Sequential Regressions for Efficient Continuous-Time Causal Inference

Johan Sebastian Ohlendorff

University of Copenhagen

- PhD student in Biostatistics at the Section of Biostatistics, University of Copenhagen.

- PhD student in Biostatistics at the Section of Biostatistics, University of Copenhagen.
- Supervisors: Thomas Alexander Gerds and Anders Munch.

- PhD student in Biostatistics at the Section of Biostatistics, University of Copenhagen.
- Supervisors: Thomas Alexander Gerds and Anders Munch.
- Work on continuous-time longitudinal causal inference using
  - Targeted learning (e.g., TMLE (van der Laan & Rubin (2006)) or double/debiased machine learning (Chernozhukov et al. (2018))).

- PhD student in Biostatistics at the Section of Biostatistics, University of Copenhagen.
- Supervisors: Thomas Alexander Gerds and Anders Munch.
- Work on continuous-time longitudinal causal inference using
  - Targeted learning (e.g., TMLE (van der Laan & Rubin (2006)) or double/debiased machine learning (Chernozhukov et al. (2018))).
- Working on a continuous-time scale is motivated by:
  - Real-world data often recorded in continuous time (e.g., electronic health records).
  - Avoiding bias due to discretization (Adams et al. (2020); Ferreira Guerra et al. (2020); Kant & Krijthe (2025); Ryalen et al. (2019); Sofrygin et al. (2019); Sun & Crawford (2023)).

- PhD student in Biostatistics at the Section of Biostatistics, University of Copenhagen.
- Supervisors: Thomas Alexander Gerds and Anders Munch.
- Work on continuous-time longitudinal causal inference using
  - Targeted learning (e.g., TMLE (van der Laan & Rubin (2006)) or double/debiased machine learning (Chernozhukov et al. (2018))).
- Working on a continuous-time scale is motivated by:
  - Real-world data often recorded in continuous time (e.g., electronic health records).
  - Avoiding bias due to discretization (Adams et al. (2020); Ferreira Guerra et al. (2020); Kant & Krijthe (2025); Ryalen et al. (2019); Sofrygin et al. (2019); Sun & Crawford (2023)).
- We follow the setting in Rytgaard et al. (2022) and are interested in the mean interventional absolute risk under a specified treatment regime in continuous time.

- PhD student in Biostatistics at the Section of Biostatistics, University of Copenhagen.
- Supervisors: Thomas Alexander Gerds and Anders Munch.
- Work on continuous-time longitudinal causal inference using
  - Targeted learning (e.g., TMLE (van der Laan & Rubin (2006)) or double/debiased machine learning (Chernozhukov et al. (2018))).
- Working on a continuous-time scale is motivated by:
  - Real-world data often recorded in continuous time (e.g., electronic health records).
  - Avoiding bias due to discretization (Adams et al. (2020); Ferreira Guerra et al. (2020); Kant & Krijthe (2025); Ryalen et al. (2019); Sofrygin et al. (2019); Sun & Crawford (2023)).
- We follow the setting in Rytgaard et al. (2022) and are interested in the mean interventional absolute risk under a specified treatment regime in continuous time.
- **Problem**: Rytgaard et al. (2022) do not provide a feasibly implementable procedure for estimation.

- **Study period**: $[0, \tau_{\text{end}}]$.

- **Study period**: $[0, \tau_{\text{end}}]$.
- **Càdlàg, jump processes**
  - Treatment: $A(t) \in \{0, 1\}$
  - Covariates: $L(t) \in \mathbb{R}^d$

- **Study period**: $[0, \tau_{\text{end}}]$.
- **Càdlàg, jump processes**
  - Treatment: $A(t) \in \{0, 1\}$
  - Covariates: $L(t) \in \mathbb{R}^d$
- **Counting processes**
  - $N^x; x \in \{a, \ell, y\}$ (treatment, covariate, death)
  - Observed until censoring time C
  - Terminal event time: T_e

- **Study period**: $[0, \tau_{\text{end}}]$.
- **Càdlàg, jump processes**
  - Treatment: $A(t) \in \{0, 1\}$
  - Covariates: $L(t) \in \mathbb{R}^d$
- **Counting processes**
  - $N^x; x \in \{a, \ell, y\}$ (treatment, covariate, death)
  - Observed until censoring time C
  - Terminal event time: T_e
- **Assumptions**
  - Jumps in $A(t)$ or $L(t)$ only when corresponding N^a or (N^ℓ or N^a) jump.
  - No simultaneous jumps + orthogonal martingales

