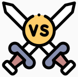
Estimation of ATE in Causal Survival: Comparison, Applications and Practical Recommendations

Julie Josse (Inria) & Bernard Sebastien (Sanofi R&D)

Charlotte VOINOT

May 30, 2024

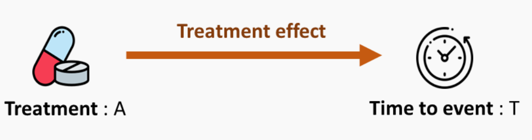
Causal Survival analysis: example of questions

Survival analysis  Causal inference

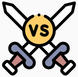
⇒ **Effect of a policy/intervention/treatment A on an time to event outcome T**

What is the impact of an oncology medicine on long term mortality ?

What is the impact of a medecine all along the time (it can be beneficial in the short term but harmful in the long term) ?



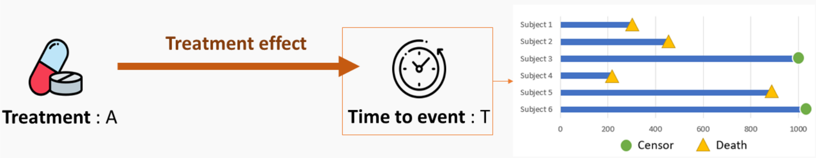
Causal Survival analysis: example of questions

Survival analysis  Causal inference

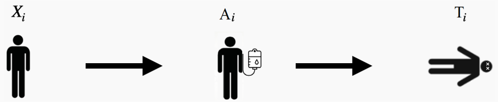
⇒ **Effect of a policy/intervention/treatment A on an time to event outcome T**

What is the impact of an oncology medicine on long term mortality ?

What is the impact of a medecine all along the time (it can be beneficial in the short term but harmful in the long term) ?



Potential outcomes



$$\Rightarrow n \textbf{ i.i.d. } (\underbrace{X_i}_{\text{covariates}}, \overbrace{A_i}^{\text{treatment}}, \underbrace{T_i}_{\text{outcome}}) \in \mathbb{R}^d \times \{0,1\} \times \mathbb{R} \times \mathbb{R}$$

Let's say that in our example X_1 = sex and X_2 = age.

Covariates		Treatment	Censoring	Status	Outcomes		
X_1	X_2	**A**	**C**	Δ	**T(0)**	**T(1)**	$\tilde{T}$
1	24	1	?	1	?	200	200
2	52	0	?	1	100	?	100
1	33	1	200	0	?	?	200

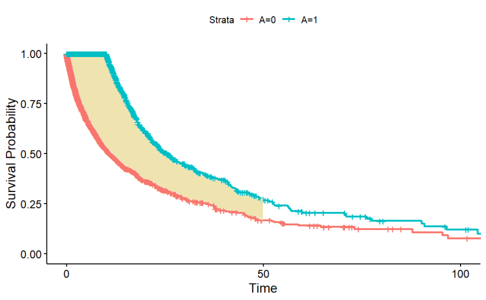
In grey, the observed data : $(X_i, A_i, \Delta_i, \widetilde{T}_i)$ with $\widetilde{T}_i = min(T_i, C_i)$

$\Rightarrow$ T is not directly observed

Difference in RMST : Average treatment effect in survival analysis

$$\theta_{RMST}(\tau) = E[min(T(1),\tau) - min(T(0),\tau)]$$
$$= \int_0^{\tau} (S_1(t) - S_0(t))\, dt$$

RMST can be defined as a measure of average survival from time 0 to time τ a **fixed time horizon**



$\theta_{RMST}(\tau = 50) = 10$ means that on average the treatment increases the survival time by 10 days at 50 days.

