3. Single Decision Treatment Regimes: Fundamentals

Basic set-up

For simplicity: Focus mostly on two treatment options coded as 0 and 1

$$\mathcal{A}_1 = \{0, 1\}$$

Recall: With $K = 1$

- Baseline information $x_1 \in \mathcal{X}$, history $h_1 = x_1 \in \mathcal{H}_1$
- Treatment regime

  $$d = \{d_1(h_1)\}$$

  comprising a single rule $d_1(h_1)$ such that $d_1 : \mathcal{H}_1 \rightarrow \mathcal{A}_1$
- I.e., for history h_1, $d_1(h_1) = 0$ selects option 0, $d_1(h_1) = 1$ selects option 1

Convention: Larger outcomes Y are more beneficial (without loss of generality)

Class of all possible regimes

Clearly: An infinitude of possible rules d_1 and thus regimes d

- $\mathcal{D}$ = class of all single decision treatment regimes
- Two possible static rules, $d_1(h_1) \equiv 1$, $d_1(h_1) \equiv 0$ for all $h_1 \in \mathcal{H}_1$
- All other rules and thus regimes are dynamic, and $\mathcal{D}$ is likely an infinite class

Example: Decision 1, acute leukemia

- $h_1 = x_1$ includes age (years), baseline white blood cell count (WBC $\times 10^3/\mu l$)
- $\mathcal{A}_1 = \{C_1, C_2\} = \{0, 1\}$

Examples of rules

Example 1: Rules involving thresholds (rectangular region); e.g., 'If age $<$ 50 and WBC $<$ 10, then give C_2; otherwise, give C_1"

$$d_1(h_1) = \mathrm{I}(\text{age} < 50 \text{ and WBC} < 10)$$

Example 2: Rules involving linear combinations (hyperplane); e.g., "if age + 8.7 $\log$(WBC) $-$ 60 $>$ 0, give C_2; otherwise give C_1"

$$d_1(h_1) = \mathrm{I}\{\text{age} + 8.7 \log(\text{WBC}) - 60 > 0\}$$

Infinitude:

- Change thresholds, linear coefficients
- Other functions of age and WBC (and other components of h_1)
- Exception: h_1 contains c binary components; 2^c possible rules
- Almost always: h_1 contains continuous and discrete variables, can be high-dimensional

Potential outcomes framework

For a randomly chosen individual from the population:

- History $H_1 = X_1$, potential outcomes $Y^*(0)$ and $Y^*(1)$ that would be achieved under options 0 and 1

Potential outcome for regime $d \in \mathcal{D}$: The outcome such an individual would achieve if assigned treatment according to (the rule d_1 in) regime d

$$
\begin{aligned}
Y^*(d) &= Y^*(1)\,\mathrm{I}\{d_1(H_1) = 1\} + Y^*(0)\,\mathrm{I}\{d_1(H_1) = 0\} \qquad (3.1)\\
&= Y^*(1)d_1(H_1) + Y^*(0)\,\{1 - d_1(H_1)\}
\end{aligned}
$$

- I.e., if d dictates option 1, $Y^*(d) = Y^*(1)$, similarly for option 0, $Y^*(0)$
- For static regime with $d_1(h_1) \equiv 1$, $Y^*(d) = Y^*(1)$, similarly for option 0, $Y^*(0)$

Potential outcomes framework

$\mathcal{A}_1$ with more than two options:

- $d_1(h_1)$ returns options $a_1 \in \mathcal{A}_1$
- $Y^*(a_1)$ = potential outcome that would be achieved by an individual with history H_1 if she were to receive option $a_1 \in \mathcal{A}_1$

$$Y^*(d) = \sum_{a_1 \in \mathcal{A}_1} Y^*(a_1)\, \mathrm{I}\{d_1(H_1) = a_1\} \tag{3.2}$$

Value of a treatment regime

For regime $d \in \mathcal{D}$: With potential outcome $Y^*(d)$ as in (3.1) or (3.2)

- $E\{Y^*(d)\}$ = expected outcome if all individuals in the population were to receive treatment according to rule d_1 in d
- Referred to as the *value* of regime $d \in \mathcal{D}$, denoted here as

$$\mathcal{V}(d) = E\{Y^*(d)\}$$

Remarks:

- For static regimes with rules $d_1(h_1) \equiv 1$ and $d_1(h_1) \equiv 0$, $\mathcal{V}(d) = E\{Y^*(1)\}$ and $E\{Y^*(0)\}$, and the average causal treatment effect δ^* is the difference in their values
- Is it more beneficial on average to select treatment using a dynamic regime $d \in \mathcal{D}$ relative to always administering option 1 regardless of history?

$$E\{Y^*(d)\} - E\{Y^*(1)\} > 0?$$

3. Single Decision Treatment Regimes: Fundamentals

Estimation of $\mathcal{V}(d)$

For a given (fixed) regime $d \in \mathcal{D}$: Estimate the value $\mathcal{V}(d)$ from i.i.d. observed data

$$(X_{1i}, A_{1i}, Y_i), \ \ i = 1, \ldots, n \tag{3.3}$$

- $H_{1i} = X_{1i}$ = history for individual i
- A_{1i} = treatment option in $\mathcal{A}_1$ actually received by i

Challenge: Estimate the quantity $\mathcal{V}(d) = E\{Y^*(d)\}$ defined in terms of potential outcomes using observed data (3.3)

- Under what conditions can we deduce the distribution of $Y^*(d)$, which depends on that of $\{X_1, Y^*(1), Y^*(0)\}$, from the distribution of (X_1, A_1, Y)?
- Possible under the following assumptions

Identifiability assumptions

SUTVA (consistency):

$$Y_i = Y_i^*(1)A_{1i} + Y_i^*(0)(1 - A_{1i}), \;\; i = 1, \ldots, n \tag{3.4}$$

No unmeasured confounders assumption (NUC):

$$\{Y^*(1), Y^*(0)\} \perp\!\!\!\perp A_1 | H_1 \tag{3.5}$$

Positivity assumption:

$$P(A_1 = a_1 | H_1 = h_1) > 0, \quad a_1 = 0, 1 \tag{3.6}$$

for all $h_1 \in \mathcal{H}_1$ such that $P(H_1 = h_1) > 0$

- Generalize in obvious way to $\mathcal{A}_1$ with more than two options
- We adopt these assumptions hold in what follows

Outcome regression estimator

Similar to manipulations in (2.14)-(2.15):

$$
\begin{aligned}
E\{Y^*(d)\} &= E\left(E\left[Y^*(1)\mathrm{I}\{d_1(H_1)=1\} + Y^*(0)\mathrm{I}\{d_1(H_1)=0\}\,\middle|\, H_1\right]\right) \\
&= E\Big[E\{Y^*(1)|H_1\}\mathrm{I}\{d_1(H_1)=1\} + E\{Y^*(0)|H_1\}\mathrm{I}\{d_1(H_1)=0\}\Big] \\
&= E\Big[E\{Y^*(1)|H_1, A_1=1\}\mathrm{I}\{d_1(H_1)=1\} \\
&\qquad + E\{Y^*(0)|H_1, A_1=0\}\mathrm{I}\{d_1(H_1)=0\}\Big] \quad \text{by NUC (3.5)} \\
&= E\left[E(Y|H_1, A_1=1)\mathrm{I}\{d_1(H_1)=1\} + E(Y|H_1, A_1=0)\mathrm{I}\{d_1(H_1)=0\}\right] \\
&= E\left[E(Y|H_1, A_1=1)d_1(H_1) + E(Y|H_1, A_1=0)\{1-d_1(H_1)\}\right] \\
&\qquad\qquad \text{by SUTVA (3.4)}
\end{aligned}
$$

- Conditional expectations are well defined by the positivity assumption (3.6)

Outcome regression estimator

$\mathcal{A}_1$ with more than two options: By similar manipulations

$$E\{Y^*(d)\} = E\left[\sum_{a_1 \in \mathcal{A}_1} E(Y|H_1, A_1 = a_1)\,\mathrm{I}\{d_1(H_1) = a_1\}\right]$$

Result: The value $\mathcal{V}(d) = E\{Y^*(d)\}$ of a regime $d \in \mathcal{D}$ can be represented in terms of the observed data (X_1, A_1, Y)

- In terms of the regression of outcome on history and treatment received

$$E(Y|H_1 = h_1, A_1 = a_1) = Q_1(h_1, a_1)$$

- E.g., for $\mathcal{A}_1 = \{0, 1\}$

$$E\{Y^*(d)\} = E\,[Q_1(H_1, 1)\mathrm{I}\{d_1(H_1) = 1\} + Q_1(H_1, 0)\mathrm{I}\{d_1(H_1) = 0\}]$$

Outcome regression estimator

Suggests: If $Q_1(h_1, a_1)$ were known, natural estimators for $\mathcal{V}(d)$

$$n^{-1} \sum_{i=1}^{n} \Big[Q_1(H_{1i}, 1)\mathrm{I}\{d_1(H_{1i}) = 1\} + Q_1(H_{1i}, 0)\mathrm{I}\{d_1(H_{1i}) = 0\} \Big]$$

$$n^{-1} \sum_{i=1}^{n} \left[\sum_{a_1 \in \mathcal{A}_1} Q_1(H_{1i}, a_1)\, \mathrm{I}\{d_1(H_{1i}) = a_1\} \right]$$