- **Study period**: $[0, \tau_{\text{end}}]$.
- **Càdlàg, jump processes**
  - Treatment: $A(t) \in \{0, 1\}$
  - Covariates: $L(t) \in \mathbb{R}^d$
- **Counting processes**
  - $N^x; x \in \{a, \ell, y\}$ (treatment, covariate, death)
  - Observed until censoring time C
  - Terminal event time: T_e
- **Assumptions**
  - Jumps in $A(t)$ or $L(t)$ only when corresponding N^a or (N^ℓ or N^a) jump.
  - No simultaneous jumps + orthogonal martingales
- **Treatment decisions**
  - Doctor chooses treatment at jump times of N^a
  - Intervention specifies decision, not timing

- **Study period**: $[0, \tau_{\text{end}}]$.
- **Càdlàg, jump processes**
  - Treatment: $A(t) \in \{0, 1\}$
  - Covariates: $L(t) \in \mathbb{R}^d$
- **Counting processes**
  - $N^x; x \in \{a, \ell, y\}$ (treatment, covariate, death)
  - Observed until censoring time C
  - Terminal event time: T_e
- **Assumptions**
  - Jumps in $A(t)$ or $L(t)$ only when corresponding N^a or (N^ℓ or N^a) jump.
  - No simultaneous jumps + orthogonal martingales
- **Treatment decisions**
  - Doctor chooses treatment at jump times of N^a
  - Intervention specifies decision, not timing
- **Bounded events**
  - Each individual has at most K events in $[0, \tau_{\text{end}}]$

$$\mathcal{F}_t = \sigma\Big(\big(A(s), L(s), N^a(s), N^\ell(s), N^y(s)\big) : s \leq t\Big) \vee \sigma((A(0), L(0)))$$

- $\mathcal{F}_t$: natural filtration for the processes without censoring

$$\mathcal{F}_t = \sigma\big(\big(A(s), L(s), N^a(s), N^\ell(s), N^y(s)\big) : s \leq t\big) \vee \sigma((A(0), L(0)))$$

$$\mathcal{F}_t^{\text{full}} = \sigma\big(\big(A(s), L(s), N^a(s), N^\ell(s), N^y(s), N^c(s)\big) : s \leq t\big) \vee \sigma((A(0), L(0)))$$

- $\mathcal{F}_t$: natural filtration for the processes without censoring
- $\mathcal{F}_t^{\text{full}}$: natural filtration for the processes including censoring

$$\mathcal{F}_t = \sigma\Big(\big(A(s), L(s), N^a(s), N^\ell(s), N^y(s)\big) : s \leq t\Big) \vee \sigma((A(0), L(0)))$$

$$\mathcal{F}_t^{\text{full}} = \sigma\Big(\big(A(s), L(s), N^a(s), N^\ell(s), N^y(s), N^c(s)\big) : s \leq t\Big) \vee \sigma((A(0), L(0)))$$

$$\bar{\mathcal{F}}_t = \sigma\Big(\big(A(s \wedge C), L(s \wedge C), N^a(s \wedge C), N^\ell(s \wedge C), N^y(s \wedge C), N^c(s \wedge T^e)\big) : s \leq t\Big) \vee \sigma((A(0), L(0)))$$

- $\mathcal{F}_t$: natural filtration for the processes without censoring
- $\mathcal{F}_t^{\text{full}}$: natural filtration for the processes including censoring
- $\bar{\mathcal{F}}_t$: observed (natural) filtration

$$\mathcal{F}_t = \sigma\big(\big(A(s), L(s), N^a(s), N^\ell(s), N^y(s)\big) : s \le t\big) \vee \sigma((A(0), L(0)))$$
$$\mathcal{F}_t^{\text{full}} = \sigma\big(\big(A(s), L(s), N^a(s), N^\ell(s), N^y(s), N^c(s)\big) : s \le t\big) \vee \sigma((A(0), L(0)))$$
$$\bar{\mathcal{F}}_t = \sigma\big(\big(A(s \wedge C), L(s \wedge C), N^a(s \wedge C), N^\ell(s \wedge C), N^y(s \wedge C), N^c(s \wedge T^e)\big) : s \le t\big) \vee \sigma((A(0), L(0)))$$