Figure 1: Plot of stratified kaplan meier survival function and the representation of $\theta_{RMST}(\tau = 50)$ (in yellow)

Identifiability assumptions

S.T.U.V.A. $T = AT(1) + (1 - A)T(0)$

RCT & Independent censoring

- **Random treatment assignment**
  $A \perp\!\!\!\perp (T(0), T(1), C, X)$
- **Independent censoring**
  $C \perp\!\!\!\perp T(0), T(1), X, A$

Identifiability assumptions

S.T.U.V.A. $T = AT(1) + (1 - A)T(0)$

RCT & Independent censoring

- **Random treatment assignment**
  $A \perp\!\!\!\perp (T(0), T(1), C, X)$
- **Independent censoring**
  $C \perp\!\!\!\perp T(0), T(1), X, A$

RCT & Dependent censoring

- **Random treatment assignment**
  $A \perp\!\!\!\perp (T(0), T(1), C, X)$
- **Conditionally independent censoring**
  $C \perp\!\!\!\perp T(0), T(1) | X, A$
- **Positivity for censoring**
  $0 < P(C > t \mid X = x, A = a) < 1$

Identifiability assumptions

S.T.U.V.A. $T = AT(1) + (1-A)T(0)$

RCT & Independent censoring

- **Random treatment assignment** $A \perp\!\!\!\perp (T(0), T(1), C, X)$
- **Independent censoring** $C \perp\!\!\!\perp T(0), T(1), X, A$

RCT & Dependent censoring

- **Random treatment assignment** $A \perp\!\!\!\perp (T(0), T(1), C, X)$
- **Conditionally independent censoring** $C \perp\!\!\!\perp T(0), T(1) | X, A$
- **Positivity for censoring** $0 < P(C > t \mid X = x, A = a) < 1$

Obs & Independent censoring

- **Unconfoundedness** $A \perp\!\!\!\perp (T(0), T(1)) | X$
- **Positivity for treatment** $1 > P(A = a \mid X = x) > 0$
- **Independent censoring** $C \perp\!\!\!\perp T(0), T(1), X, A$

Identifiability assumptions

S.T.U.V.A. $T = AT(1) + (1-A)T(0)$

RCT & Independent censoring

- **Random treatment assignment** $A \perp\!\!\!\perp (T(0), T(1), C, X)$
- **Independent censoring** $C \perp\!\!\!\perp T(0), T(1), X, A$

Obs & Independent censoring

- **Unconfoundedness** $A \perp\!\!\!\perp (T(0), T(1))|X$
- **Positivity for treatment** $1 > P(A = a \mid X = x) > 0$
- **Independent censoring** $C \perp\!\!\!\perp T(0), T(1), X, A$

RCT & Dependent censoring

- **Random treatment assignment** $A \perp\!\!\!\perp (T(0), T(1), C, X)$
- **Conditionally independent censoring** $C \perp\!\!\!\perp T(0), T(1)|X, A$
- **Positivity for censoring** $0 < P(C > t \mid X = x, A = a) < 1$

Obs & Dependent censoring

- **Unconfoundedness** $A \perp\!\!\!\perp (T(0), T(1))|X$
- **Positivity for treatment** $1 > P(A = a \mid X = x) > 0$
- **Conditionally independent censoring** $C \perp\!\!\!\perp T(0), T(1)|X, A$
- **Positivity for censoring** $0 < P(C > t \mid X = x, A = a) < 1$

RCT & Independent censoring

$$
\begin{aligned}
\theta_{RMST} &= \mathbb{E}[T(1) \wedge \tau - T(0) \wedge \tau] \\
&= \int_0^\tau \mathbb{E}[I\{T(1) > t\} - I\{T(0) > t\}]dt && \text{(By definition)} \\
&= \int_0^\tau \mathbb{E}[I\{T(1) > t\}] - \mathbb{E}[I\{T(0) > t\}]dt && \text{(By linearity)} \\
&= \int_0^\tau \mathbb{E}[I\{T(1) > t | A = 1\}] - \mathbb{E}[I\{T(0) > t | A = 0\}]dt && \text{(Random treatment assignment)} \\
&= \int_0^\tau \mathbb{E}[I\{T > t | A = 1\}] - \mathbb{E}[I\{T > t | A = 0\}]dt && \text{(By consistency)} \\
&= \int_0^\tau \mathbb{P}(T > t | A = 1) - \mathbb{P}(T > t | A = 0)dt \\
&= \int_0^\tau S(t | A = 1) - S(t | A = 0)dt
\end{aligned}
$$

Non-adjusted Kaplan-Meier estimator [1]