- Obvious strategy: Posit a model $Q_1(h_1, a_1; \beta_1)$ with parameter β_1; e.g., with $\mathcal{A}_1 = \{0, 1\}$, continuous Y

$$Q_1(h_1, a_1; \beta_1) = \beta_{11} + \beta_{12}^T h_1 + \beta_{13} a_1 + \beta_{14}^T h_1 a_1, \quad \beta_1 = (\beta_{11}, \beta_{12}^T, \beta_{13}, \beta_{14}^T)^T$$

  and similarly for binary Y using logistic regression
- Fit using suitable M-estimation techniques and substitute the fitted model $Q_1(h_1, a_1; \widehat{\beta}_1)$ in the above expressions

Outcome regression estimator

Result: Outcome regression estimator for the value $\mathcal{V}(d)$ of $d \in \mathcal{D}$

$$\widehat{\mathcal{V}}_Q(d) \tag{3.7}$$

$$= n^{-1} \sum_{i=1}^{n} \Big[Q_1(H_{1i}, 1; \widehat{\beta}_1) \mathrm{I}\{d_1(H_{1i}) = 1\} + Q_1(H_{1i}, 0; \widehat{\beta}_1) \mathrm{I}\{d_1(H_{1i}) = 0\} \Big]$$

and similarly for general $\mathcal{A}_1$

- If $Q_1(h_1, a_1; \beta_1)$ is correctly specified, with true value $\beta_{1,0}$ of β_1, under SUTVA, NUC, and positivity assumption, $\widehat{\mathcal{V}}_Q(d)$ is a consistent estimator for $\mathcal{V}(d)$
- Approximate large sample distribution for $\widehat{\mathcal{V}}_Q(d)$ obtained by stacking estimating equations and appealing to M-estimation theory

Inverse probability weighted estimator

For fixed $d \in \mathcal{D}$: If we could observe $Y_i^*(d)$, $i = 1, \ldots, n$, obvious estimator for $\mathcal{V}(d) = E\{Y^*(d)\}$

$$n^{-1} \sum_{i=1}^{n} Y_i^*(d)$$

- Clearly, by the definition of $Y^*(d)$ and SUTVA, if $A_{1i} = d_1(H_{1i})$, then $Y_i = Y_i^*(d)$, so $Y_i^*(d)$ is observed
- E.g., if $d_1(H_{1i}) = 1$ and $A_{1i} = 1$, then $Y_i^*(d) = Y_i^*(1)$ and $Y_i = Y_i^*(1)$
- If $A_i \neq d_1(H_{1i})$, then $Y_i \neq Y_i^*(d)$, and $Y_i^*(d)$ is "missing"
- Suggests an inverse weighting strategy similar to that used to estimate the average causal treatment effect δ^*

Inverse probability weighted estimator

Regime consistency indicator: Define

$$\mathcal{C}_d = \mathrm{I}\{A_1 = d_1(H_1)\} = A_1 \mathrm{I}\{d_1(H_1) = 1\} + (1 - A_1)\mathrm{I}\{d_1(H_1) = 0\} \tag{3.8}$$

- If $\mathcal{C}_d = 1$, $Y^*(d)$ is observed; else, it is missing

Propensity for treatment consistent with d:

$$\pi_{d,1}(H_1) = P(\mathcal{C}_d = 1 | H_1) \tag{3.9}$$

Suggests: Inverse probability weighted estimator for $\mathcal{V}(d)$

$$\widehat{\mathcal{V}}_{IPW}(d) = n^{-1} \sum_{i=1}^{n} \frac{\mathcal{C}_{d,i} Y_i}{\pi_{d,1}(H_{1i})}. \tag{3.10}$$

- Weight outcomes from individuals with particular H_1 who received treatment consistent with d by $1/\pi_{d,1}(H_1)$

Inverse probability weighted estimator

From (3.8) and (3.9):

$$\begin{aligned}\pi_{d,1}(H_1) &= E\left[A_1 \mathrm{I}\{d_1(H_1)=1\} + (1-A_1)\mathrm{I}\{d_1(H_1)=0\} | H_1\right] \\ &= \pi_1(H_1)\mathrm{I}\{d_1(H_1)=1\} + \{1-\pi_1(H_1)\}\mathrm{I}\{d_1(H_1)=0\} \qquad (3.11) \\ &= \pi_1(H_1)^{d_1(H_1)}\{1-\pi_1(H_1)\}^{1-d_1(H_1)}\end{aligned}$$

- Must have $\pi_{d,1}(H_1) > 0$ for all $d \in \mathcal{D}$, which holds under the positivity assumption

Can show: $\widehat{\mathcal{V}}_{IPW}(d)$ in (3.10) is an unbiased estimator for $\mathcal{V}(d)$

Inverse probability weighted estimator

$$
\begin{aligned}
E\left\{\frac{\mathcal{C}_d Y}{\pi_{d,1}(H_1)}\right\} &= E\left\{\frac{\mathcal{C}_d Y^*(d)}{\pi_{d,1}(H_1)}\right\} \\
&= E\left[E\left\{\left.\frac{\mathcal{C}_d Y^*(d)}{\pi_{d,1}(H_1)}\right| Y^*(1), Y^*(0), H_1\right\}\right] \\
&= E\left[\frac{E\{\mathcal{C}_d | Y^*(1), Y^*(0), H_1\} Y^*(d)}{\pi_1(H_1)\mathrm{I}\{d_1(H_1)=1\} + \{1-\pi_1(H_1)\}\mathrm{I}\{d_1(H_1)=0\}}\right] \\
&= E\{Y^*(d)\}
\end{aligned}
$$

because

$$
\begin{aligned}
&E\{\mathcal{C}_d | Y^*(1), Y^*(0), H_1\} \\
&\quad = E\left[A_1\mathrm{I}\{d_1(H_1)=1\} + (1-A_1)\mathrm{I}\{d_1(H_1)=0\} \,\middle|\, Y^*(1)\,, Y^*(0), H_1\right] \\
&\quad = E(A_1|H_1)\mathrm{I}\{d_1(H_1)=1\} + E(1-A_1|H_1)\mathrm{I}\{d_1(H_1)=0\} \\
&\quad = \pi_1(H_1)\mathrm{I}\{d_1(H_1)=1\} + \{1-\pi_1(H_1)\}\mathrm{I}\{d_1(H_1)=0\}
\end{aligned}
$$

Inverse probability weighted estimator

Randomized study: $\pi_1(h_1)$ and thus $\pi_{d,1}(h_1)$ is known

Observational study: Posit model $\pi_1(h_1; \gamma_1)$, e.g., logistic model as in (2.28), and obtain maximum likelihood estimator $\widehat{\gamma}_1$

$$\pi_1(h_1; \gamma_1) = \frac{\exp(\gamma_{11} + \gamma_{12}^T h_1)}{1 + \exp(\gamma_{11} + \gamma_{12}^T h_1)}, \quad \gamma_1 = (\gamma_{11}, \gamma_{12}^T)^T \tag{3.12}$$

- Induces a model $\pi_{d,1}(h_1; \gamma_1)$, correctly specified if $\pi_1(h_1; \gamma_1)$ is

IPW estimator: Substitute in (3.10)

$$\widehat{\mathcal{V}}_{IPW}(d) = n^{-1} \sum_{i=1}^{n} \frac{\mathcal{C}_{d,i} Y_i}{\pi_{d,1}(H_{1i}; \widehat{\gamma}_1)} \tag{3.13}$$

- Consistent estimator for $\mathcal{V}(d)$ if $\pi_1(h_1; \gamma_1)$ correctly specified

Alternative inverse probability weighted estimator

$$\widehat{\mathcal{V}}_{IPW*}(d) = \left\{ \sum_{i=1}^{n} \frac{\mathcal{C}_{d,i}}{\pi_{d,1}(H_{1i}; \widehat{\gamma}_1)} \right\}^{-1} \sum_{i=1}^{n} \frac{\mathcal{C}_{d,i} Y_i}{\pi_{d,1}(H_{1i}; \widehat{\gamma}_1)} \qquad (3.14)$$

- Can be shown by manipulations similar to those above that a summand of the first term in (3.14) has expectation = 1, so $\widehat{\mathcal{V}}_{IPW*}(d)$ is a consistent estimator for $\mathcal{V}(d)$ if $\pi_1(h_1; \gamma_1)$ is correctly specified
- Exhibits considerably smaller sampling variation than $\widehat{\mathcal{V}}_{IPW}(d)$ in practice (relatively more efficient)

Approximate large sample distributions: For either of (3.13) or (3.14), can be obtained by stacking estimating equations and appealing to M-estimation theory

Equivalent representations

Possibly simpler expressions for $\widehat{\mathcal{V}}_{IPW}(d)$ and $\widehat{\mathcal{V}}_{IPW*}(d)$:

- When $\mathcal{C}_d = 1$, $A_1 = d_1(H_1)$
- Straightforward: $\widehat{\mathcal{V}}_{IPW}(d)$ and $\widehat{\mathcal{V}}_{IPW*}(d)$ are unchanged if $\pi_{d,1}(H_{1i}; \widehat{\gamma}_1)$ is replaced for each i by

$$\begin{aligned} &\pi_1(H_{1i}; \widehat{\gamma}_1)\mathrm{I}(A_{1i} = 1) + \{1 - \pi_1(H_{1i}; \widehat{\gamma}_1)\}\mathrm{I}(A_{1i} = 0) \\ &\quad = \pi_1(H_{1i}; \widehat{\gamma}_1)A_{1i} + \{1 - \pi_1(H_{1i}; \widehat{\gamma}_1)\}(1 - A_{1i}) \qquad (3.15) \\ &\quad = \pi_1(H_{1i}; \widehat{\gamma}_1)^{A_{1i}}\{1 - \pi_1(H_{1i}; \widehat{\gamma}_1)\}^{(1-A_{1i})} \end{aligned}$$