- $\mathcal{F}_t$: natural filtration for the processes without censoring
- $\mathcal{F}_t^{\text{full}}$: natural filtration for the processes including censoring
- $\bar{\mathcal{F}}_t$: observed (natural) filtration
- Data format (uncensored)

$$\left(T_{(K)}, \Delta_{(K)}, A\big(T_{(K-1)}\big), L\big(T_{(K-1)}\big), \underbrace{T_{(K-1)}}_{\text{ordered event time}}, \underbrace{\Delta_{(K-1)}}_{\substack{\in\{a,y,\ell\} \\ \text{status indicator}}}, ..., A(0), L(0)\right)$$

- Data format (censored)

$$\big(\bar{T}_{(K)}, \bar{\Delta}_{(K)}, A\big(\bar{T}_{K-1}\big), L\big(\bar{T}_{K-1}\big), \bar{T}_{(K-1)}, \bar{\Delta}_{(K-1)}, ..., A(0), L(0)\big)$$

- **Random measure** N_t^{a*}: random measure associated to N_a and A given by

$$N_t^{a*} = \sum_{k:\Delta_{(k)}=a} \delta_{\left(T_{(k)}, A\left(T_{(k)}\right)\right)}$$

- **Random measure** N_t^{a*}: random measure associated to N_a and A given by

$$N_t^{a*} = \sum_{k:\Delta_{(k)}=a} \delta_{\left(T_{(k)}, A\left(T_{(k)}\right)\right)}$$

- **Intervention**
  - Modify compensator: $\Lambda_t^{a*}(\cdot) = \pi_t(\cdot)\Lambda^a(t)$
  - Replace treatment mechanism
    - $\pi_t(\{x\}) = P(A(t) = x \mid \mathcal{F}_{t-})$
    - Under new law P^{G^*}, compensator of N^a is $\pi_t^*(\cdot)\Lambda^a(t)$ with respect to $\mathcal{F}_t$

- **Random measure** N^{a*}_t: random measure associated to N_a and A given by

$$N^{a*}_t = \sum_{k:\Delta_{(k)}=a} \delta_{\left(T_{(k)}, A\left(T_{(k)}\right)\right)}$$

- **Intervention**
  - Modify compensator: $\Lambda^{a*}_t(\cdot) = \pi_t(\cdot)\Lambda^a(t)$
  - Replace treatment mechanism
    - $\pi_t(\{x\}) = P(A(t) = x \mid \mathcal{F}_{t-})$
    - Under new law P^{G^*}, compensator of N^a is $\pi^*_t(\cdot)\Lambda^a(t)$ with respect to $\mathcal{F}_t$
- **Special case**
  - $\pi^*_t(\{x\}) \equiv \mathbb{1}\{x = 1\}$ (stay on treatment)

- **Random measure** N_t^{a*}: random measure associated to N_a and A given by

$$N_t^{a*} = \sum_{k:\Delta_{(k)}=a} \delta_{\left(T_{(k)}, A\left(T_{(k)}\right)\right)}$$

- **Intervention**
  - Modify compensator: $\Lambda_t^{a*}(\cdot) = \pi_t(\cdot)\Lambda^a(t)$
  - Replace treatment mechanism
    - $\pi_t(\{x\}) = P(A(t) = x \mid \mathcal{F}_{t-})$
    - Under new law P^{G^*}, compensator of N^a is $\pi_t^*(\cdot)\Lambda^a(t)$ with respect to $\mathcal{F}_t$
- **Special case**
  - $\pi_t^*(\{x\}) \equiv \mathbb{1}\{x = 1\}$ (stay on treatment)
- **Target parameter**
  - $\Psi_\tau(P) = \mathbb{E}_P\left[\frac{\mathrm{d}P^{G^*}}{\mathrm{d}P}(\tau)N^y(\tau)\right] = \mathbb{E}_{P^{G^*}}[N^y(\tau)], \tau < \tau_{\text{end}}$

