

Double robust estimation in Cox models using a dynamic propensity score

H. Sitter¹, K. H. Schild²

¹Institute of Theoretical Surgery, University Marburg, Germany

²Department of Economics; University Marburg, Germany

email: sitter@mail.uni-marburg.de

Introduction

- Non-randomized trials in survival analysis with varying treatment initialisation
- Standard approach: inverse probability weighting (Robins et al. 2000, logistic regression)
- Determination of causal treatment-effect by propensity score (Rosenbaum, Rubin 1983)
- Our approach: Use propensity score as additional covariate in Cox regression

Framework for Causal Inference

A = Treatment assignment (binary)

Y = Factual outcome

Y^a = Potential outcome: $Y = Y^A$

$E[Y^1 - Y^0]$ = Average causal treatment (ATE)

X = Pre-treatment covariates

CUA Conditional unconfoundedness assumption $Y^a \perp A \mid X$

P_s = Propensity Score: $P_s(\mathbf{x}) = P(A = 1 \mid X = \mathbf{x})$

Regression of factual outcome on factual treatment and propensity score

CUA implies for this regression

$$\mathbf{E}[Y|A = a, P = p] = \mathbf{E}[Y^a|P = p]$$

From this:

Regression-coefficient of P measures selection bias, because:

$$\mathbf{E}[Y|A = 0, P = p_1] - \mathbf{E}[Y|A = 0, P = p_0] = \mathbf{E}[Y^0|P = p_1] - \mathbf{E}[Y^0|P = p_0]$$

Regression-coefficient of factual treatment assignment A measures causal treatment effect, because:

$$\mathbf{E}[Y|A = 1, P = p] - \mathbf{E}[Y|A = 0, P = p] = \mathbf{E}[Y^1 - Y^0|P = p]$$

Application to survival analysis

- Assume now:
 - T survival time (possibly right-censored) is outcome
 - X covariates (possibly time-dependent)
- For each individual there is a treatment initiation time S
- Assume: proportional hazards model, if treatment starts (possibly counterfactually) at s

$$\mathbf{h}(t, s | \mathbf{X}_i) = e^{\beta' \mathbf{X}_i + \alpha 1_{[t \geq s]}} \mathbf{h}_0(t) = \begin{cases} e^{\beta' \mathbf{X}_i} \mathbf{h}_0(t) & \text{if } t < s \\ e^{\alpha} e^{\beta' \mathbf{X}_i} \mathbf{h}_0(t) & \text{if } t \geq s \end{cases}$$

(eqn. 1)

Objective

Estimation of causal treatment effect α by Cox regression (for factual outcome).

Note: Previous equation is the potential outcome model; it describes the outcome-hazard, if the treatment would start at (an arbitrary) time s

Why include propensity score in outcome regression ?

- Assume the CUA in the present application: $\mathbf{S} \perp \mathbf{T}^s \mid \mathbf{X} \quad \forall s$.
- In virtue of CUA: the Cox regression for (1) – with s replaced by \mathbf{S} - can be expected to provide a reliable (consistent) estimate of α , provided that the covariates \mathbf{X} are exactly those under which CUA holds.
- To avoid these problems with **wrong selection of variables**: We include the propensity score as a control function in eqn. 1.

How to include the propensity score?

- Propensity score is the prob. of receiving treatment. In the present application, this requires a model for \mathbf{S}
- Assumption: \mathbf{S} follows a proportional hazard model (let $\mathbf{h}^{\mathbf{S}}(\mathbf{t}) :=$ the ‘hazard’-rate of \mathbf{S}):

$$\mathbf{h}^{\mathbf{S}}(\mathbf{t}) = \mathbf{h}^{\mathbf{S}}(\mathbf{t}, \mathbf{X}) = e^{\gamma \mathbf{X}} \mathbf{h}_0^{\mathbf{S}}(\mathbf{t})$$

How to define the propensity score ?

