i)},$$

  where $\widehat{S}_C(D|A, X)$ is the estimated conditional survival function of C given (A, X).

- $E(T|A, X)$ is modeled using $\Phi(X, A)\theta$, where $\Phi(X, A)$ is the basis, e.g., $\Phi(X, A) = (1, X, A, XA)$.

- The estimated optimal decision rule is

$$\widehat{d}(x) = \underset{a \in \{0,1\}}{\operatorname{argmin}} \Phi(x, a)\widehat{\theta}.$$

- Regularization can be applied.

Goldberg & Kosorok (*Annals of Stat.*, 2012)

Outcome Weighted Learning Approach

- Modification of outcome weighted learning
- Minimize

$$n^{-1} \sum_{i=1}^{n} \frac{\Omega_i D_i}{\hat{S}_C(D_i|A_i, X_i)} \frac{\phi\{Af(X_i)\}}{P\{d(X)|X\}} + \lambda_n \|f\|^2,$$

- Replace D_i by $\Omega_i D_i / \hat{S}_C(D_i|A_i, X_i)$ in the outcome weighted learning for uncensored data.

Zhao et al (*Biometrika*, 2015)

Accommodating a Time-to-event Outcome

Let T be the event time. Let $D = I(T < t_0)$ be an indicator that the event occurs before a landmark time t_0.

The methods above can be used to model D, with the following modifications:

1. Estimate $E(D|A, X)$ using a regression method suitable for time-to-event outcomes (e.g. Cox regression). This may need to be paired with a baseline hazard estimate.
2. Estimate performance measures empirically using inverse-probability-of-censoring weights. (Model-based estimates require no modification.)
3. Consider performing analyses for different choices of t_0; typically X more weakly predicts treatment effect for larger t_0.

Multicategory Treatment

Multicategory Treatment

- Multiple treatments of interest, $A = 0, 1, \ldots, K$, e.g., $K = 2$ in depression data

- $d^*(x) = \text{argmin}_{k=0,\ldots,K}\, \mu(k, x)$.

- Posit a regression model

$$E(D|A, X) = \mu(A, X; \beta)$$

  and estimate $\hat{\beta}$.

- The estimator for the optimal treatment regime

$$\hat{d}_n(x) = \underset{k=0,\ldots,K}{\text{argmin}}\, \mu(k, x; \hat{\beta}_n).$$

- Other methods under development.

Observational Study

Observational Study

- Suppose the data are observational: a random sample from (X, A, D) where X is a vector of pre-treatment covariates.
- Patients receiving treatment 1 may not be prognostically similar to those receiving treatment 0.

Observational Study

- The intuition behind no unmeasured confounder assumption (NUCA, $D(0), D(1) \amalg A|X$): we have measured enough covariates X, so that within levels of X, the data mimics a randomized trial with the randomization probabilities now allowed to depend on X.
- This can be achieved only if we are able to measure all common predictors of A and D.
- Standard but unverifiable assumption for observational studies.

Observational Study

- Regression modeling: Model $E(D|X, A, Z)$, where Z is the confounder.

- Estimate propensity scores $P(A = 1|X, Z)$ and apply methods introduced in Section 2.

Dynamic Treatment Regimes

Dynamic Treatment Regimes (DTRs)

- Motivation : treatment of chronic illness
  - Some examples: HIV/AIDS, cancer, depression, schizophrenia, drug and alcohol addiction, ADHD, etc.
  - Multistage decision making problem
  - Longer-term treatment requires consideration and tradeoff of present versus longer term benefit.