It corresponds to a simple Kaplan meier estimator:

$$\hat{S}_{KM}(t \mid a) = \prod_{j=1, t_j <= t} \left(1 - \frac{\sum_i I\left\{ \tilde{T}_i = t_j, \Delta_i = 1, A_i = a \right\}}{\sum_k I\left\{ \tilde{T}_k \geq t_j, A_i = a \right\}} \right)$$

$$\theta_{RMST}(\tau) = \int_0^\tau \left(S_1(t) - S_0(t) \right) dt$$

RCT & Dependent censoring

Non-adjusted Kaplan-Meier is biased in this context:

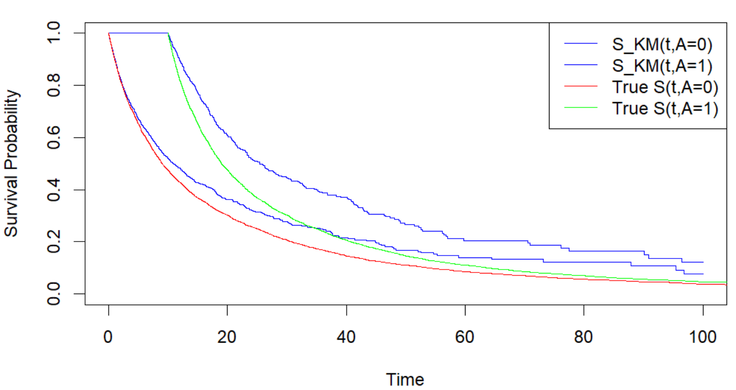


Figure 2: Plot of stratified kaplan meier survival function (A=1 and A=0) and the true survival function

⇒ The probability of survival is overestimated in using Non adjusted KM [3].

RCT & Dependent censoring

$E[T \wedge \tau | X]$ is no longer straightforwardly identifiable
$\Rightarrow$ But $E(T \wedge \tau | X, A)$ can be writen as $E(T^* | X, A)$:

$$T^*(\tau) = \frac{\widetilde{T} \wedge \tau * \Delta^\tau}{S_c(\widetilde{T} \wedge \tau | X, A)}$$

with $\Delta^\tau = I\{T \wedge \tau < C | A = 1\}$ and $\hat{S}_c(t | X_i, A_i)$ is the probability of remain uncensored given the covariates.

IPCW adjusted Kaplan-Meier estimator [2]

$$\hat{S}_{IPCW}(t \mid A = a) = \prod_{j=1, t_j \leq t} \left(1 - \frac{\sum_i \hat{w}_i(t_j, X_i) \cdot I\left\{\widetilde{T}_i = t_j, C_i \geq t_j, A_i = a\right\}}{\sum_k \hat{w}_k(t_j, X_k) \cdot I\left\{\widetilde{T_k} \geq t_j, C_k \geq t_j, A_k = a\right\}} \right)$$

with $\hat{w}_i(t, X_i) = \frac{\Delta_i^\tau}{\hat{S}_c(t | X_i, A_i)}$: every uncensored observation is weighted by the inverse of the probability of remain uncensored given the covariates.

$$\theta_{RMST} = \int_0^\tau \hat{S}_{IPCW}(t, A = 1) - \hat{S}_{IPCW}(t, A = 0) dt$$

Obs & Independent censoring

Risk of confounding bias

$\Rightarrow$ Need for balancing differences between the two groups.

IPTW adjusted Kaplan-Meier estimator [4]

$$\hat{S}_{IPTW}(t \mid A = a) = \prod_{j=1, t_j \leq t} \left(1 - \frac{\sum_i \hat{w}_i(t_j, X_i) \cdot I\left\{ \widetilde{T}_i = t_j, C_i \geq t_j, A_i = a \right\}}{\sum_k \hat{w}_k(t_j, X_k) \cdot I\left\{ \widetilde{T_k} \geq t_j, C_k \geq t_j, A_k = a \right\}} \right)$$

with $\hat{w}_i(t, X_i) = \frac{A_i}{\hat{e}(X_i)} + \frac{1-A_i}{1-\hat{e}(X_i)}$: every observation is weighted by the inverse of the propensity score (probability of being treated) given the covariates.

$$\theta_{RMST} = \int_0^{\tau} \hat{S}_{IPTW}(t, A = 1) - \hat{S}_{IPTW}(t, A = 0) dt$$