- General definition is: Propensity score is the probability of receiving treatment (given the covariates).
- In the present application: This probability is time-dependent.
- Thus: We obtain a ,dynamic' propensity score in the variable \mathbf{X} at each fixed point in time \mathbf{t} :

$$p(\mathbf{t}, \mathbf{X}) = 1 - F(\mathbf{t}, \mathbf{X})$$

where

$$F(\mathbf{t}, \mathbf{X}_i) = P(S_i \geq \mathbf{t} | \mathbf{X}_i)$$

is the 'survival function' of the treatment initiation model.

Note: Fairly easy to estimate by most survival analysis software packages.

How to include the dynamic ps in the outcome regression

- Add the ps as a control function in the outcome Cox regression.
- For the proportional hazard rates $\mathbf{h}(t|\mathbf{S}, \mathbf{X})$ of the survival time \mathbf{T}_i we have:

$$\mathbf{h}(t|\mathbf{S}, \mathbf{X}) = e^{\alpha 1[t \geq S] + \delta \hat{p}(t, \mathbf{X})} \cdot \mathbf{h}_0(t) \quad (\text{eqn. 2})$$

- Preceding considerations suggest:
 - The Cox regression for this model will produce a reliable **estimate of the causal treatment effect α** .
 - The coefficient δ of the ps will yield an **estimate of the selection bias** that would confound the outcomes.
Additional benefit, direct estimation of selection bias.

Two-step procedure

To summarize, the whole procedure consists of two steps:

1. Estimate a Cox regression for the treatment initiation using the full set of covariates \mathbf{X} . Obtain the propensity score $\hat{p}(t, \mathbf{X})$ as $1 - F^S(t, \mathbf{X})$
2. Run a Cox regression for the outcome model using (at least) the treatment indicator $\mathbf{1}[t \geq S]$ and the estimated propensity score as **time-dependent regressors** $\hat{p}(t, \mathbf{X})$

Simulation (I)

- Covariates $X = (X_1, X_2, X_3, X_4)$, where X_1 is binary with $P(X_1 = 1) = 0,5$, X_2 is binary with $P(X_2 = 1) = 0,75$, X_3 is uniformly distributed on $[-1, 1]$ and X_4 has a standard normal distribution
- Weibull-distributed (potential) survival times T^0 with shape parameter $\kappa = 3$ and scale parameter $(1/\lambda_1)^{1/\kappa}$

$$\lambda_1 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$$

$$\beta_0 = 12, \beta_1 = 0.5, \beta_2 = 0.5, \beta_3 = 0.5, \beta_4 = 0.5$$

Simulation (II)

- Weibull-distributed treatment initiation times S with shape parameter $\kappa = 2$ and scale parameter $(1/\lambda_2)^{1/\kappa}$

$$\lambda_2 = \gamma_0 + \gamma_1 \mathbf{X}_1 + \gamma_2 \mathbf{X}_2 + \gamma_3 \mathbf{X}_3 + \gamma_4 \mathbf{X}_4$$