- Some literature accounts present these estimators directly in this form

Outcome regression vs. IPW

Tradeoff: Is the same as for estimators for average causal treatment effect δ^*

- $\widehat{\mathcal{V}}_Q(d)$ requires correct modeling of outcome regression $Q(h_1, a_1)$
- $\widehat{\mathcal{V}}_{IPW}(d)$ and $\widehat{\mathcal{V}}_{IPW*}(d)$ require correct modeling of propensity score $\pi_1(h_1)$
- Randomized study: IPW estimators are guaranteed to be consistent because $\pi_1(h_1)$ is known, while $\widehat{\mathcal{V}}_Q(d)$ still requires a correct regression model

Counterintuitive result persists: It is preferable on efficiency grounds to estimate $\pi_1(h_1)$ even if it is known as on Slide 90

Augmented inverse probability weighted estimator

Analogous to the class of AIPW estimators for the average causal treatment effect: If $\pi_1(h_1; \gamma_1)$ is correctly specified, from semiparametric theory (Robins et al., 1994; Tsiatis, 2006), all consistent and asymptotically normal estimators for $\mathcal{V}(d)$ for fixed $d \in \mathcal{D}$ are asymptotically equivalent to an estimator of form

$$\widehat{\mathcal{V}}_{AIPW}(d) = n^{-1} \sum_{i=1}^{n} \left[\frac{\mathcal{C}_{d,i} Y_i}{\pi_{d,1}(H_{1i}; \widehat{\gamma}_1)} - \frac{\mathcal{C}_{d,i} - \pi_{d,1}(H_{1i}; \widehat{\gamma}_1)}{\pi_{d,1}(H_{1i}; \widehat{\gamma}_1)} L_1(H_{1i}) \right] \tag{3.16}$$

- $L_1(H_1)$ is an arbitrary function of H_1
- The "augmentation term" can be shown to have conditional expectation given H_{1i} equal to zero when evaluated at the true $\gamma_{1,0}$ and serves to increase efficiency over $\widehat{\mathcal{V}}_{IPW}(d)$ in (3.13)

Optimal AIPW estimator

Among class (3.16): The optimal, efficient estimator; i.e., with smallest asymptotic variance, is obtained with

$$\begin{aligned} L_1(H_1) &= E\{Y^*(d)|H_1\} \\ &= Q_1(H_1,1)\mathrm{I}\{d_1(H_1)=1\} + Q_1(H_1,0)\mathrm{I}\{d_1(H_1)=0\} \end{aligned} \quad (3.17)$$

- Follows using SUTVA, NUC, positivity assumption
- Suggests: Posit a model $Q_1(h_1, a_1; \beta_1)$ for $Q_1(h_1, a_1)$ and represent (3.17) as

$$\mathcal{Q}_{d,1}(H_1;\beta_1) = Q_1(H_1,1;\beta_1)\mathrm{I}\{d_1(H_1)=1\}+Q_1(H_1,0;\beta_1)\mathrm{I}\{d_1(H_1)=0\}$$

- Estimate $\widehat{\beta}_1$ by β_1

Optimal AIPW estimator

Leads to:

$$\widehat{\mathcal{V}}_{AIPW}(d) = n^{-1} \sum_{i=1}^{n} \left[\frac{\mathcal{C}_{d,i} Y_i}{\pi_{d,1}(H_{1i}; \widehat{\gamma}_1)} - \frac{\mathcal{C}_{d,i} - \pi_{d,1}(H_{1i}; \widehat{\gamma}_1)}{\pi_{d,1}(H_{1i}; \widehat{\gamma}_1)} \mathcal{Q}_{d,1}(H_{1i}; \widehat{\beta}_1) \right] \tag{3.18}$$

- The augmentation term attempts to gain precision by recovering information from individuals for whom $\mathcal{C}_d = 0$ (so did not receive treatment consistent with d)
- Can be shown: $\widehat{\mathcal{V}}_{AIPW}(d)$ is unchanged by replacing $\pi_{d,1}(H_{1i}; \widehat{\gamma}_1)$ by (3.15)

$$\pi_1(H_{1i}; \widehat{\gamma}_1) \mathrm{I}(A_{1i} = 1) + \{1 - \pi_1(H_{1i}; \widehat{\gamma}_1)\} \mathrm{I}(A_{1i} = 0)$$

Double robustness

Can be shown: By an argument similar to that on Slides 95-99, $\widehat{\mathcal{V}}_{AIPW}(d)$ in (3.18) is doubly robust

- Consistent estimator for $\mathcal{V}(d)$ if either the propensity score model $\pi_1(h_1;\gamma_1)$ or the outcome regression model $Q_1(h_1,a_1;\beta_1)$ is correctly specified
- Randomized study: Form of $\pi_1(h_1;\gamma_1)$ is known, so $\widehat{\mathcal{V}}_{AIPW}(d)$ is consistent regardless of whether or not $Q_1(h_1,a_1;\beta_1)$ is correctly specified and is relatively more efficient than $\widehat{\mathcal{V}}_{IPW}(d)$

If both propensity and outcome regression models are correctly specified: $\widehat{\mathcal{V}}_{AIPW}(d)$ in (3.18) achieves the smallest asymptotic variance among all AIPW estimators of the form (3.16)

- Locally efficient estimator

Large sample properties: (3.18) is an M-estimator, so can derive based on stacked estimating equations

Outcome regression vs. locally efficient AIPW

- Outcome regression estimator (3.7) requires $Q_1(h_1, a_1; \beta_1)$ correctly specified
- If it is, $\widehat{\mathcal{V}}_Q(d)$, which is outside the class (3.16), is more efficient than (3.18) even if both propensity and outcome regression models are correctly specified
- In practice: Gain in efficiency of $\widehat{\mathcal{V}}_Q(d)$ over $\widehat{\mathcal{V}}_{AIPW}(d)$ is often negligible
- Doubly robust, AIPW estimator is attractive alternative
- Zhang, Tsiatis, Laber, and Davidian (2012)

3. Single Decision Treatment Regimes: Fundamentals

Key goal: An optimal regime

Recall: A key goal of precision medicine is to identify an *optimal treatment regime*

$$d^{opt} \in \mathcal{D},$$

where

$$d^{opt} = \{d_1^{opt}(h_1)\}$$

and leads to the "best" decision and most beneficial expected outcome

- Formalize this definition

Intuitively:

- Conventional treatment comparisons are based on comparing the associated values
- Suggests: An optimal regime should lead to the maximum value among all regimes $d \in \mathcal{D}$
- As we will see shortly, this definition leads to the "best" decisions for individuals given their histories

Formal definition

An optimal regime $d^{opt} \in \mathcal{D}$ satisfies

$$d^{opt} = \arg\max_{d \in \mathcal{D}} E\{Y^{*}(d)\} = \arg\max_{d \in \mathcal{D}} \mathcal{V}(d)$$

or, equivalently,

$$E\{Y^{*}(d^{opt})\} \geq E\{Y^{*}(d)\} \text{ for all } d \in \mathcal{D} \tag{3.19}$$

- In principle, is possible that there is more than one regime d^{opt} satisfying (3.19), discussed shortly

Form of an optimal regime

Intuitive demonstration, $\mathcal{A}_1 = \{0, 1\}$: From (3.1) can write

$$\mathcal{V}(d) = E\{Y^*(d)\} = E\left[E\{Y^*(d)|H_1\}\right]$$
$$= E\left[E\{Y^*(1)|H_1\}\mathrm{I}\{d_1(H_1) = 1\} + E\{Y^*(0)|H_1\}\mathrm{I}\{d_1(H_1) = 0\}\right]$$

- Maximizing the expression inside the outer expectation (which is wrt to distribution of H_1) at any h_1 leads to $E\{Y^*(d)\}$ as large as possible
- This expression is as large as possible if

$$d_1(h_1) = 1 \;\text{ when }\; E\{Y^*(1)|H_1 = h_1\} > E\{Y^*(0)|H_1 = h_1\}$$
$$d_1(h_1) = 0 \;\text{ when }\; E\{Y^*(1)|H_1 = h_1\} < E\{Y^*(0)|H_1 = h_1\}$$

- I.e., $d_1(h_1)$ chooses $a_1 \in \mathcal{A}_1$ that maximizes $E\{Y^*(a_1)|H_1 = h_1\}$ for all h_1
- This definition extends straightforwardly to general $\mathcal{A}_1$

Form of an optimal regime

Result: An optimal regime d^{opt} is characterized by the rule

$$d_1^{opt}(h_1) = \underset{a_1 \in \mathcal{A}_1}{\arg\max}\, E\{Y^*(a_1)|H_1 = h_1\} \tag{3.20}$$

for all h_1 for which $P(H_1 = h_1) > 0$

- Chooses the option in $\mathcal{A}_1$ having the maximum expected outcome conditional on history
- The best decision for an individual patient with realized history h_1 is to choose the option that maximizes the expected value of the outcome that would be achieved for such a patient
- In this sense individualizing the decision to the patient

Unique representation

Consider $\mathcal{A}_1 = \{0, 1\}$: If for some h_1

$$E\{Y^*(1)|H_1 = h_1\} = E\{Y^*(0)|H_1 = h_1\}$$

a rule (3.20) that chooses option 1 for this h_1 and another that chooses option 0 for this h_1 define regimes that both achieve the maximum value

