Optimal Treatment Regimes for Survival Endpoints from a Classification Perspective

Anastasios (Butch) Tsiatis
and
Xiaofei Bai

Department of Statistics
North Carolina State University

Motivation for Problem

- Coronary artery bypass grafting (CABG) versus percutaneous coronary intervention (PCI) for patients with coronary artery disease (CAD)
- **Research Question:** Based on the measured baseline covariates what treatment (PCI or CABG, coded as 0 and 1 respectively) should be recommended?
- Data from the ASCERT study–a retrospective study of patients with 2 or 3 vessel CAD treated with CABG or PCI
  - Observational study
  - We consider 7,391 patients from a substudy in 54 hospitals
  - Primary endpoint survival time
  - Baseline covariates measured prior to treatment

Treatment regime

- This problem can be cast by considering treatment regimes
- A treatment regime is a decision rule which takes an individual's baseline information and dictates which treatment to be given
- **Formally:** Letting X denote the vector of baseline covariates taking values $x \in \mathcal{X}$, then a treatment regime

  $$d : \mathcal{X} \to (0, 1)$$

  I.e., if $X = x$ then patient treated according to regime d receives
  - treatment 1 if $d(x) = 1$
  - treatment 0 if $d(x) = 0$
- Denote by $\mathcal{D}$ the class of all treatment regimes. Within this class which is the optimal regime (i.e., best in some sense)?

Potential outcomes

- Let $T^*(1)$ denote the survival time of an arbitrary patient if (possibly contrary to fact) they received treatment 1; similarly, we define $T^*(0)$
- In the population there is some unobservable distribution of $\{X, T^*(1), T^*(0)\}$
- $T^*(d) = d(X)T^*(1) + \{1 - d(X)\}T^*(0)$, the survival time for patient treated according to regime d
- We define primary outcome as $f\{T^*(d)\}$
    - $f\{T^*(d)\} = I\{T^*(d) \geq u\}$
    - $f\{T^*(d)\} = \min\{T^*(d), L\}$
- The **value** of d, denoted by $V(d) = E[f\{T^*(d)\}]$, the mean outcome for a population treated according to regime d; e.g., $P\{T^*(d) \geq u\}$ or $E[\min\{T^*(d), L\}]$. **Note** we can also define $V(1)$ and $V(0)$

Goal

- Optimal treatment regime $d^{opt} \in \mathcal{D}$ satisfies

$$V(d) \leq V(d^{opt}) \text{ for all } d \in \mathcal{D}$$

- **Statistical goal:** Estimate d^{opt} from data

Assumptions

- We observe censored survival data $(U_i, \Delta_i, A_i, X_i), i = 1, \ldots, N$,
  - $U_i = \min(T_i, C_i)$ (i.e., minimum of observed survival time T_i and censoring time C_i)
  - $\Delta_i = I(T_i \leq C_i)$ (failure indicator)
  - A_i is assigned treatment indicator
  - X_i is vector of baseline covariates

Assumptions

- Observed survival time for patient i, $T_i = A_i T_i^*(1) + (1 - A_i) T_i^*(0)$
- $C \perp\!\!\!\perp T | X, A$ (non-informative censoring)
- $A \perp\!\!\!\perp \{T^*(1), T^*(0)\} | X$ (no unmeasured confounders)

Optimal treatment regime

- Under the previous assumptions the **optimal treatment regime** is given by

$$d^{opt}(x) = I[E\{f(T)|A = 1, X = x\} \geq E\{f(T)|A = 0, X = x\}]$$

- or equivalently

$$d^{opt}(x) = I\{CF(x) \geq 0\},$$

  where the **contrast function**
  $CF(x) = E\{f(T)|A = 1, X = x\} - E\{f(T)|A = 0, X = x\}$.

Optimal treatment regime

- **Regression estimator:** An obvious strategy is to develop a model for the conditional distribution of T given A and X say, with parameters θ; derive $E\{f(T)|A, X\} = Q(X, A; \theta)$; estimate θ from the data; then

$$\widehat{d}^{opt}(x) = I\{Q(x, 1; \hat{\theta}) \geq Q(x, 0; \hat{\theta})\}$$

- or equivalently

$$\widehat{d}^{opt}(x) = I\{\widehat{CF}(x, \hat{\theta}) \geq 0\},$$

where $\widehat{CF}(x) = Q(x, 1; \hat{\theta}) - Q(x, 0; \hat{\theta})$

Regression estimator

- E.g., consider proportional hazards regression model:

$$\lambda(t|A, X) = \lambda_0(t) \exp\{\gamma_0 + \gamma_1^T X - A(\eta_0 + \eta_1^T X)\}$$