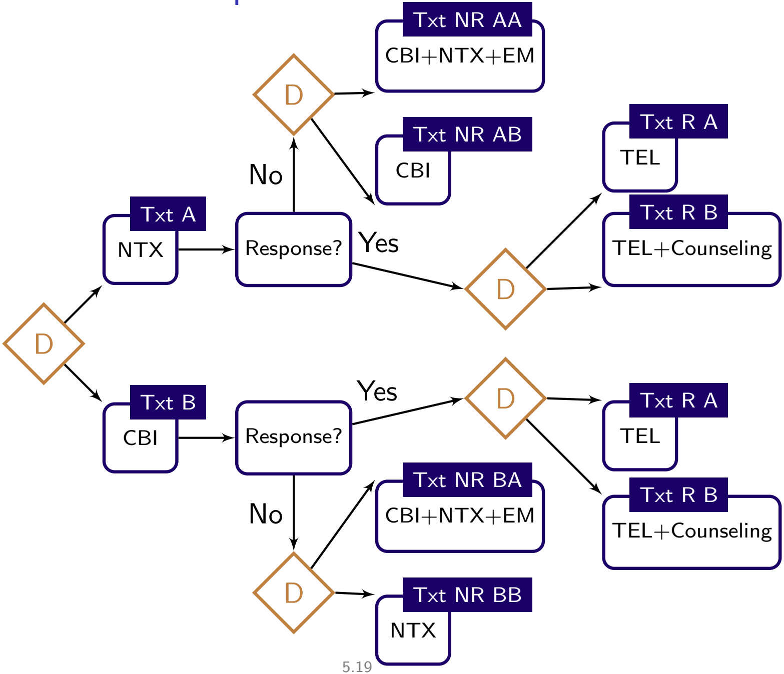
Dynamic Treatment Regimes

- Operationalize multistage decision making via as sequence of decision rules
  - One decision rule for each time (decision) point
  - A decision rule is a function inputs patient history and outputs a recommended treatment
- Aim to optimize some cumulative clinical outcome
  - Survival time
  - Depression test scores
  - Indicator of no myocardial infarction within 30 days ...

Dramatized Example

- Addiction management example inspired by the ExTENd and COMBINE trials (Murphy et al, 2007)
- Devising two-time point treatment strategy for alcohol dependent patients.
  - Initial treatment choices Naltrexone (NTX) and Combined Behavioral Intervention (CBI).
  - At six-months responders classified as responders or non-responders.
  - For responders to initial treatment, followup treatment choices are telephone monitoring (TEL) and telephone monitoring + counseling (TEL+Counseling).
  - For non-responders to initial treatment, followup treatment choices are switch initial treatments (NTX $\leftrightarrow$ CBI), or step-up initial treatment CBI + NTX + Enhanced monitoring (CBI + NTX +EM).

Dramatized Example



Dramatized Example

- H_j denote history at stage j.
- At presentation: Baseline variables x_1; accrued information $h_1 = x_1$
  - Decision point 1: Two treatment options {NTX, CBI}; rule 1: $d_1(h_1) \Rightarrow d_1 : h_1 \rightarrow \{\text{NTX, CBI}\}$
  - Between decisions 1 and 2: Collect additional information x_2, including responder status
  - Accrued information $h_2 = \{x_1, \text{treatment at decision } 1, x_2\}$
  - Decision point 2: Four options

Optimal Dynamic Treatment Regimes

- Examples of treatment regimes: Prescribe NTX initially; then assign TEL to responders; and assign step-up to non-responders.
- Optimal DTR d^* leads to the lowest expected outcome among all possible regimes