- Can designate one of the options as the default when both are equally beneficial for any h_1
- Convention: Option 0 is the default; 0 often corresponds to control or standard of care, while 1 corresponds to experimental treatment
- Then (3.20) is equivalent to

$$d_1^{opt}(h_1) = \mathrm{I}\left[E\{Y^*(1)|H_1 = h_1\} > E\{Y^*(0)|H_1 = h_1\}\right] \quad (3.21)$$

Formal argument

Proposition: The regime d^{opt} with rule d_1^{opt} given in (3.20)

$$d_1^{opt}(h_1) = \underset{a_1 \in \mathcal{A}_1}{\arg\max}\, E\{Y^*(a_1)|H_1 = h_1\} \quad \text{for all } h_1$$

satisfies (3.19)

$$E\{Y^*(d^{opt})\} \geq E\{Y^*(d)\} \text{ for all } d \in \mathcal{D}$$

and is thus an optimal treatment regime

Proof: Choose arbitrary $d \in \mathcal{D}$. Because

$$E\{Y^*(d)\} = E\left[E\{Y^*(d)|H_1\}\right] \quad \text{and} \quad E\{Y^*(d^{opt})\} = E\left[E\{Y^*(d^{opt})|H_1\}\right]$$

the result follows if we show that

$$E\{Y^*(d^{opt})|H_1 = h_1\} \geq E\{Y^*(d)|H_1 = h_1\} \quad \text{for all } h_1 \qquad (3.22)$$

Formal argument

From (3.20), it follows for any h_1 that

$$E\{Y^*(d^{opt})|H_1 = h_1\} = \max_{a_1 \in \mathcal{A}_1} E\{Y^*(a_1)|H_1 = h_1\} = V_1(h_1)$$

Using this and the definition of $Y^*(d)$ (3.1)

$$\begin{aligned}
E\{Y^*(d^{opt})|H_1 = h_1\} &= \max_{a_1 \in \mathcal{A}_1} E\{Y^*(a_1)|H_1 = h_1\} \\
&= \max_{a_1 \in \mathcal{A}_1} E\{Y^*(a_1)|H_1 = h_1\}\, [\mathrm{I}\{d_1(h_1) = 1\} + \mathrm{I}\{d_1(h_1) = 0\}] \\
&\geq E\{Y^*(1)|H_1 = h_1\}\, \mathrm{I}\{d_1(h_1) = 1\} + E\{Y^*(0)|H_1 = h_1\}\, \mathrm{I}\{d_1(h_1) = 0\} \\
&= E\left[Y^*(1)\mathrm{I}\{d_1(h_1) = 1\} + Y^*(0)\,\mathrm{I}\{d_1(h_1) = 0\} \,\middle|\, H_1 = h_1 \right] \\
&= E\{Y^*(d)|H_1 = h_1\} \quad \text{which is (3.22)} \qquad (3.23)
\end{aligned}$$

Value function: $V_1(h_1)$ = expected outcome using the option selected by $d_1^{opt}(h_1)$ for given h_1 and satisfies

$$E\{V_1(H_1)\} = E\left[E\{Y^*(d^{opt})|H_1\}\right] = E\{Y^*(d^{opt})\} = \mathcal{V}(d^{opt})$$

Optimal treatment option vs. optimal decision

For a randomly chosen individual with history H_1:

- The *optimal option* for this individual is

$$\underset{a_1 \in \mathcal{A}_1}{\arg\max}\, Y^*(a_1)$$

  corresponding to the largest (potential) outcome he can achieve
- Potential outcomes are not known at time of treatment decision, so this option is unknown in practice
- All that is known at the time of the decision is H_1, and $d_1^{opt}(H_1)$ selects the option corresponding to the largest expected outcome given knowledge of this history
- Because $Y^*(d) \leq \max\{Y^*(1), Y^*(0)\}$ for all $d \in \mathcal{D}$,

$$Y^*(d^{opt}) \leq \max\{Y^*(1), Y^*(0)\}$$

  so an optimal regime might not select the optimal option
- Rather, d^{opt} dictates the *optimal decision* that can be made given what is known at the time of the decision

Characterization in terms of observed data

This characterization is in terms of potential outcomes:

- To estimate an optimal regime in practice, it must be possible to identify an optimal regime from the observed data (X, A, Y)

Optimal regime in terms of observed data: Under SUTVA, NUC, positivity assumption, for any $a_1 \in \mathcal{A}_1$

$$E\{Y^*(a_1)|H_1\} = E\{Y^*(a_1)|H_1, A_1 = a_1\} = E(Y|H_1, A_1 = a_1) = Q_1(h_1, a_1)$$

Applying this to (3.20) yields the equivalent representation

$$d_1^{opt}(h_1) = \arg\max_{a_1 \in \mathcal{A}_1} E(Y|H_1 = h_1, A_1 = a_1) = \arg\max_{a_1 \in \mathcal{A}_1} Q_1(h_1, a_1) \quad (3.24)$$

$$d_1^{opt}(h_1) = \mathrm{I}\{Q_1(h_1, 1) > Q_1(h_1, 0)\} \quad \text{for } \mathcal{A}_1 = \{0, 1\} \quad (3.25)$$

$$\text{and} \quad \mathcal{V}(d^{opt}) = E\{V_1(H_1)\} = E\left\{\max_{a_1 \in \mathcal{A}_1} Q_1(H_1, a_1)\right\}$$

3. Single Decision Treatment Regimes: Fundamentals

Regression-based estimation

Obvious approach: To estimate an optimal regime d^{opt} from i.i.d. observed data (X_{1i}, A_{1i}, Y_i), $i = 1, \ldots, n$, under SUTVA, NUC, and positivity, (3.24) and (3.25) suggest

- Posit a parametric model $Q_1(h_1, a_1; \beta_1)$ (linear, logistic, etc, depending on Y), e.g., for continuous Y, $\mathcal{A}_1 = \{0, 1\}$,

$$Q_1(h_1, a_1; \beta_1) = \beta_{11} + \beta_{12}^T h_1 + \beta_{13} a_1 + \beta_{14}^T h_1 a_1 = \beta_{11} + \beta_{12}^T h_1 + (\beta_{13} + \beta_{14}^T h_1) a_1$$

  and obtain $\widehat{\beta}_1$ by an M-estimation technique
- Assuming a correct model, obtain the estimated rule

$$\widehat{d}_{Q,1}^{opt}(h_1) = \arg\max_{a_1 \in \mathcal{A}_1} Q_1(h_1, a_1; \widehat{\beta}_1)$$

  which for $\mathcal{A}_1 = \{0, 1\}$ and option 0 the default is

$$\widehat{d}_{Q,1}^{opt}(h_1) = \mathrm{I}\{Q_1(h_1, 1; \widehat{\beta}_1) > Q_1(h_1, 0; \widehat{\beta}_1)\}$$

Regression-based estimation

Regression-based estimators for d^{opt} and $\mathcal{V}(d^{opt})$:

$$\widehat{d}_Q^{opt} = \{\widehat{d}_{Q,1}^{opt}(h_1)\}, \tag{3.26}$$

$$\widehat{\mathcal{V}}_Q(d^{opt}) = n^{-1}\sum_{i=1}^{n} \max_{a_1 \in \mathcal{A}_1} Q_1(H_{1i}, a_1; \widehat{\beta}_1)$$

Example: Linear model with $\mathcal{A}_1 = \{0, 1\}$

$$\widehat{d}_{Q,1}^{opt}(h_1) = \mathrm{I}(\widehat{\beta}_{13} + \widehat{\beta}_{14}^T h_1 > 0)$$

$$\max_{a_1 \in \mathcal{A}_1} Q_1(H_1, a_1; \widehat{\beta}_1) = \widehat{\beta}_{11} + \widehat{\beta}_{12}^T H_1 + (\widehat{\beta}_{13} + \widehat{\beta}_{14}^T H_1)\mathrm{I}(\widehat{\beta}_{13} + \widehat{\beta}_{14}^T H_1 > 0)$$

$$\widehat{\mathcal{V}}_Q(d^{opt}) = n^{-1}\sum_{i=1}^{n}\Big\{\widehat{\beta}_{11} + \widehat{\beta}_{12}^T H_{1i} + (\widehat{\beta}_{13} + \widehat{\beta}_{14}^T H_{1i})\mathrm{I}(\widehat{\beta}_{13} + \widehat{\beta}_{14}^T H_{1i} > 0)\Big\}$$

Regression-based estimation

Terminology: The regression-based approach to estimation of an optimal regime and its value is a special case in the single decision setting of the method of *Q-learning* for estimation of an optimal multiple decision regime and its value

Large sample approximation: As for estimators for the value of a fixed $d \in \mathcal{D}$, would like large sample properties of $\widehat{\mathcal{V}}_Q(d^{opt})$

- First thought: View $\widehat{\mathcal{V}}_Q(d^{opt})$ and OLS $\widehat{\beta}_1$ as solving stacked estimating equations and use usual M-estimation theory

$$\sum_{i=1}^{n} \left\{ \max_{a_1 \in \mathcal{A}_1} Q_1(H_{1i}, a_1; \beta_1) - \mathcal{V}(d^{opt}) \right\} = 0 \tag{3.27}$$

$$\sum_{i=1}^{n} \frac{\partial Q_1(H_{1i}, A_i; \beta_1)}{\partial \beta_1} \{ Y_i - Q_1(H_{1i}, A_i; \beta_1) \} = 0 \tag{3.28}$$