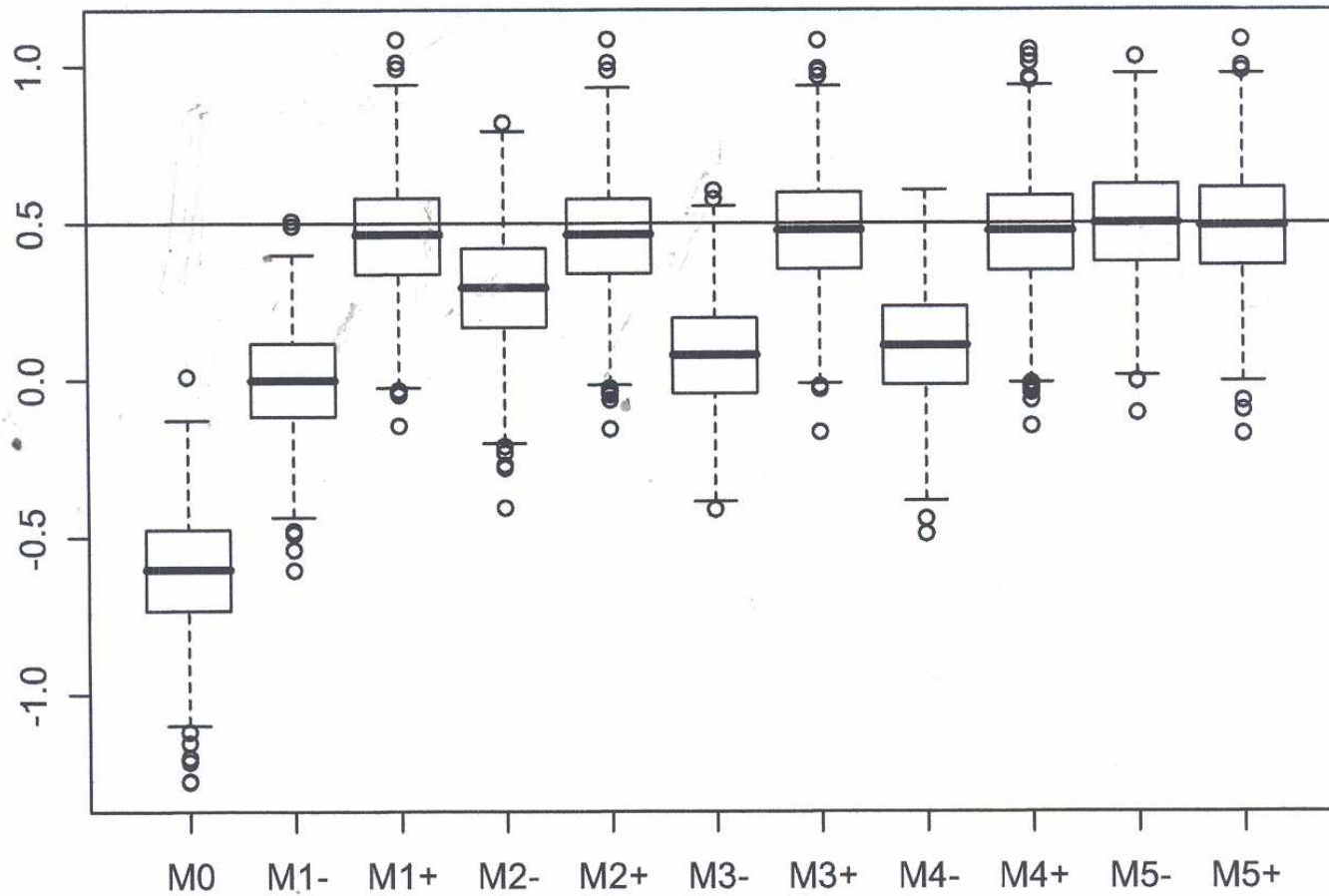
$$\gamma_0 = -8.5, \gamma_1 = -0.3, \gamma_2 = -0.3, \gamma_3 = -0.4, \gamma_4 = -0.8$$

- $\mathbf{T} = \sqrt[\kappa]{e^{-\alpha} (\mathbf{T}^0)^\kappa + (1 - e^{-\alpha}) \mathbf{S}^\kappa}$; $\alpha = 0.5$; $\kappa = 3$
- The duration times T and S comply with the proportional hazards model.
- 1000 simulations with $N = 200$ patients each.

Specification for Outcome Models

- M0: Time-independent covariates (correct form)
- M1: No other covariates
- M2: Only \mathbf{X}_3
- M3: \mathbf{X}_1 and \mathbf{X}_3
- M4: $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3^2, \mathbf{X}_4^2$
- M5: All covariates (in correct functional form)

Results of simulations



Conclusion

- False specification of outcome model leads to strongly biased estimate of α .
- Including dynamic propensity score resolves this problem.
- Residual bias may be due to linearity assumptions or double censoring.