- **Efficient influence function (EIF)** for $\Psi_\tau(P)$ in the nonparametric model is given by (Rytgaard et al. (2022))

$$
\begin{aligned}
\varphi_\tau^*(P) &= \mathbb{E}_{P^{G^*}}\left[N_y(\tau) \mid \mathcal{F}_0\right] - \Psi_\tau(P) \\
&+ \int_0^\tau \frac{\mathrm{d}P^{G^*}}{\mathrm{d}P}(t-)\left(\mathbb{E}_{P^{G^*}}\left[N_y(\tau) \mid L(t), N^\ell(t), \mathcal{F}_{t-}\right] - \mathbb{E}_{P^{G^*}}\left[N_y(\tau) \mid N^\ell(t), \mathcal{F}_{t-}\right]\right)\widetilde{N}^\ell(\mathrm{d}t) \\
&+ \int_0^\tau \frac{\mathrm{d}P^{G^*}}{\mathrm{d}P}(t-)\left(\mathbb{E}_{P^{G^*}}\left[N_y(\tau) \mid \Delta N^\ell(t) = 1, \mathcal{F}_{t-}\right] - \mathbb{E}_{P^{G^*}}\left[N_y(\tau) \mid \Delta N^\ell(t) = 0, \mathcal{F}_{t-}\right]\right)\widetilde{M}^\ell(\mathrm{d}t) \\
&+ \int_0^\tau \frac{\mathrm{d}P^{G^*}}{\mathrm{d}P}(t-)\left(\mathbb{E}_{P^{G^*}}\left[N_y(\tau) \mid \Delta N^a(t) = 1, \mathcal{F}_{t-}\right] - \mathbb{E}_{P^{G^*}}\left[N_y(\tau) \mid \Delta N^a(t) = 0, \mathcal{F}_{t-}\right]\right)\widetilde{M}^a(\mathrm{d}t) \\
&+ \int_0^\tau \frac{\mathrm{d}P^{G^*}}{\mathrm{d}P}(t-)\left(1 - \mathbb{E}_{P^{G^*}}\left[N_y(\tau) \mid \Delta N^y(t) = 0, \mathcal{F}_{t-}\right]\right)\widetilde{M}^y(\mathrm{d}t).
\end{aligned}
$$

- $\widetilde{M}^x(t) = \widetilde{N}^x(t) - \Lambda^x(t)$ is the P-$\bar{\mathcal{F}}_t$ martingale for $\widetilde{N}^x(t) = N^x(t \wedge C)$.

5.1. Efficient influence function (continued)

- To work within the targeted learning framework, we need the efficient influence function.

5.1. Efficient influence function (continued)

- To work within the targeted learning framework, we need the efficient influence function.
- Rytgaard et al. (2022) propose sequential regressions for estimating terms in $\varphi_\tau^*(P)$.

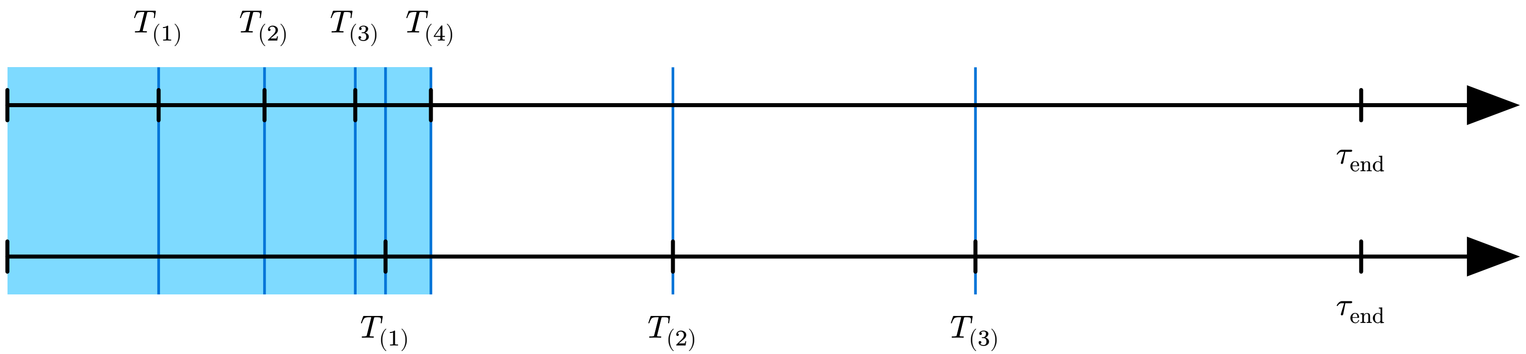
5.1. Efficient influence function (continued)

- To work within the targeted learning framework, we need the efficient influence function.
- Rytgaard et al. (2022) propose sequential regressions for estimating terms in $\varphi_\tau^*(P)$.
- Implementation is unclear and may require thousands of iterations (iterate through all unique event times in the sample).