Nonregularity of $\widehat{\mathcal{V}}_Q(d^{opt})$

Difficulty: The max operator in (3.27)

- Recall: The standard M-estimation argument to demonstrate asymptotic normality is based on a linear Taylor series
- This argument implicitly assumes differentiability of the estimating function with respect to its parameters
- The max operator is not differentiable everywhere

Demonstration in a simple special case: H_1 is one-dimensional, $\mathcal{A}_1 = \{0, 1\}$ with option 0 the default, and correctly specified model

$$Q_1(h_1, a_1; \beta_1) = \beta_{11} + \beta_{12} h_1 + \beta_{13} a_1 \qquad (3.29)$$

with true value $\beta_{1,0} = (\beta_{11,0}, \beta_{12,0}, \beta_{13,0})^T$, $\widehat{\beta}_1$ solves the OLS estimating equation (3.28)

Nonregularity of $\widehat{\mathcal{V}}_Q(d^{opt})$

Thus: When $\beta_{13} = 0$ (null hypothesis of no treatment difference)

$$\max_{a_1 \in \mathcal{A}_1} Q_1(h_1, a_1; \beta_1) = \beta_{11} + \beta_{12} h_1 + \beta_{13} \mathrm{I}(\beta_{13} > 0) \tag{3.30}$$

is *not differentiable* in β_{13}

- Because (3.29) is correctly specified

$$\begin{aligned} \mathcal{V}(d^{opt}) &= E\left\{ \max_{a_1 \in \mathcal{A}_1} Q_1(H_1, a_1; \beta_{1,0}) \right\} \\ &= \beta_{11,0} + \beta_{12,0}\, E(H_1) + \beta_{13,0} \mathrm{I}(\beta_{13,0} > 0) \end{aligned}$$

- And the estimator for $\mathcal{V}(d^{opt})$ is

$$\begin{aligned} \widehat{\mathcal{V}}_Q(d^{opt}) &= n^{-1} \sum_{i=1}^{n} \left\{ \widehat{\beta}_{11} + \widehat{\beta}_{12} H_{1i} + \widehat{\beta}_{13} \mathrm{I}(\widehat{\beta}_{13} > 0) \right\} \\ &= \widehat{\beta}_{11} + \widehat{\beta}_{12} \overline{H}_1 + \widehat{\beta}_{13} \mathrm{I}(\widehat{\beta}_{13} > 0) \quad \overline{H}_1 = n^{-1} \sum_{i=1}^{n} H_{1i}, \end{aligned}$$

Nonregularity of $\widehat{\mathcal{V}}_Q(d^{opt})$

$$
\begin{aligned}
& n^{1/2}\left\{\widehat{\mathcal{V}}_Q(d^{opt}) - \mathcal{V}(d^{opt})\right\} \\
& \quad = n^{1/2}(\widehat{\beta}_{11} - \beta_{11,0}) + n^{1/2}(\widehat{\beta}_{12} - \beta_{12,0})E(H_1) && (3.31) \\
& \quad + n^{1/2}(\widehat{\beta}_{12} - \beta_{12,0})\{\overline{H}_1 - E(H_1)\} && (3.32) \\
& \quad + n^{1/2}\{\widehat{\beta}_{13} I(\widehat{\beta}_{13} > 0) - \beta_{13,0} I(\beta_{13,0} > 0)\} && (3.33)
\end{aligned}
$$

and by the usual M-estimation theory

$$
n^{1/2}\begin{pmatrix} \widehat{\beta}_{11} - \beta_{11,0} \\ \widehat{\beta}_{12} - \beta_{12,0} \\ \widehat{\beta}_{13} - \beta_{13,0} \end{pmatrix} \xrightarrow{\mathcal{D}} \begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \end{pmatrix} \sim \mathcal{N}(0, \Sigma)
$$

- (3.32) $\xrightarrow{p}$ 0 because $\overline{H}_1 - E(H_1) \xrightarrow{p} 0$
- Terms in (3.31) $\xrightarrow{\mathcal{D}}$ Z_1 and $Z_2 E(H_1)$

Nonregularity of $\widehat{\mathcal{V}}_Q(d^{opt})$

Term (3.33): $n^{1/2}\{\widehat{\beta}_{13} I(\widehat{\beta}_{13} > 0) - \beta_{13,0} I(\beta_{13,0} > 0)\}$

- $g(u) = u\,\mathrm{I}(u > 0)$ continuous in u but not differentiable at $u = 0$

Case 1: $\beta_{13,0} \neq 0$: $g(u)$ is differentiable in an open interval containing $\beta_{13,0}$, and standard Taylor series can be used to show

$$(3.33) \xrightarrow{\mathcal{D}} g'(\beta_{13,0}) Z_3, \quad g'(\beta_{13,0}) = \{dg(u)/du\}|_{u=\beta_{13,0}} = \mathrm{I}(\beta_{13,0} > 0)$$

- Thus, $n^{1/2}(\widehat{\beta}_{11} - \beta_{11,0})$, $n^{1/2}(\widehat{\beta}_{12} - \beta_{12,0})$, and (3.33) jointly $\xrightarrow{\mathcal{D}}$ $\{Z_1, Z_2, Z_3 \mathrm{I}(\beta_{13,0} > 0)\}^T$
- Continuous mapping theorem, Slutsky's theorem yield

$$n^{1/2}\left\{\widehat{\mathcal{V}}_Q(d^{opt}) - \mathcal{V}(d^{opt})\right\} \xrightarrow{\mathcal{D}} Z_1 + Z_2 E(H_1) + Z_3\,\mathrm{I}(\beta_{13,0} > 0)$$

- Linear combination of jointly $\mathcal{N}(0, \Sigma)$ random variables is normal, so $n^{1/2}\left\{\widehat{\mathcal{V}}_Q(d^{opt}) - \mathcal{V}(d^{opt})\right\}$ is asymptotically normal with this distribution

Nonregularity of $\widehat{\mathcal{V}}_Q(d^{opt})$

Term (3.33): $n^{1/2}\{\widehat{\beta}_{13} I(\widehat{\beta}_{13} > 0) - \beta_{13,0} I(\beta_{13,0} > 0)\}$

Case 2: $\beta_{13,0} = 0$: $(3.33) = n^{1/2}\widehat{\beta}_{13}\mathrm{I}(\widehat{\beta}_{13} > 0)$

- $n^{1/2}\widehat{\beta}_{13} \xrightarrow{\mathcal{D}} Z_3$
- $\mathrm{I}(\widehat{\beta}_{13} > 0) = \mathrm{I}(n^{1/2}\widehat{\beta}_{13} > 0)$, so by the continuous mapping and Slutsky's theorems

$$n^{1/2}\left\{\widehat{\mathcal{V}}_Q(d^{opt}) - \mathcal{V}(d^{opt})\right\} \xrightarrow{\mathcal{D}} Z_1 + Z_2 E(H_1) + Z_3\,\mathrm{I}(Z_3 > 0)$$

- Even though Z_1, Z_2, and Z_3 are jointly normal, the distribution of $Z_1 + Z_2 E(H_1) + Z_3 \mathrm{I}(Z_3 > 0)$ is *not normal*
- Thus: When $\beta_{13,0} = 0$, $\widehat{\mathcal{V}}(d^{opt})$ does not follow standard asymptotic theory

Nonregularity of $\widehat{\mathcal{V}}_Q(d^{opt})$

Result: $\widehat{\mathcal{V}}_Q(d^{opt})$ is an example of a *nonregular estimator*

- Although $\widehat{\mathcal{V}}_Q(d^{opt})$ follows standard asymptotic theory when $\beta_{13,0} \neq 0$, the usual large sample normal approximation to its sampling distribution is not valid when $\beta_{13,0} = 0$
- In (3.29), $\beta_{13,0} = 0$ corresponds to no difference in expected outcome between options 0 and 1 for any h_1, which cannot be ruled out in practice
- Technically, cannot disregard this behavior at $\beta_{13,0} = 0$ and appeal to standard theory to obtain measures of uncertainty
- Even if $\beta_{13,0} \neq 0$, where standard theory holds, if $\beta_{13,0}$ is close to zero, using the standard normal approximation can be poor

General phenomenon: Due to nonsmoothness of the max operator

- Because finding d^{opt} involves a max operation, all estimators for d^{opt} and $\mathcal{V}(d^{opt})$ are subject to this issue
- Nonstandard inferential approaches are required

Estimation via A-learning

Consider $\mathcal{A}_1 = \{0, 1\}$: An optimal regime has rule

$$d_1^{opt}(h_1) = \text{I}\{Q_1(h_1, 1) > Q_1(h_1, 0)\} = \text{I}\{Q_1(h_1, 1) - Q_1(h_1, 0) > 0\}$$

Definition: The *contrast function* is given by

$$C_1(h_1) = Q_1(h_1, 1) - Q_1(h_1, 0) \tag{3.34}$$

- Thus, the rule $d_1^{opt}(h_1)$ can be written as

$$d_1^{opt}(h_1) = \text{I}\{C_1(h_1) > 0\} \tag{3.35}$$

- From (3.35), full knowledge of $Q_1(h_1, a_1)$ is not required to characterize and estimate an optimal regime
- Premise of the class of methods for estimation of an optimal regime referred to as *advantage* or *A-learning*

Estimation via A-learning

Because a_1 is binary: Any arbitrary function $Q_1(h_1, a_1)$ can be written as

$$Q_1(h_1, a_1) = \nu_1(h_1) + a_1 C_1(h_1), \quad \nu_1(h_1) = Q_1(h_1, 0) \tag{3.36}$$