Obs & Dependent censoring

Risk of censoring bias & confounding bias

IPTW-IPCW adjusted Kaplan-Meier estimator

$$\hat{S}_{IPTW-IPCW}(t \mid A=a) = \prod_{j=1, t_j<=t} \left(1 - \frac{\sum_i \hat{w}_i(t, X_i) * I\{T_i = t_j, C_i \geq t_j, A_i = a\}}{\sum_i \hat{w}_i(t, X_i) * I\{T_i \geq t_j, C_i \geq t_j, A_i = a\}}\right)$$

with $\hat{w}_i(t, X_i) = \frac{\Delta_i^\tau}{\hat{S}_C\left(\widetilde{T} \wedge \tau | A_i, X_i\right)} * \left(\frac{A_i}{\hat{e}(X_i)} + \frac{1-A_i}{1-\hat{e}(X_i)}\right)$: every uncensored observation is weighted by the inverse of remain uncensored and by the inverse propensity score given the covariates.

$$\theta_{RMST} = \int_0^\tau \hat{S}_{IPTW-IPCW}(t, A=1) - \hat{S}_{IPTW-IPCW}(t, A=0) dt$$

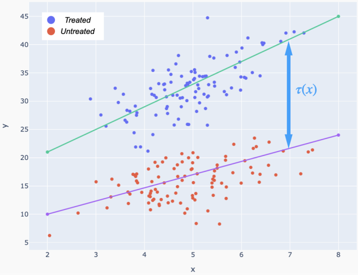
Obs & Dependent censoring

$$\begin{aligned}\theta &= \mathbb{E}\left[T(1)\wedge\tau - T(0)\wedge\tau\right]\\ &= \mathbb{E}\left[\mathbb{E}\left[T(1)\wedge\tau - T(0)\wedge\tau \mid X\right]\right]\\ &= \mathbb{E}\left[\mathbb{E}\left[T(1)\wedge\tau \mid X, A=1\right] - \mathbb{E}\left[T(0)\wedge\tau \mid X, A=0\right]\right]\end{aligned}$$

(Uncounfoundedness)

$$= \mathbb{E}\left[\mathbb{E}\left[T\wedge\tau \mid X, A=1\right] - \mathbb{E}\left[T\wedge\tau \mid X, A=0\right]\right]$$

(Consistency)



G-formula estimator

$$\widehat{\theta}_{\text{g-formula}}(\tau) = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{F}\left(X_i, 1\right) - \hat{F}\left(X_i, 0\right)\right)$$

with $\hat{F}(x,a) \triangleq \mathbb{E}[T\wedge\tau \mid X = x, A = a]$

$\hat{F}(x,a)$ can be obtained by fitting a parametric model (e.g., Weibull distribution), a semi-parametric model (e.g., Cox model), or a non-parametric model (e.g., survival random forest).

Obs & Dependent censoring

Augmented estimator

$$\hat{\theta}_{\text{AIPTW-AIPCW}} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{A_i}{\hat{e}\left(X_i\right)} - \frac{1-A_i}{1-\hat{e}\left(X_i\right)}\right)\hat{T}^*_{\text{DR}}$$
$$+ \hat{F}\left(X_i, A=1\right)\left(1-\frac{A_i}{\hat{e}\left(X_i\right)}\right) - \hat{F}\left(X_i, A=0\right)\left(1-\frac{1-A_i}{1-\hat{e}\left(X_i\right)}\right)$$

with $\hat{T}^*_{\text{DR}} = \frac{\widetilde{T}_i\wedge\tau\cdot\Delta^\tau}{\hat{S}_C\left(\widetilde{T}_i\wedge\tau|X_i\right)} + \frac{Q_{\hat{S}}(\tilde{T}_i\wedge\tau|X,A)\cdot(1-\Delta^\tau)}{\hat{S}_C(\tilde{T}_i\wedge\tau|X_i)} - \int_0^{\widetilde{T}_i\wedge\tau}\frac{Q_{\hat{S}}(c|X_i,A_i)}{\hat{S}^2_C(c|X_i,A_i)}d\hat{S}_C(c \mid X_i, A_i)$ and

$Q_S(t|x,a) = \mathbb{E}[T\wedge\tau|X=x, A=a, T\wedge\tau > t]$

$\Rightarrow$ 3 nuisance parameters to compute :