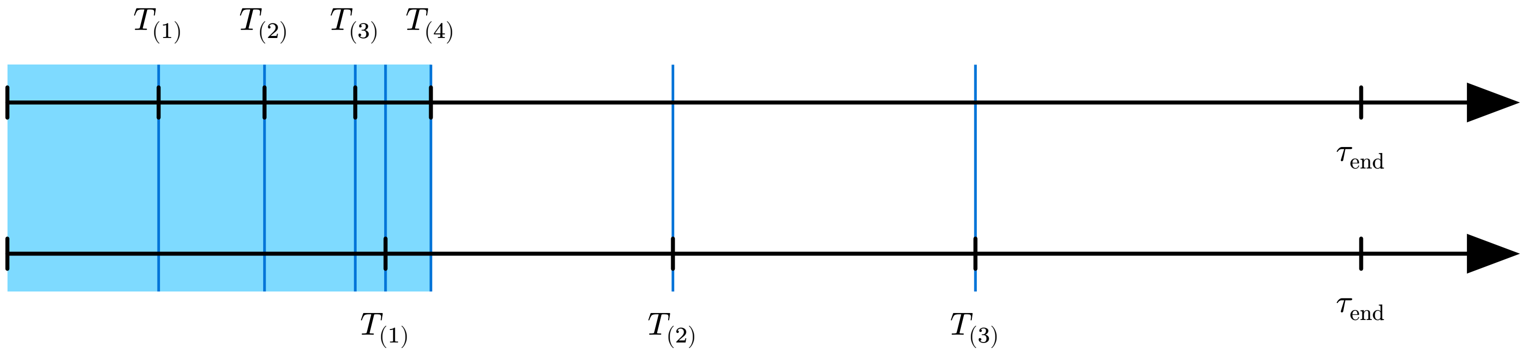
5.1. Efficient influence function (continued)

- To work within the targeted learning framework, we need the efficient influence function.
- Rytgaard et al. (2022) propose sequential regressions for estimating terms in $\varphi_\tau^*(P)$.
- Implementation is unclear and may require thousands of iterations (iterate through all unique event times in the sample).
- **Idea**: Replace $\mathcal{F}_{t-}$ with simpler histories:
  - $\mathcal{F}_{T_{(k)}} = \sigma\left(A\left(T_{(j)}\right), L\left(T_{(j)}\right), T_{(j)}, \Delta_{(j)} : j \leq k\right) \vee \sigma((A(0), L(0)))$
  - Censored versions: $\bar{\mathcal{F}}_{\bar{T}_{(k)}} = \sigma\left(A\left(\bar{T}_j\right), L\left(\bar{T}_j\right), \bar{T}_{(j)}, \bar{\Delta}_{(j)} : j \leq k\right) \vee \sigma((A(0), L(0)))$

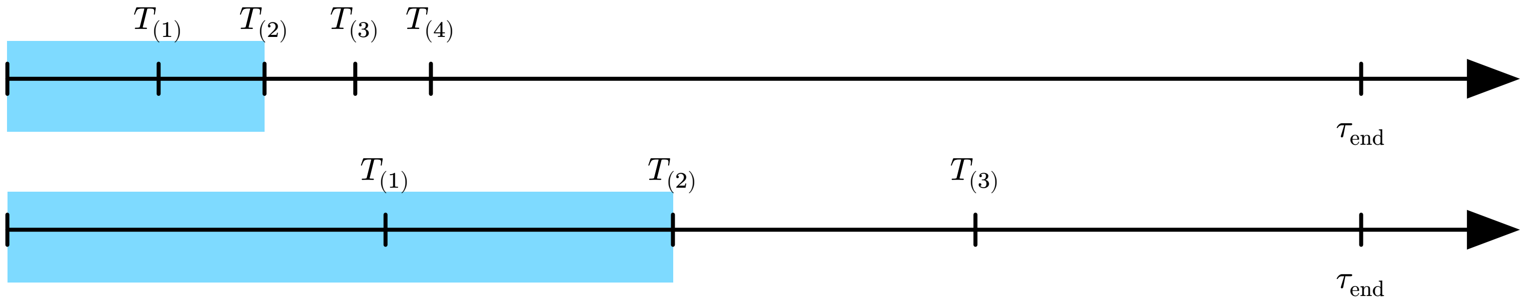
Rytgaard et al. (2022):