- (3.36) shows $Q_1(h_1, a_1)$ is maximized by $a_1 = \text{I}\{C_1(h_1) > 0\}$ with maximum

$$V_1(h_1) = \nu_1(h_1) + C_1(h_1)\text{I}\{C_1(h_1) > 0\}$$

- Robins (2004) refers to $Q_1(h_1, a_1) - Q_1(h_1, 0) = a_1 C_1(h_1)$ as the *optimal blip to zero function* comparing difference in expected outcome between using option 0 (control or reference option) and using a_1 among individuals with history h_1

Estimation via A-learning

Suggests: Posit a model $C_1(h_1; \psi_1)$ for the contrast function; equivalently, a semiparametric model for $Q_1(h_1, a_1)$

$$\nu_1(h_1) + a_1 C_1(h_1; \psi_1) \tag{3.37}$$

for arbitrary function $\nu_1(h_1)$ of h_1 and finite-dimensional parameter ψ_1

- May be more robust to misspecification than the regression method, as only $C_1(h_1; \psi_1)$ must be correctly specified for valid estimation of d^{opt}
- (3.37) preserves the *causal null hypothesis* because $C_1(h_1; \psi_1) = 0$ implies $E\{Q_1(H_1, 1) - Q_1(H_1, 0)\} = 0$

Goal: Estimate ψ_1 in (3.37) based on (X_{1i}, A_{1i}, Y_i), $i = 1, \ldots, n$ and substitute fitted contrast function in (3.35)

Estimation via A-learning

G-estimation: By semiparametric theory, Robins (2004) showed that all consistent and asymptotically normal estimators for ψ_1 solve an estimating equation of form

$$\sum_{i=1}^{n} \lambda_1(H_{1i}) \{A_{1i} - \pi_1(H_{1i})\} \{Y_i - A_{1i} C_1(H_{1i}; \psi_1) + \theta_1(H_{1i})\} = 0 \tag{3.38}$$

for arbitrary dim(ψ_1)-dimensional $\lambda_1(h_1)$ and real-valued $\theta_1(h_1)$

- Can show: The estimating function in (3.38) is unbiased; i.e.,

  $$E_{\psi_1}[\lambda_1(H_1) \{A_1 - \pi_1(H_1)\}\{Y - A_1 C_1(H_1; \psi_1) + \theta_1(H_1)\}] = 0$$

  so that $\widehat{\psi}_1$ solving (3.38) is an M-estimator
- When var($Y|H_1, A_1$) is constant, optimal choices

  $$\lambda_1(h_1) = \partial C_1(h_1; \psi_1)/\partial\psi_1 \quad \text{and} \quad \theta_1(h_1) = -\nu_1(h_1)$$

Estimation via A-learning

Suggests: $\nu_1(h_1)$ is arbitrary, but can proceed adaptively and posit a model $\nu_1(h_1;\phi_1)$ for parameter ϕ_1 and estimate ϕ_1 jointly with ψ_1 jointly by solving

$$\sum_{i=1}^{n} \frac{\partial C_1(H_{1i};\psi_1)}{\partial \psi_1} \{A_{1i} - \pi_1(H_{1i})\}$$
$$\times \{Y_i - A_{1i}C_1(H_{1i};\psi_1) - \nu_1(H_{1i};\phi_1)\} = 0$$
$$\sum_{i=1}^{n} \frac{\partial \nu_1(H_{1i};\phi_1)}{\partial \phi_1} \{Y_i - A_{1i}C_1(H_{1i};\psi_1) - \nu_1(H_{1i};\phi_1)\} = 0$$

- Can show: If $C_1(h_1;\psi_1)$ is correctly specified, with true value $\psi_{1,0}$, but $\nu_1(h_1;\phi_1)$ is not, $\widehat{\psi}_1$ solving these equations is consistent for $\psi_{1,0}$
- And thus $C_1(h_1;\widehat{\psi}_1)$ is consistent for $C_1(h_1)$

Estimation via A-learning

Unknown $\pi_1(h_1)$: Posit a model $\pi_1(h_1;\gamma_1)$ (e.g., logistic) and jointly solve in $(\psi_1^T, \phi_1^T, \gamma_1^T)^T$ the stacked estimating equations

$$\sum_{i=1}^{n} \frac{\partial C_1(H_{1i};\psi_1)}{\partial \psi_1} \{A_{1i} - \pi_1(H_{1i};\gamma_1)\} \times \{Y_i - A_{1i}C_1(H_{1i};\psi_1) - \nu_1(H_{1i};\phi_1)\} = 0$$

$$\sum_{i=1}^{n} \frac{\partial \nu_1(H_{1i};\phi_1)}{\partial \phi_1} \{Y_i - A_{1i}C_1(H_{1i};\psi_1) - \nu_1(H_{1i};\phi_1)\} = 0$$

$$\sum_{i=1}^{n} \begin{pmatrix} 1 \\ H_{1i} \end{pmatrix} \left\{ A_{1i} - \frac{\exp(\gamma_{11} + \gamma_{12}^T H_{1i})}{1 + \exp(\gamma_{11} + \gamma_{12}^T H_{1i})} \right\} = 0$$

- Can show: If $C_1(h_1;\psi_1)$ is correctly specified but either $\pi_1(H_1;\gamma_1)$ or $\nu_1(H_1;\phi_1)$ (but not both) is misspecified, $\widehat{\psi}_1$ solving these equations is consistent for $\psi_{1,0}$ so is doubly robust in this sense

Estimation via A-learning

Estimator for d^{opt}: Given $\widehat{\psi}_1$, from (3.35), estimate d^{opt} by

$$\widehat{d}_A^{opt} = \{\widehat{d}_{A,1}^{opt}(h_1)\}, \quad \widehat{d}_{A,1}^{opt}(h_1) = \mathrm{I}\{C_1(h_1; \widehat{\psi}_1) > 0\} \tag{3.39}$$

- Alternative approach: Murphy (2003) instead propose an A-learning approach based on the *advantage or regret function*

$$C_1(H_1)\,[\mathrm{I}\{C_1(H_1) > 0\} - A_1]$$

- Can show: If $\pi_1(h_1) = P(A_1 = 1 | H_1 = h_1)$ does not depend on h_1, $Q_1(h_1, a_1; \beta_1)$ is linear in h_1, and $\nu_1(h_1; \phi_1) + a_1 C_1(h_1; \psi_1)$ is of the same form, A-learning and Q-learning are identical

Estimation via A-learning

Recall: $\mathcal{V}(d^{opt}) = E\{V_1(H_1)\} = E\left\{\max_{a_1 \in \mathcal{A}_1} Q(H_1, a_1)\right\}$

$$
\begin{aligned}
&E\Big(Y + C_1(H_1)[\mathrm{I}\{C_1(H_1) > 0\} - A_1]\Big|H_1\Big) \\
&\quad = E\Big\{E\Big(Y + C_1(H_1)[\mathrm{I}\{C_1(H_1) > 0\} - A_1]\Big|H_1, A_1\Big)\Big|H_1\Big\} \\
&\quad = E\Big(E(Y|H_1, A_1) + C_1(H_1)[\mathrm{I}\{C_1(H_1) > 0\} - A_1]\Big|H_1\Big) \\
&\quad = E\Big(Q_1(H_1, 0) + A_1 C_1(H_1) + C_1(H_1)[\mathrm{I}\{C_1(H_1) > 0\} - A_1]\Big|H_1\Big) \\
&\quad = E[Q_1(H_1, 0) + C_1(H_1)\mathrm{I}\{C_1(H_1) > 0\}|H_1] \\
&\quad = Q_1(H_1, 0) + C_1(H_1)\mathrm{I}\{C_1(H_1) > 0\} = V_1(H_1)
\end{aligned}
$$

- Suggests the estimator for $\mathcal{V}(d^{opt})$

$$
\widehat{\mathcal{V}}_A(d^{opt}) = n^{-1} \sum_{i=1}^{n} \Big(Y_i + C_1(H_{1i}; \widehat{\psi}_1)\Big[\mathrm{I}\{C_1(H_{1i}; \widehat{\psi}_1) > 0\} - A_{1i}\Big]\Big)
$$

Is also a nonregular estimator

Restricted class of regimes

Continue to consider $\mathcal{A}_1 = \{0, 1\}$: For the previous approaches, the form $Q_1(h_1, a_1; \beta_1)$ or $C_1(h_1; \psi_1)$ dictates the form of the rules d_1^{opt}

- Example: With $h_1 = x_1 = (x_{11}, x_{12})^T$

$$Q_1(h_1, a_1; \beta_1) = \beta_{11} + \beta_{12} x_{11} + \beta_{13} x_{12} + a_1(\beta_{14} + \beta_{15} x_{11} + \beta_{16} x_{12}) \quad (3.40)$$

  implies rules d_1^{opt} of form

$$d_1^{opt}(h_1) = \mathrm{I}(\beta_{14} + \beta_{15} x_{11} + \beta_{16} x_{12} > 0)$$

  and similarly for a linear contrast function
- Result: Posited models for regression or constrast function induce a *class of regimes*, indexed by parameters in the model, to which the search for d^{opt} is restricted
- In the example, the *restricted class* $\mathcal{D}_\eta$ indexed by η is the class of regimes with rules of the form

$$d_1(h_1; \eta_1) = \mathrm{I}(\eta_{11} + \eta_{12} x_{11} + \eta_{13} x_{12} > 0), \quad \eta_1 = (\eta_{11}, \eta_{12}, \eta_{13})^T, \quad \eta = \eta_1 \quad (3.41)$$