- Censoring model : $C \sim A + X$
- Propensity score model : $A \sim X$
- Conditional survival : $T \sim A + X$

Obs & Dependent censoring

Augmented estimator

$$\hat{\theta}_{\text{AIPTW-AIPCW}} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{A_i}{\hat{e}\left(X_i\right)} - \frac{1-A_i}{1-\hat{e}\left(X_i\right)}\right)\hat{T}^*_{\text{DR}}$$
$$+ \hat{F}\left(X_i, A=1\right)\left(1-\frac{A_i}{\hat{e}\left(X_i\right)}\right) - \hat{F}\left(X_i, A=0\right)\left(1-\frac{1-A_i}{1-\hat{e}\left(X_i\right)}\right)$$

with $\hat{T}^*_{\text{DR}} = \frac{\widetilde{T}_i \wedge \tau \cdot \Delta^\tau}{\hat{S}_C\left(\widetilde{T}_i \wedge \tau | X_i\right)} + \frac{Q_{\hat{S}}(\tilde{T}_i \wedge \tau | X, A)\cdot(1-\Delta^\tau)}{\hat{S}_C(\tilde{T}_i \wedge \tau | X_i)} - \int_0^{\widetilde{T}_i \wedge \tau} \frac{Q_{\hat{S}}(c|X_i, A_i)}{\hat{S}^2_C(c|X_i, A_i)} d\hat{S}_C(c \mid X_i, A_i)$ and

$Q_S(t|x, a) = \mathbb{E}[T \wedge \tau | X = x, A = a, T \wedge \tau > t]$

T^*_{DR} corresponds to the augmented censoring transformation (AIPCW):

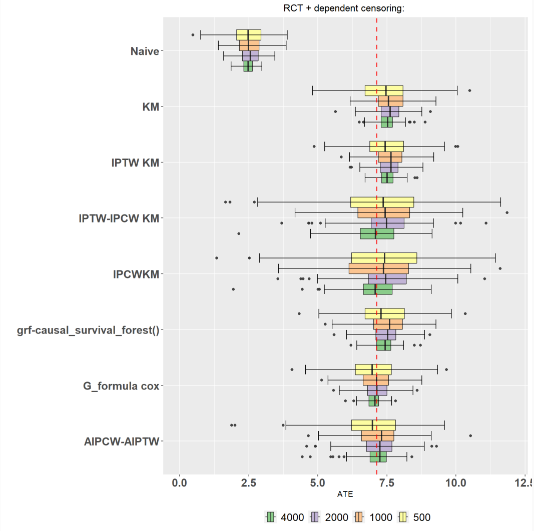
- The first term : IPCW (it weights the uncensored observation by the inverse probability of remain uncensored)
- The second term : it weights in the same way the censored observation by using an estimation of survival
- The third one is an augmentation term

Obs & Dependent censoring

Estimator	mis. S	mis. S_C	mis. e	mis. S and S_C	mis. S and e	mis. S_C and e
G-formula		✓	✓			✓
IPTW-IPCW	✓					
AIPTW-AIPCW	✓	✓	✓			✓

Table 1: Consistency under model mis-specification. When all the nuisances models are mis-specified none of the estimators is consistent. ✓indicates consistency of the estimator.

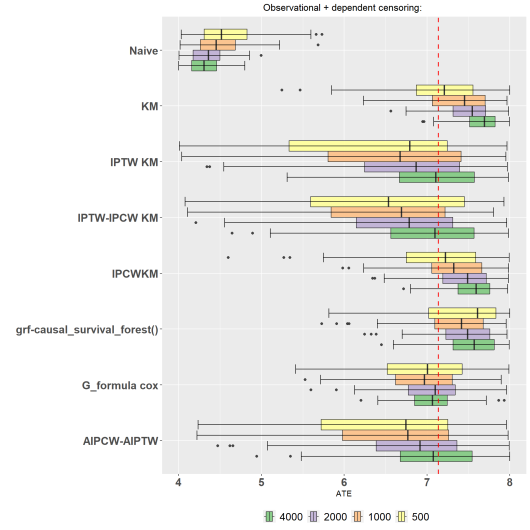
Good specification of nuisance model



- As expected 'Naive' is still completely biased
- KM and IPTW estimators are biased because of dependent censoring
- All estimator with censoring transformation converge
- Surprisingly, causal survival forest from grf is biased (maybe with more sample size)
- G-formula seems to be the best estimators