- For such a model $\widehat{d}^{opt}(x) = I(\widehat{\eta}_0 + \widehat{\eta}_1^T x \geq 0)$ (true for any function $f(\cdot)$)
- **Difficulty:** If model is misspecified then the regime $\widehat{d}^{opt}$ may not be that good

Restricted regimes

- Searching for an optimal treatment regime among all possible regimes may be too ambitious
- For practical reasons and ease of interpretability we may want to consider a class of restricted regimes $\mathcal{D}_\eta$, indexed by a finite parameter η; e.g.,
    - $\mathcal{D}_\eta = I(\eta_0 + \eta_1^T x \geq 0)$ (hyperplanes)
    - $\mathcal{D}_\eta = I(x_1 < \eta_1, x_2 < \eta_2)$ (rectangular regions)
- The optimal restricted regime $d_\eta^{opt} = d(x, \eta^{opt})$, is such that $V(d_\eta) \leq V(d_{\eta^{opt}})$ for all η

Restricted regimes

- We note that the proportional hazards model with interaction terms led to $\mathcal{D}_\eta$ in the form of hyperplanes
- Also note that the regression estimator $\widehat{d}^{opt}(x) = I(\widehat{\eta}_0 + \widehat{\eta}_1^T x \geq 0)$ may be a poor estimator of d_η^{opt} within the class $\mathcal{D}_\eta$; i.e., $(\widehat{\eta}_0, \widehat{\eta}_1)$ may be a poor estimator of $(\eta_0^{opt}, \eta_1^{opt})$ if the model is misspecified

Value search estimator

- For any regime d find a **robust** estimator for $V(d) = E[f\{T^*(d)\}]$, say $\widehat{V}(d)$
- Directly search for optimal estimator within the class $\mathcal{D}_\eta$

$$\widehat{\eta}^{opt} = \arg\max_\eta \widehat{V}(d_\eta)$$

Complete-case estimator

- If we were able to observe the potential outcomes $\{T_i^*(d), i = 1, \ldots, N\}$, then a nonparametric unbiased estimator for $V(d) = E[f\{T^*(d)\}]$ would be

$$\widehat{V}(d) = N^{-1} \sum_{i=1}^{N} f\{T_i^*(d)\}$$

- Of course, we cannot observe potential outcomes and our estimator must be based on the observed data $(U_i, \Delta_i, X_i), i = 1, \ldots, N$.
- Using missing data analogy we propose the **augmented inverse probability weighted complete case estimator** (AIPWCC)

AIPWCC estimator

We first define the following notation:

- **Propensity score** $\pi(X) = P(A = 1|X)$
- Denote by $\mathcal{C}(d, X) = Ad(X) + (1 - A)\{1 - d(X)\}$ to be the ***d*-consistency indicator**; that is $\mathcal{C}(d, X) = 1$ if patient with baseline covariate X actually receives treatment consistent with treatment regime d, and 0 otherwise
- Propensity for receiving treatment regime d,

$$\pi(d, X) = \pi(X)d(X) + \{1 - \pi(X)\}\{1 - d(X)\}$$

Additionally define

- Failure time distribution given A and X
  $H(r, a, X) = P(T \geq r | A = a, X)$,
  $H(r, d, X) = H(r, 1, X)d(X) + H(r, 0, X)\{1 - d(X)\}$
- Censoring distribution given A and X
  $K(r, a, X) = P(C \geq r | A = a, X)$,
  $K(r, d, X) = K(r, 1, X)d(X) + K(r, 0, X)\{1 - d(X)\}$

AIPWCC estimator

If we take the point of view that the propensity score and censoring distribution are known or correctly specified then using semiparametric theory for **monotone missing data**, all estimators of $V(d)$ can be written as

$$\widehat{V}(d) = N^{-1} \sum_{i=1}^{N} IF_i(d),$$

where

$$\begin{aligned} IF_i(d) &= \frac{\mathcal{C}(d, X_i)\Delta_i f(U_i)}{\pi(d, X_i)K(U_i, d, X_i)} \\ &- \left\{ \frac{\mathcal{C}(d, X_i) - \pi(d, X_i)}{\pi(d, X_i)} \right\} h_1(X_i) \\ &+ \frac{\mathcal{C}(d, X_i)}{\pi(d, X_i)} \int_0^\infty \frac{\mathrm{d}M_c(r, d, X_i)}{K(r, d, X_i)} h_2(r, X_i), \end{aligned}$$

for arbitrary functions $h_1(X_i)$ and $h_2(r, X_i)$

AIPWCC estimator

The optimal choice for $h_1(X_i)$ and $h_2(r, X_i)$ are

$$
\begin{aligned}
IF_i(d) &= \frac{\mathcal{C}(d, X_i)\Delta_i f(U_i)}{\pi(d, X_i)K(U_i, d, X_i)} \\
&- \left\{\frac{\mathcal{C}(d, X_i) - \pi(d, X_i)}{\pi(d, X_i)}\right\} \mathrm{E}\Big\{f(T_i)\Big|X_i, A_i = d(X_i)\Big\} \\
&+ \frac{\mathcal{C}(d, X_i)}{\pi(d, X_i)} \int_0^\infty \frac{\mathrm{d}M_c(r, d, X_i)}{K(r, d, X_i)} \mathrm{E}\Big\{f(T_i)\Big|T_i \geq r, X_i, A_i = d(X_i)\Big\},
\end{aligned}
$$