Effect of model misspecification

$\mathcal{D}_\eta$ may or may not contain $d^{opt} \in \mathcal{D}$:

- Example, continued: Suppose the true regression relationship is
  $$Q_1(h_1, a_1) = \exp\{1 + x_{11} + 2x_{12} + 3x_{11}x_{12} + a_1(1 - 2x_{11} + x_{12})\}$$
  so that
  $$d_1^{opt}(h_1) = \mathrm{I}(1 - 2x_{11} + x_{12} > 0),$$
  which is of the form (3.41)
- Here, although the model (3.40) is misspecified, $d^{opt} \in \mathcal{D}_\eta$
- However: If we fit (3.40) by OLS, $\widehat{d}_Q^{opt}$ with
  $$\widehat{d}_{Q,1}^{opt}(h_1) = \mathrm{I}\{\widehat{\beta}_{14} + \widehat{\beta}_{15}x_{11} + \widehat{\beta}_{16}x_{12}) > 0\}$$
  may be a poor estimator for d^{opt} because $\widehat{\beta}$ is likely far from the values of the coefficients in the true relationship
- Of course: If $Q_1(h_1, a_1; \beta_1)$ or $C_1(h_1; \psi_1)$ does not imply a restricted class containing d^{opt}, the estimated regime can be quite far from d^{opt}

Alternative perspective

Suggests: Deliberately restrict attention to a class $\mathcal{D}_\eta \subset \mathcal{D}$ of regimes d_η with rules of form $d_1(h_1; \eta_1)$

- $\mathcal{D}_\eta$ may be chosen based on cost, feasibility in practice, *interpretability* (by clinicians and patients)
- E.g., with $h_1 = (x_{11}, x_{12})^T$, $\mathcal{D}_\eta$ comprises regimes with rules

  $$d_1(h_1; \eta_1) = \mathrm{I}(x_{11} < \eta_{11}, x_{12} < \eta_{12}), \qquad \eta_1 = (\eta_{11}, \eta_{12})^T$$

  involving rectangular regions with thresholds
- Or rules involving linear combinations as in (3.41)
- $\mathcal{D}_\eta$ may or may not contain d^{opt} but still of interest
- This perspective of course extends to general $\mathcal{A}_1$ with > 2 options

Value search estimation

Optimal restricted regime: $d_\eta^{opt} \in \mathcal{D}_\eta$ with rule

$$d_1(h_1; \eta_1^{opt}), \quad \eta_1^{opt} = \arg\max_{\eta_1} \mathcal{V}(d_\eta), \tag{3.42}$$

$$d_\eta^{opt} = \{d_1(h_1; \eta_1^{opt})\}$$

Approach: Given an estimator $\widehat{\mathcal{V}}(d)$ for the value of fixed $d \in \mathcal{D}$

- Estimate $\mathcal{V}(d_\eta)$ by $\widehat{\mathcal{V}}(d_\eta)$ for fixed $\eta = \eta_1$
- Regard $\widehat{\mathcal{V}}(d_\eta)$ as a function of η_1, maximize in η_1 to obtain

$$\widehat{\eta}_1^{opt} = \arg\max_{\eta_1} \widehat{\mathcal{V}}(d_\eta)$$

  and estimate d_η^{opt} by

$$\widehat{d}_\eta^{opt} = \{d_1(h_1, \widehat{\eta}_1^{opt})\}$$

- *Value search* or *policy or direct search* estimation

Value search estimation

Natural choices for $\widehat{\mathcal{V}}(d_\eta)$: IPW or AIPW estimators

- Analogous to (3.8) and (3.9) define for fixed $\eta = \eta_1$

$$\mathcal{C}_{d_\eta} = \mathrm{I}\{A_1 = d_1(H_1; \eta_1)\}$$

$$\pi_{d_\eta,1}(H_1; \eta_1, \gamma_1)$$
$$= \pi_1(H_1; \gamma_1)\mathrm{I}\{d_1(H_1; \eta_1) = 1\} + \{1 - \pi_1(H_1; \gamma_1)\}\mathrm{I}\{d_1(H_1; \eta_1) = 0\}$$

- From (3.13) IPW estimator

$$\widehat{\mathcal{V}}_{IPW}(d_\eta) = n^{-1} \sum_{i=1}^{n} \frac{\mathcal{C}_{d_\eta,i} Y_i}{\pi_{d_\eta,1}(H_{1i}; \eta_1, \widehat{\gamma}_1)} \tag{3.43}$$

Value search estimation

- From (3.18), AIPW estimator

$$\widehat{\mathcal{V}}_{AIPW}(d_\eta) \tag{3.44}$$
$$= n^{-1} \sum_{i=1}^{n} \left[\frac{\mathcal{C}_{d_\eta,i} Y_i}{\pi_{d_\eta,1}(H_{1i}; \eta_1, \widehat{\gamma}_1)} - \frac{\mathcal{C}_{d_\eta,i} - \pi_{d_\eta,1}(H_{1i}; \eta_1, \widehat{\gamma}_1)}{\pi_{d_\eta,1}(H_{1i}; \eta_1, \widehat{\gamma}_1)} \mathcal{Q}_{d_\eta,1}(H_{1i}; \eta_1, \widehat{\beta}_1) \right]$$

$$\mathcal{Q}_{d_\eta,1}(H_1; \eta_1, \beta_1)$$
$$= Q_1(H_1, 1; \beta_1) \mathrm{I}\{d_1(H_1; \eta_1) = 1\} + Q_1(H_1, 0; \beta_1) \mathrm{I}\{d_1(H_1; \eta_1) = 0\}$$

- Also: Alternative estimator $\widehat{\mathcal{V}}_{IPW*}(d_\eta)$
- As before, $\widehat{\mathcal{V}}_{IPW}(d_\eta)$ and $\widehat{\mathcal{V}}_{AIPW}(d_\eta)$ are consistent estimators for $\mathcal{V}(d_\eta)$ for fixed $\eta = \eta_1$, and $\widehat{\mathcal{V}}_{AIPW}(d_\eta)$ is moreover doubly robust

Value search estimation

Result: Estimators for η_1^{opt} by maximizing $\widehat{\mathcal{V}}_{IPW}(d_\eta)$ or $\widehat{\mathcal{V}}_{AIPW}(d_\eta)$ in η_1 to obtain $\widehat{\eta}_{1,IPW}^{opt}$ or $\widehat{\eta}_{1,AIPW}^{opt}$

- Estimators for optimal restricted regime $d_\eta^{opt} \in \mathcal{D}_\eta$

$$\widehat{d}_{\eta,IPW}^{opt} = \{d_1(h_1, \widehat{\eta}_{1,IPW}^{opt})\} \quad \text{and} \quad \widehat{d}_{\eta,AIPW}^{opt} = \{d_1(h_1, \widehat{\eta}_{1,AIPW}^{opt})\}$$

- Estimators for $\mathcal{V}(d_\eta^{opt})$ by substituting $\widehat{\eta}_{1,IPW}^{opt}$ or $\widehat{\eta}_{1,AIPW}^{opt}$ for η_1 in (3.43) or (3.44) to yield estimators $\widehat{\mathcal{V}}_{IPW}(d_\eta^{opt})$ and $\widehat{\mathcal{V}}_{AIPW}(d_\eta^{opt})$
- Challenge: Maximization of (3.43) or (3.44) is a *nonsmooth* optimization problem; standard optimization techniques cannot be used
- Intuition: $\widehat{d}_{\eta,AIPW}^{opt}$ should be of higher quality than $\widehat{d}_{\eta,IPW}^{opt}$ because $\widehat{\mathcal{V}}_{AIPW}(d_\eta)$ is more efficient and stable than $\widehat{\mathcal{V}}_{IPW}(d_\eta)$
- Similarly, expect $\widehat{\mathcal{V}}_{AIPW}(d_\eta^{opt})$ to be more efficient than $\widehat{\mathcal{V}}_{IPW}(d_\eta^{opt})$

Nonregularity of $\widehat{\mathcal{V}}_{IPW}(d_\eta^{opt})$ and $\widehat{\mathcal{V}}_{AIPW}(d_\eta^{opt})$

Not surprisingly: $\widehat{\mathcal{V}}_{IPW}(d_\eta^{opt})$ and $\widehat{\mathcal{V}}_{AIPW}(d_\eta^{opt})$ are nonregular estimators

- But because of maximization, cannot be cast as solving stacked M-estimating equations, so cannot show by an argument similar to that for $\widehat{\mathcal{V}}_Q(d^{opt})$

Instead: Nonstandard theory suggested by behavior of the true value $\mathcal{V}(d_\eta^{opt})$ in a simple example

- $H_1 \sim \mathcal{N}(0,1)$, $\mathcal{D}_\eta$ comprises regimes with rules

$$d_1(h_1;\eta) = \mathrm{I}(h_1 > \eta_1), \quad \eta_1 \in \mathbb{R}$$

- Y continuous with true regression relationship

$$\begin{aligned} Q_1(h_1, a_1) &= E(Y|H_1 = h_1, A_1 = a_1) \\ &= \beta_{11,0} + \beta_{12,0} h_1 + \beta_{13,0} a_1 + \beta_{14,0} h_1 a_1 \end{aligned}$$

  with $\beta_{14,0} \geq 0$

Nonregularity of $\widehat{\mathcal{V}}_{IPW}(d_\eta^{opt})$ and $\widehat{\mathcal{V}}_{AIPW}(d_\eta^{opt})$