Figure 3: Boxplot of RMST estimation in the context of RCT and dependent censoring for 150 simulations ($\tau = 25$)

Good specification of nuisance model:



- 'Naive' still biased
- All estimators without both censoring transformation and inverse propensity weighting are biased : 'KM', 'IPTW KM', 'IPCW KM'.
- grf biased
- G-formula is the best estimator

Figure 4: Boxplot of RMST estimation in the context of Observational study and dependent censoring for 150 simulations ($\tau = 25$)

Conclusions

- Few packages: implementation of estimators from scratch.
- RMST is one extension of causal inference in survival analysis.
- Robust estimators such as AIPCW-AIPTW converge for all setup but sometimes G-formula outperforms this DR estimator.
- But AIPCW-AIPTW needs more data due to its complexity.

Perspectives

- Publication in Computo.
- Non parametric setting.
- Importance in variable selections : variables which influences censoring and time to event not necessary the same.

References

[1] E. L. Kaplan and Paul Meier. **“Nonparametric Estimation from Incomplete Observations”.** In: *Journal of the American Statistical Association* (1958).

[2] James M. Robins and Dianne M. Finkelstein. **“Correcting for Noncompliance and Dependent Censoring in an AIDS Clinical Trial with Inverse Probability of Censoring Weighted (IPCW) Log-Rank Tests”.** In: *Biometrics* (2000).

[3] SJW Willems et al. **“Correcting for dependent censoring in routine outcome monitoring data by applying the inverse probability censoring weighted estimator”.** In: *Statistical Methods in Medical Research* (2018).

[4] Jun Xie and Chaofeng Liu. **"Adjusted Kaplan–Meier estimator and log-rank test with inverse probability of treatment weighting for survival data".** In: *Statistics in Medicine* (2005).

Appendix

RMST and survival probability

Why $\theta_{RMST}(\tau) = E[min(T(1), \tau) - min(T(0), \tau)]$ and $\theta_{RMST}(\tau) = \int_0^\tau (S_1(t) - S_0(t))\, dt$ are equal ?
$E(T \wedge \tau) = E(\int_0^{T \wedge \tau} 1dt) = E(\int_0^\tau I\{T > t\}dt) = \int_0^\tau E(I\{T > t\})dt = \int_0^\tau S(t)dt$

Non adjusted Kaplan meier

$$
\begin{aligned}
\theta_{RMST} &= \mathbb{E}[T(1) \wedge \tau - T(0) \wedge \tau] \\
&= \int_0^{\tau} \mathbb{E}[I\{T(1) > t\} - I\{T(0) > t\}]dt && \text{(By definition)} \\
&= \int_0^{\tau} \mathbb{E}[I\{T(1) > t\}] - \mathbb{E}[I\{T(0) > t\}]dt && \text{(By linearity)} \\
&= \int_0^{\tau} \mathbb{E}[I\{T(1) > t|A = 1\}] - \mathbb{E}[I\{T(0) > t|A = 0\}]dt && \text{(Random treatment assignm} \\
&= \int_0^{\tau} \mathbb{E}[I\{T > t|A = 1\}] - \mathbb{E}[I\{T > t|A = 0\}]dt && \text{(By consistency)} \\
&= \int_0^{\tau} \mathbb{P}(T > t|A = 1) - \mathbb{P}(T > t|A = 0)dt \\
&= \int_0^{\tau} S(t|A = 1) - S(t|A = 0)dt
\end{aligned}
$$