Challenges in Estimating Optimal DTRs: Delayed Effects

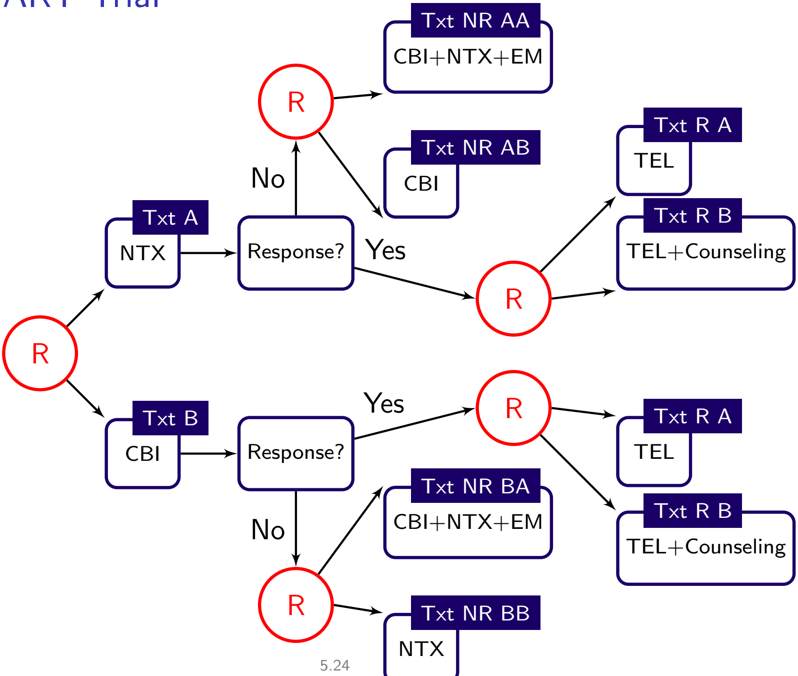
- The therapy with the higher proportion of responders might have other effects that render subsequent treatments less effective in regard to the final response.
- The therapy with lower proportion of responders may not appear best initially but may have enhanced long term effectiveness when followed by a particular maintenance treatment.
- Must consider the entire sequence of decisions
- Must accommodate intermediate information including prior treatments into current treatment choice.

Sequential Multiple Assignment Randomized Trial (SMART)

- Due the the aforementioned challenges, we need to adopt a particular design to best estimate the optimal DTRs
- SMART: designed for estimation of optimal DTRs
- Randomize subjects to the treatment options at each decision point
- Collect both initial and intermediate information on possible tailoring variables

Murphy (*Stat in Med*, 2005)

SMART Trial


R
Txt A
NTX
Response?
No
Yes
R
Txt NR AA
CBI+NTX+EM
Txt NR AB
CBI
R
Txt R A
TEL
Txt R B
TEL+Counseling
Txt B
CBI
Response?
Yes
No
R
Txt R A
TEL
Txt R B
TEL+Counseling
R
Txt NR BA
CBI+NTX+EM
Txt NR BB
NTX

Data

- (X_1, A_1, X_2, A_2, D) for each individual
  X_k: Observations available at stage k
  A_k: Treatment at stage k
  D: Primary outcome
  H_k: History at stage k, $H_1 = X_1,\ H_2 = (X_1, A_1, X_2)$

- The regime, $d = \{d_1, d_2\},\ d_k : \mathcal{H}_k \rightarrow \mathcal{A}_k$, should have the lowest $E^d(D)$, the expected outcome if all patients are assigned treatment according to d

Dynamic Programming

- Optimal regime d^* can be derived using dynamic programming (Bellman, 1957)
  - Define
    - $Q_2(h_2, a_2) \triangleq E\left(D \middle| H_2 = h_2, A_2 = a_2\right)$
    - $\tilde{D} \triangleq \min_{a_2} Q_2(H_2, a_2)$
    - $Q_1(h_1, a_1) \triangleq E\left(\tilde{D} \middle| H_1 = h_1, A_1 = a_1\right)$
  - $d_j^*(h_j) = \arg\min_{a_j \in \{0,1\}} Q_j(h_j, a_j)$

Constructing a DTR from Data: Q-learning

- When system dynamics are known dynamic programming yields the optimal DTR, but we only have data
- Q-learning: data-driven analog of dynamic programming: replaces conditional expectations with regression models
- Backwards and recursively estimates the Q-function.
- The estimated optimal sequence of decision rules

$$\hat{d}_j(h_j) = \underset{a_j \in \{0,1\}}{\operatorname{argmin}}\ \hat{Q}_j(h_j, a_j).$$

- An extension of regression to sequential treatments.

Summary

- An extremely active area of research
- Data from SMART designs can be used to construct optimal DTRs
- Q learning is a common method, though it has some drawbacks, e.g., require correct specified models
- Many other methods have been developed.