Motivation for AIPWCC estimator

- For any fixed treatment regime d we will observe the endpoint of interest $f\{T_i^*(d)\}$ if $\mathcal{C}(d, X_i) = 1$ and $\Delta_i = 1$, i.e., $(C_i > U_i)$.
- The probability of seeing such a complete case for an individual with covariate X_i is given by $P\{\mathcal{C}(d, X_i) = 1|X_i\} \times P\{C_i > U_i|A_i = d(X_i), X_i\} = \pi(d, X_i)K(d, U_i, X_i)$ and is the inverse of the weight used in the first term
- The first term is the **inverse probability weighted complete-case estimator** and is an unbiased estimator for $V(d)$

Motivation

- For a patient that does not receive treatment consistent with regime d; that is, $\mathcal{C}(d, X_i) = 0$, then we observe the baseline covariate X_i but, $f\{T_i^*(d)\}$ is missing. The second term and the **first augmentation term** has expectation equal to zero and the optimal choice is used to capture back some information for such individuals
- For patients that receive treatment consistent with regime d but are censored, $\mathcal{C}(d, X_i) = 1$ and $\Delta_i = 0$, then we observe the baseline covariates X_i and the partial information that their survival time $T_i^*(d)$ was greater than C_i. The third term and the **second augmentation term** also has expectation zero and the optimal choice captures back some information for these patients.

AIPWCC estimator

- Note that the estimator proposed needs $\pi(X)$, $K(r, a, X)$ and $H(r, a, X)$, for $a = (0, 1)$, which are not known to us and must be estimated from the data
- Often logistic regression models are used to estimate $\pi(X)$
- Often Cox's proportional hazards regression models, stratified by treatment, are used to estimate $K(r, a, X)$ and $H(r, a, X)$, for $a = (0, 1)$
- This estimator can be used to estimate $E[f\{T^*(1)\}]$ and $E[f\{T^*(0)\}]$ by taking $d(X_i)$ to be identically equal to 1 or 0 respectively

Double robustness property

The AIPWCC estimator will be a consistent estimator for $V(d)$ if either

- $\pi(X)$ and $K(r, a, X), a = 0, 1$ are correctly specified
- or $H(r, a, X), a = 0, 1$ are correctly specified
- For a randomized study the propensity score $\pi(X)$ is known by design and the censoring distribution can be estimated consistently using the Kaplan-Meier estimator reversing the role of failure and censoring.

Value search estimator

Equipped with this estimator for $\widehat{V}(d_\eta)$ the value search estimator for the optimal estimator within the class $\mathcal{D}_\eta$ is simply

$$\widehat{\eta}^{opt} = \text{arg max}_\eta \widehat{V}(d_\eta)$$

Issues

- $\widehat{V}(d_\eta)$ is a non-smooth non-monotonic function of η; hence difficult to maximize
- If the dimension of η is small we had some success with genetic algorithms

Classification perspective

- Recall $\widehat{V}(d_\eta) = N^{-1}\sum_{i=1}^N IF_i(d_\eta)$
-

$$\begin{aligned} IF_i(d_\eta) &= d(X_i,\eta)IF_i(1) + \{1 - d(X_i,\eta)\}IF_i(0) \\ &= d(X_i,\eta)\{IF_i(1) - IF_i(0)\} + IF_i(0) \end{aligned}$$

- $\widehat{V}(d_\eta) = N^{-1}\sum_{i=1}^N d(X_i,\eta)\widehat{CF}(X_i) +$ terms not involving η, where $\widehat{CF}(X_i) = IF_i(1) - IF_i(0)$
- Note that $E\{IF_1(1) - IF_i(0)|X_i\} = CF(X_i)$, where
- $CF(X_i) = E\{f(T_i)|A_i = 1, X_i\} - E\{f(T_i)|A_i = 0, X_i\}$ is the *contrast function*
- We also note that the optimal regime $d^{opt}(x) = I\{CF(x) \geq 0\}$