Rytgaard et al. (2022):



ICE-IPCW (Ohlendorff et al. (2025)):



- **Propensity score**:
  - $\pi_k\left(t, L\left(T_{(k)}\right), \mathcal{F}_{T_{(k-1)}}\right)$: $P\left(A\left(T_{(k)}\right) = 1 \mid L\left(T_{(k)}\right), T_{(k)} = t, \Delta_{(k)} = a, \mathcal{F}_{T_{(k-1)}}\right)$
- **Hazard measures**:
  - $\tilde{\Lambda}_k^c\left(t \mid \bar{\mathcal{F}}_{\bar{T}_{(k-1)}}\right)$: hazard measure for $\left(\bar{T}_{(k)}, \mathbb{1}\left\{\bar{\Delta}_{(k)} = c\right\}\right)$ given $\bar{\mathcal{F}}_{\bar{T}_{(k-1)}}$
  - $\Lambda_k^x\left(t, \mathcal{F}_{T_{(k-1)}}\right)$: hazard measure of $\left(T_{(k)}, \mathbb{1}\left\{\Delta_{(k)} = x\right\}\right)$ given $\mathcal{F}_{T_{(k-1)}}$ for $x \in \{a, \ell, y\}$

- **Propensity score**:
  - $\pi_k\left(t, L\left(T_{(k)}\right), \mathcal{F}_{T_{(k-1)}}\right)$: $P\left(A\left(T_{(k)}\right) = 1 \mid L\left(T_{(k)}\right), T_{(k)} = t, \Delta_{(k)} = a, \mathcal{F}_{T_{(k-1)}}\right)$
- **Hazard measures**:
  - $\tilde{\Lambda}_k^c\left(t \mid \bar{\mathcal{F}}_{\bar{T}_{(k-1)}}\right)$: hazard measure for $\left(\bar{T}_{(k)}, \mathbb{1}\left\{\bar{\Delta}_{(k)} = c\right\}\right)$ given $\bar{\mathcal{F}}_{\bar{T}_{(k-1)}}$
  - $\Lambda_k^x\left(t, \mathcal{F}_{T_{(k-1)}}\right)$: hazard measure of $\left(T_{(k)}, \mathbb{1}\left\{\Delta_{(k)} = x\right\}\right)$ given $\mathcal{F}_{T_{(k-1)}}$ for $x \in \{a, \ell, y\}$
- $\prod$: product integral (Gill & Johansen (1990))

- **Propensity score**:
  - $\pi_k\left(t, L\left(T_{(k)}\right), \mathcal{F}_{T_{(k-1)}}\right)$: $P\left(A\left(T_{(k)}\right) = 1 \mid L\left(T_{(k)}\right), T_{(k)} = t, \Delta_{(k)} = a, \mathcal{F}_{T_{(k-1)}}\right)$
- **Hazard measures**:
  - $\tilde{\Lambda}_k^c\left(t \mid \bar{\mathcal{F}}_{\bar{T}_{(k-1)}}\right)$: hazard measure for $\left(\bar{T}_{(k)}, \mathbb{1}\left\{\bar{\Delta}_{(k)} = c\right\}\right)$ given $\bar{\mathcal{F}}_{\bar{T}_{(k-1)}}$
  - $\Lambda_k^x\left(t, \mathcal{F}_{T_{(k-1)}}\right)$: hazard measure of $\left(T_{(k)}, \mathbb{1}\left\{\Delta_{(k)} = x\right\}\right)$ given $\mathcal{F}_{T_{(k-1)}}$ for $x \in \{a, \ell, y\}$
- $\prod$: product integral (Gill & Johansen (1990))
- **Survival functions**:
  - $\tilde{S}^c\left(t \mid \bar{\mathcal{F}}_{\bar{T}_{(k-1)}}\right) = \prod_{s \in \left(\bar{T}_{(k-1)}, t\right]}\left(1 - d\tilde{\Lambda}_k^c\left(s \mid \bar{\mathcal{F}}_{\bar{T}_{(k-1)}}\right)\right)$
  - $S\left(t \mid \mathcal{F}_{T_{(k-1)}}\right) = \prod_{s \in \left(T_{(k-1)}, t\right]}\left(1 - \sum_{x=a,\ell,y,d} d\Lambda_k^x\left(s \mid \mathcal{F}_{T_{(k-1)}}\right)\right)$