IPCW Kaplan meier

$$
\begin{aligned}
\theta_{RMST} &= \mathbb{E}[T(1) \wedge \tau - T(0) \wedge \tau] \\
&= \int_0^\tau \mathbb{E}[I\{T(1) > t\}] - \mathbb{E}[I\{T(0) > t\}]dt \\
&= \int_0^\tau \mathbb{E}\left[\mathbb{E}[I\{T > t\}|A = 1, X]\right] - \mathbb{E}\left[\mathbb{E}[I\{T > t\}|A = 0, X]\right] \\
&= \int_0^\tau \mathbb{E}\left[\frac{\mathbb{E}[I\{T > t|A = 1, X\}] * \mathbb{E}[I\{T \wedge \tau < C|A = 1, X\}]}{S_c(T(1) \wedge \tau|X, A = 1)}\right] - \\
&\quad \mathbb{E}\left[\frac{\mathbb{E}[I\{T > t|A = 0, X\}] * \mathbb{E}[I\{T \wedge \tau < C|A = 0, X\}]}{S_c(T(0) \wedge \tau|X, A = 0)}\right] dt \\
&= \int_0^\tau \mathbb{E}\left[\frac{\mathbb{E}[I\{T > t|A = 1, X\} * I\{T \wedge \tau < C|A = 1, X\}]}{S_c(T \wedge \tau|X, A = 1)}\right] - \\
&\quad \mathbb{E}\left[\frac{\mathbb{E}[I\{T > t|A = 0, X\} * I\{T \wedge \tau < C|A = 0, X\}]}{S_c(T \wedge \tau|X, A = 0)}\right] dt \\
&= \int_0^\tau \mathbb{E}\left[\frac{I\{T > t|A = 1\} * \Delta^\tau}{S_c(T \wedge \tau|X, A = 1)}\right] - \mathbb{E}\left[\frac{I\{T > t|A = 0\} * \Delta^\tau}{S_c(T \wedge \tau|X, A = 0)}\right] dt
\end{aligned}
$$

IPTW Kaplan meier

$$\theta = \mathbb{E}[T(1) \wedge \tau - T(0) \wedge \tau]$$

$$= \int_0^\tau \mathbb{E}[I\{T(1) > t\}] - \mathbb{E}[I\{T(0) > t\}]dt$$

(By linearity)

$$= \int_0^\tau \mathbb{E}\left[\mathbb{E}[I\{T(1) > t\}|X]\right] - \mathbb{E}\left[\mathbb{E}[I\{T(0) > t\}|X]\right]$$

(Law of total probability and Consistency)

$$= \int_0^\tau \mathbb{E}\left[\frac{\mathbb{E}[I\{T(1) > t|X\}] * \color{blue}{\mathbb{E}[A|X\}]}}{\color{blue}{e(X)}}\right] - \mathbb{E}\left[\frac{\mathbb{E}[I\{T(0) > t|X\}] * \color{red}{\mathbb{E}[1 - A|X\}]}}{\color{red}{1 - e(X)}}\right] dt$$

(In color, the terms are equal)

$$= \int_0^\tau \mathbb{E}\left[\frac{\mathbb{E}[I\{T(1) > t\} * A|X\}}{e(X)}\right] - \mathbb{E}\left[\frac{\mathbb{E}[I\{T(0) > t\} * (1 - A)|X]]}{1 - e(X)}\right] dt$$

(By unconfoundedness)

$$= \int_0^\tau \frac{\mathbb{E}[I\{T(1) > t\} * A]}{e(X)} - \frac{\mathbb{E}[I\{T(0) > t\} * (1 - A)]}{1 - e(X)} dt$$

(Law of total probability)

$$= \int_0^\tau \frac{\mathbb{E}[I\{T(1) > t|A = 1\} * A]}{e(X)} - \frac{\mathbb{E}[I\{T(0) > t|A = 0\} * (1 - A)]}{1 - e(X)} dt$$

$$= \int_0^\tau \frac{\mathbb{E}[I\{T > t|A = 1\} * A]}{e(X)} - \frac{\mathbb{E}[I\{T > t|A = 0\} * (1 - A)]}{1 - e(X)} dt$$

(By consistency)

$$= \int_0^\tau p(T \geq t|A = 1) * \left(\frac{A}{e(X)}\right) - p(T \geq t|A = 0) * \left(\frac{1 - A}{1 - e(X)}\right) dt$$

RCTs simulations

For the simulation, 2000 samples $(X_i, A_i, C, T_i(0), T_i(1))$ are generated in the following way:

- $X \sim \mathcal{N}\left(\mu = [1, 1, -1, 1]^\top, \Sigma = I_4\right)$
- $e(X) = 0.5$ (constant) for the propensity score (A)
- $\lambda(0)(X) = 0.01 \cdot \exp\{0.5X_1 + 0.5X_2 - 0.5X_3 + 0.5X_4\}$ hazard for the event time $T(0)$
- The hazard for the censoring time C:
  - For scenario 1:$\lambda_c = 0.09$.
  - For scenario 2:$\lambda_c(X) = 0.03 \cdot \exp\{0.1X_1 + 0.1X_2 - 0.2X_3 - 0.2X_4\}$.
- $T(1) = T(0) + 10$
- the event time is $T = AT(1) + (1 - A)T(0)$
- The observed time is $\widetilde{T} = \min(T, C)$
- The status is $\Delta = 1(T \leq C)$

Obs simulations

For the simulation, 2000 samples $(X_i, A_i, C, T_i(0), T_i(1))$ are generated in the following way:

- $X \sim \mathcal{N}\left(\mu = [1, 1, -1, 1]^\top, \Sigma = I_4\right)$

- $\text{logit}\{e(X)\} = -1X_1 - 0.5X_2 + 2X_3 + 1X_4$ for the propensity score (A)

- $\lambda(0)(X) = 0.1 \cdot \exp\{0.5X_1 - 0.1X_2 + 0.3X_3 + 0.2X_4\}$ hazard for the event time $T(0)$

- The hazard for the censoring time C:
  - For scenario 1:$\lambda_c = 0.09$.
  - For scenario 2:$\lambda_c(X) = 0.03 \cdot \exp\{0.1X_1 + 0.1X_2 - 0.2X_3 - 0.2X_4\}$.

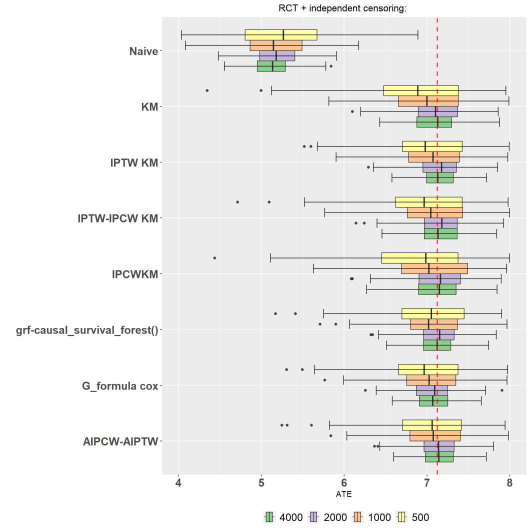
- $T(1) = T(0) + 10$

- the event time is $T = AT(1) + (1-A)T(0)$

- The observed time is $\widetilde{T} = \min(T, C)$

- The status is $\Delta = 1(T \leq C)$

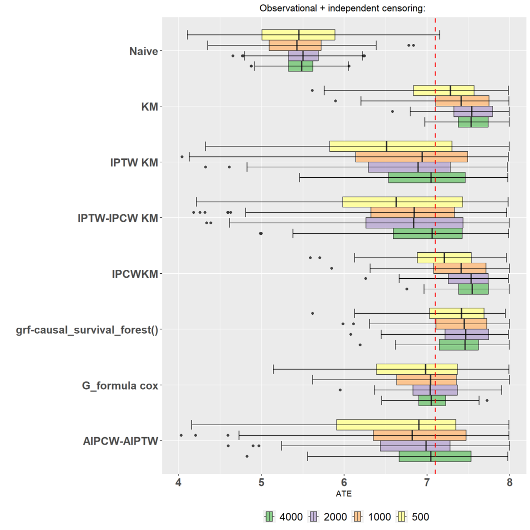
Good specification of nuisance model: RCT + independent censoring



- All estimators converge except 'naive'
- Convergence starting from 2,000 observations
- Small bias even for small sample size
- The best estimators (smaller variance + smaller bias at small sample size) is AIPCW-AIPTW estimator

Figure 5: Boxplot of RMST estimation in the context of RCT and independent

Good specification of nuisance model



- 'Naive' is still biased
- Estimators without inverse propensity weighting are biased : KM, IPCW KM.
- Causal survival forest is biased also (too small sample size)
- G-formula seems to be the best estimator

Figure 6: Boxplot of RMST estimation in the context of Observational study and independent censoring for 150 simulations ($\tau = 25$)