Classification perspective

Some further algebra

- $d(X_i,\eta)\widehat{CF}(X_i) = -|\widehat{CF}(X_i)|[I\{\widehat{CF}(X_i) > 0\} - d(X_i,\eta)]^2 +$ terms not involving η
- Hence $\widehat{\eta}^{opt} = \arg\min_\eta \sum_{i=1}^{N} |\widehat{CF}(X_i)|[I\{\widehat{CF}(X_i) > 0\} - d(X_i,\eta)]^2$

Classification Methods

Generic classification situation:

- Z = *outcome*, *class*, *label*; here, $Z = \{0, 1\}$ (*binary*)
- X = vector of covariates, *features* taking values in $\mathcal{X}$, the *feature space*
- d is a *classifier:* $d : \mathcal{X} \rightarrow \{0, 1\}$
- $\mathcal{D}$ is a *family of classifiers*, e.g.,
  - *Hyperplanes* of the form

$$I(\eta_0 + \eta_1 X_1 + \eta_2 X_2 > 0)$$

  - *Rectangular regions* of the form

$$I(X_1 < a_1) + I(X_1 \geq a_1, X_2 < a_2)$$

Classification Methods

Generic classification problem:

- *Training set:* $(X_i, Z_i), i = 1, \ldots, N$
- *Find* classifier $d \in \mathcal{D}$ that minimizes
  - *Classification error*

$$\sum_{i=1}^{N} \{Z_i - d(X_i)\}^2$$

  - *Weighted classification error*

$$\sum_{i=1}^{N} w_i \{Z_i - d(X_i)\}^2$$

Classification Methods

Approaches:

- This problem has been studied extensively by *statisticians* and *computer scientists*
- *Machine learning* (*supervised* learning)
- Many methods and software are available
- *Recursive partitioning* (*CART*): Rectangular regions
- *Support vector machines:* Hyperplanes, etc.

Classification perspective

From this perspective the value search estimator

$$\widehat{\eta}^{opt} = \arg\min_{\eta} \sum_{i=1}^{N} |\widehat{CF}(X_i)| [I\{\widehat{CF}(X_i) > 0\} - d(X_i, \eta)]^2$$

is a weighted classification problem with

- X the feature space
- The class label $Z_i = I\{\widehat{CF}(X_i) > 0\}$
- The weight $w_i = |\widehat{CF}(X_i)|$
- $\mathcal{D}_\eta$ is the family of classifiers

ASCERT study

- Retrospective analysis of patients with two or three vessel coronary artery disease treated by PCI (0) or CABG (1)
- 7,391 patients from a substudy from 54 hospitals were used for this analysis
- 28 baseline covariates were used including
  - demographics (e.g., age, gender)
  - risk factors (e.g., body mass index, smoking)
  - symptoms and history of heart disease (e.g., chest pain, congestive heart failure
  - comorbidities (e.g., diabetes)

ASCERT study

- **Primary outcome** was survival at four years $f\{T^*(d)\} = I\{T^*(d) \geq 4)$
- Using all 28 covariates **propensity score** $\pi(X)$ was estimated using **logistic regression model**
- $H(r, a, X)$ and $K(r, a, X)$ for $a = 0, 1$ were estimated using **proportional hazards models** stratified by treatment

ASCERT study

- We considered regimes in the form of hyperplanes; i.e., $d(X, \eta) = I(\eta_0 + \eta_1^T X \geq 0)$
- We used **support vector machines** with L_1 norm where we wrote our own software using linear programming to estimate η

Results from the ASCERT study

Estimators of the value function $P\{T^*(d) \geq 4\}$

- $\widehat{V}(\widehat{d}^{opt}) = .862$
- $\widehat{V}(1) = .841$
- $\widehat{V}(0) = .816$
- CI of $\{\widehat{V}(\widehat{d}^{opt}) - \widehat{V}(1)\} = (0.005, 0.036)$
- CI of $\{\widehat{V}(\widehat{d}^{opt}) - \widehat{V}(0)\} = (0.028, 0.064)$

Results from the ASCERT study

Table: ASCERT analysis with original contrast function.

Treatment	Number of patients	Survival Probability (%)	
		CABG	PCI
CABG	5024	86.9	80.2
PCI	2367	76.5	84.3

Concluding remarks

- Estimating optimal treatment regime
  - Regression estimator versus value search estimator
  - Bias-variance tradeoff
  - AIPWCC estimator is guaranteed to be a consistent estimator of the value function for a randomized study and is doubly-robust for an observational study
- Generalize to more than one decision and consider dynamic treatment regimes

Reference

Bai et al. (2016). Optimal treatment regimes for survival endpoints using a locally-efficient doubly-robust estimator from a classification perspective. *In revision Lifetime Data Analysis*.