- **Propensity score**:
  - $\pi_k\left(t, L\left(T_{(k)}\right), \mathcal{F}_{T_{(k-1)}}\right): P\left(A\left(T_{(k)}\right) = 1 \mid L\left(T_{(k)}\right), T_{(k)} = t, \Delta_{(k)} = a, \mathcal{F}_{T_{(k-1)}}\right)$
- **Hazard measures**:
  - $\tilde{\Lambda}^c_k\left(t \mid \bar{\mathcal{F}}_{\bar{T}_{(k-1)}}\right)$: hazard measure for $\left(\bar{T}_{(k)}, \mathbb{1}\left\{\bar{\Delta}_{(k)} = c\right\}\right)$ given $\bar{\mathcal{F}}_{\bar{T}_{(k-1)}}$
  - $\Lambda^x_k\left(t, \mathcal{F}_{T_{(k-1)}}\right)$: hazard measure of $\left(T_{(k)}, \mathbb{1}\left\{\Delta_{(k)} = x\right\}\right)$ given $\mathcal{F}_{T_{(k-1)}}$ for $x \in \{a, \ell, y\}$
- $\prod$: product integral (Gill & Johansen (1990))
- **Survival functions**:
  - $\tilde{S}^c\left(t \mid \bar{\mathcal{F}}_{\bar{T}_{(k-1)}}\right) = \prod_{s \in \left(\bar{T}_{(k-1)}, t\right]}\left(1 - d\tilde{\Lambda}^c_k\left(s \mid \bar{\mathcal{F}}_{\bar{T}_{(k-1)}}\right)\right)$
  - $S\left(t \mid \mathcal{F}_{T_{(k-1)}}\right) = \prod_{s \in \left(T_{(k-1)}, t\right]}\left(1 - \sum_{x=a,\ell,y,d} d\Lambda^x_k\left(s \mid \mathcal{F}_{T_{(k-1)}}\right)\right)$
- **Random measure**:
  - $N = \sum_k \delta_{\left(T_{(k)}, \Delta_{(k)}, A\left(T_{(k)}\right), L\left(T_{(k)}\right)\right)}$
  - Turns out that natural filtration of N is $\mathcal{F}^{\text{full}}_t$ (Chapter 2.5 of Last & Brandt (1995))

7.1. Independent censoring conditions

- Let $\bar{Q}^g_{K,\tau} : (a_k, h_k) \mapsto 0$

7.1. Independent censoring conditions

- Let $\bar{Q}^g_{K,\tau} : (a_k, h_k) \mapsto 0$
- Define recursively, for $k = K, ..., 0$,

$$\bar{Z}^a_{k,\tau}(u) = \frac{1}{\tilde{S}^c\left(\bar{T}_{(k)} - \mid A(\bar{T}_{k-1}), \bar{H}_{k-1}\right)}\Big(\mathbb{1}\left\{\bar{T}_{(k)} \leq u, \bar{T}_{(k)} < \tau, \bar{\Delta}_{(k)} = a\right\}\bar{Q}^g_{k,\tau}\left(1, \bar{H}_k\right)$$
$$+\mathbb{1}\left\{\bar{T}_{(k)} \leq u, \bar{T}_{(k)} < \tau, \bar{\Delta}_{(k)} = \ell\right\}\bar{Q}^g_{k,\tau}\left(A\left(\bar{T}_k\right), \bar{H}_k\right)$$
$$+\mathbb{1}\left\{\bar{T}_{(k)} \leq u, \bar{\Delta}_{(k)} = y\right\}\Big),$$

and

$$\bar{Q}^g_{k,\tau} : (u, a_k, h_k) \mapsto \mathbb{E}_P\left[\bar{Z}^a_{k+1,\tau}(u) \mid A\left(\bar{T}_k\right) = a_k, \bar{H}_k = h_k\right], \quad u \leq \tau$$

where $h_k = (a_k, l_k, t_k, d_k, ..., a_0, l_0)$.