True value for fixed $\eta = \eta_1$:

$$\begin{aligned}\mathcal{V}(d_\eta) = E\Big[&\{\beta_{11,0} + \beta_{12,0}H_1 + \beta_{13,0} + \beta_{14,0}H_1\}\mathrm{I}(H_1 > \eta_1) \\ &+ \{\beta_{11,0} + \beta_{12,0}H_1\}\{1 - \mathrm{I}(H_1 > \eta_1)\}\Big] \\ = \beta_{11,0} &+ \beta_{12,0}E(H_1) + \beta_{13,0}E\{\mathrm{I}(H_1 > \eta_1)\} + \beta_{14,0}E\{H_1\mathrm{I}(H_1 > \eta_1)\} \\ = \beta_{11,0} &+ \beta_{13,0}\{1 - \Phi(\eta_1)\} + \beta_{14,0}\varphi(\eta_1)\end{aligned}$$

- $\Phi(\cdot)$ and $\varphi(\cdot)$ cdf and density of $\mathcal{N}(0,1)$

Nonregularity of $\widehat{\mathcal{V}}_{IPW}(d_\eta^{opt})$ and $\widehat{\mathcal{V}}_{AIPW}(d_\eta^{opt})$

Case 1: $\beta_{14,0} = 0$: $\quad \mathcal{V}(d_\eta) = \beta_{11,0} + \beta_{13,0}\{1 - \Phi(\eta_1)\}$

- If $\beta_{13,0} > 0$, $\Phi(\eta_1) \to 0$ as $\eta_1 \to -\infty$, $\mathcal{V}(d_\eta) \to$ its max, so no unique maximum in $-\infty < \eta_1 < \infty$, and all individuals receive option 1
- Similarly, if $\beta_{13,0} < 0$, $\Phi(\eta_1) \to 1$ as $\eta_1 \to \infty$, and all individuals receive option 0
- If $\beta_{13,0} = 0$, $\mathcal{V}(d_\eta)$ is constant with no unique maximum, treatment selection ambiguous
- Result: $\mathcal{V}(d_\eta)$ does not have a unique maximum in η_1, so η_1^{opt} and thus d_η^{opt} are not well defined, and standard asymptotic theory does not apply to $\widehat{\mathcal{V}}_{IPW}(d_\eta^{opt})$ and $\widehat{\mathcal{V}}_{AIPW}(d_\eta^{opt})$

Nonregularity of $\widehat{\mathcal{V}}_{IPW}(d_\eta^{opt})$ and $\widehat{\mathcal{V}}_{AIPW}(d_\eta^{opt})$

Case 2: $\beta_{14,0} > 0$**:** $\mathcal{V}(d_\eta)$ is a smooth function in η_1 with

$$\partial \mathcal{V}(d_\eta)/\partial \eta_1 = -(\beta_{13,0} + \beta_{14,0}\eta_1)\varphi(\eta_1) \tag{3.45}$$

$$\partial^2 \mathcal{V}(d_\eta)/\partial \eta_1^2 = (\beta_{14,0}\eta_1^2 + \beta_{13,0}\eta_1 - \beta_{14,0})\varphi(\eta_1) \tag{3.46}$$

- Setting (3.45) = 0 yields $\eta_1 = -\beta_{13,0}/\beta_{14,0}$, at which (3.46) < 0
- So $\mathcal{V}(d_\eta)$ has a unique maximum, and thus η_1^{opt}, $d_\eta^{opt} \in \mathcal{D}_\eta$, and $\mathcal{V}(d_\eta^{opt})$ are well defined
- Standard asymptotic theory applies

However: Standard theory does not apply for all $\beta_{14,0} \geq 0$

- Zhang et al. (2012): Apply standard asymptotic theory anyway when in a "Case 2" situation
- Can work well under this condition, but can fail if not or if $\beta_{14,0} \neq 0$ but is close to 0

Discussion

Implementation:

- Regression methods/Q-learning straightforward using established methods and software
- A-learning methods similarly
- Value search methods involve maximization of nonsmooth objective functions, require special techniques; e.g., a genetic algorithm (as in R `rgenoud`) or grid search, becomes untenable for η_1 of higher dimension
- Q-learning and value search are available in R package `DynTxRegime`

Discussion

Practical performance: Estimation of an optimal regime

- No uniformly "best" method
- $\widehat{d}^{opt}_{Q}$ can achieve performance of true optimal regime if d^{opt} is in the induced class of regimes, but can be very poor if the outcome regression model is misspecified
- $\widehat{d}^{opt}_{\eta,AIPW}$ is comparable if $d^{opt} \in \mathcal{D}_{\eta}$ if $Q_1(h_1, a_1; \beta_1)$ is correct, even if $\pi_1(h_1; \gamma_1)$ is misspecified
- And $\widehat{d}^{opt}_{\eta,AIPW}$ is much better than $\widehat{d}^{opt}_{Q}$ when the regression model is misspecified but propensity is correct
- $\widehat{d}^{opt}_{\eta,IPW}$ not recommends on inefficiency and instability grounds
- Schulte, Tsiatis, Laber, and Davidian (2014) compare Q- and A-learning

More than two treatment options

$\mathcal{A}_1 = \{1, \ldots, m_1\}$: With appropriate versions of SUTVA, NUC, positivity

- Outcome regression methods require no modification, just a suitable model $Q_1(h_1, a_1; \beta_1)$
- Inverse probability weighted methods: With

$$\omega_1(h_1, a_1) = P(A_1 = a_1 | H_1 = h_1), \quad \omega_1(h_1, m_1) = 1 - \sum_{a_1=1}^{m_1-1} \omega_1(h_1, a_1)$$

can adopt a multinomial (polytomous) logistic model, e.g.,

$$\omega_1(h_1, a_1; \gamma_1) = \frac{\exp(\widetilde{h}_1^T \gamma_{1,a_1})}{1 + \sum_{j=1}^{m_1-1} \exp(\widetilde{h}_1^T \gamma_{1j})}, \quad a_1 = 1, \ldots, m_1 - 1$$

$\widetilde{h}_1 = (1, h_1^T)^T$, $\gamma_1 = (\gamma_{11}^T, \ldots, \gamma_{1,m_1-1}^T)^T$

More than two treatment options

- Redefine

$$\mathcal{C}_{d_\eta} = \mathrm{I}\{A_1 = d_1(H_1;\eta_1)\}$$

$$\pi_{d_\eta,1}(H_1;\eta_1,\gamma_1) = \sum_{a_1=1}^{m_1} \mathrm{I}\{d_1(H_1;\eta_1) = a_1\}\,\omega_1(H_1,a_1;\gamma_1)$$

$$\mathcal{Q}_{d_\eta,1}(H_1;\eta_1,\beta_1) = \sum_{a_1=1}^{m_1} \mathrm{I}\{d_1(H_1;\eta_1) = a_1\}\,Q_1(H_1,a_1;\beta_1)$$

- AIPW estimator with these definitions

$$\widehat{\mathcal{V}}_{AIPW}(d_\eta)$$
$$= n^{-1}\sum_{i=1}^{n}\left[\frac{\mathcal{C}_{d_\eta,i}Y_i}{\pi_{d_\eta,1}(H_{1i};\eta_1,\widehat{\gamma}_1)} - \frac{\mathcal{C}_{d_\eta,i} - \pi_{d_\eta,1}(H_{1i};\eta_1,\widehat{\gamma}_1)}{\pi_{d_\eta,1}(H_{1i};\eta_1,\widehat{\gamma}_1)}\mathcal{Q}_{d_\eta,1}(H_{1i};\eta_1,\widehat{\beta}_1)\right]$$

More than two treatment options

- A-learning: Take $\mathcal{A}_1 = \{0, 1, \ldots, m_1 - 1\}$, analogous to (3.34) define

$$C_{1j}(h_1) = Q_1(h_1, j) - Q_1(h_1, 0), \quad j = 0, 1, \ldots, m_1 - 1$$

so that

$$d_1^{opt}(h_1) = \underset{j \in \{0,1,\ldots,m_1-1\}}{\arg\max} \; C_{1j}(h_1)$$

- Posit models $C_{1j}(h_1; \psi_1)$, $j = 1, \ldots, m_1 - 1$

3. Single Decision Treatment Regimes: Fundamentals

References

Moodie, E. E. M., Richardson, T. S., and Stephens, D. A. (2007). Demystifying optimal dynamic treatment regimes. *Biometrics* **63,** 447–455.

Robins, J. M., Rotnitzky, A., and Zhao, L. P. (1994). Estimation of regression coeffcients when some regressors are not always observed. *Journal of the American Statistical Association*, 89, 846–866.

Robins, J. M. (2004). Optimal structural nested models for optimal sequential decisions. In Lin, D. Y. and Heagerty, P., editors, *Pro- ceedings of the Second Seattle Symposium on Biostatistics*, pages 189–326 . Springer.

Schulte, P. J., Tsiatis, A. A., Laber, E. B., and Davidian, M. (2014). Robust estimation of optimal dynamic treatment regimes for sequential treatment decisions. *Statistical Science*, 29, 640–661.

Tsiatis, A. A. (2006). *Semiparametric Theory and Missing Data*. Springer.

Zhang, B., Tsiatis, A. A., Laber, E. B., and Davidian, M. (2012). A robust method for estimating optimal treatment regimes. *Biometrics* **68,** 1010–1018.