7.2. ICE-IPCW procedure: Consistency

Theorem 7.2.1 Assume that the P-$\mathcal{F}_t^{\text{full}}$ compensator Λ of N is also the P-$\mathcal{F}_t$ compensator of N.

7.2. ICE-IPCW procedure: Consistency

Theorem 7.2.1 Assume that the P-$\mathcal{F}_t^{\text{full}}$ compensator Λ of N is also the P-$\mathcal{F}_t$ compensator of N. If

1. $\Delta\tilde{\Lambda}_k^c\left(\cdot, \mathcal{F}_{\bar{T}_{(k-1)}}\right)\Delta\Lambda_k^x\left(\cdot, \mathcal{F}_{T_{(k-1)}}\right) \equiv 0$ for $x \in \{a, \ell, y\}$ and $k \in \{1, \ldots, K\}$.
2. $\tilde{S}^c\left(t \mid \mathcal{F}_{\bar{T}_{(k-1)}}\right) > \eta$ for all $t \in (0, \tau]$ and $k \in \{1, \ldots, K\}$ P-a.s. for some $\eta > 0$.

7.2. ICE-IPCW procedure: Consistency

Theorem 7.2.1 Assume that the P-$\mathcal{F}_t^{\text{full}}$ compensator Λ of N is also the P-$\mathcal{F}_t$ compensator of N. If

1. $\Delta\tilde{\Lambda}_k^c\left(\cdot, \mathcal{F}_{\bar{T}_{(k-1)}}\right)\Delta\Lambda_k^x\left(\cdot, \mathcal{F}_{T_{(k-1)}}\right) \equiv 0$ for $x \in \{a, \ell, y\}$ and $k \in \{1, \dots, K\}$.
2. $\tilde{S}^c\left(t \mid \mathcal{F}_{\bar{T}_{(k-1)}}\right) > \eta$ for all $t \in (0, \tau]$ and $k \in \{1, \dots, K\}$ P-a.s. for some $\eta > 0$.

It holds that

$$\Psi_\tau^g(P) = \mathbb{E}_P\left[\bar{Q}_{0,\tau}^g(\tau, 1, L(0))\right].$$

7.2. ICE-IPCW procedure: Consistency

Theorem 7.2.1 Assume that the P-$\mathcal{F}_t^{\text{full}}$ compensator Λ of N is also the P-$\mathcal{F}_t$ compensator of N. If

1. $\Delta\tilde{\Lambda}_k^c\left(\cdot, \mathcal{F}_{\bar{T}_{(k-1)}}\right)\Delta\Lambda_k^x\left(\cdot, \mathcal{F}_{T_{(k-1)}}\right) \equiv 0$ for $x \in \{a, \ell, y\}$ and $k \in \{1, \dots, K\}$.
2. $\tilde{S}^c\left(t \mid \mathcal{F}_{\bar{T}_{(k-1)}}\right) > \eta$ for all $t \in (0, \tau]$ and $k \in \{1, \dots, K\}$ P-a.s. for some $\eta > 0$.

It holds that

$$\Psi_\tau^g(P) = \mathbb{E}_P\left[\bar{Q}_{0,\tau}^g(\tau, 1, L(0))\right].$$

- We make *explicit* use of the fact that the compensator can be explicitly written in terms of the regular conditional distributions of the variables $\left(\bar{T}_{(k)}, \bar{\Delta}_{(k)}, A\left(\bar{T}_k\right), L\left(\bar{T}_k\right)\right)$, $k = 1, \dots, K$ and $(A(0), L(0))$.

Theorem 8.1 Under suitable regularity conditions, $\varphi_\tau^*(P)$ can be rewritten as

$$\varphi_\tau^*(P) = \frac{\mathbb{1}\{A(0)=1\}}{\pi_0(L(0))} \sum_{k=1}^{K} \prod_{j=1}^{k-1} \left(\frac{\mathbb{1}\{A(\bar{T}_j)=1\}}{\pi_j\left(\bar{T}_{(j)}, L(\bar{T}_j), \bar{\mathcal{F}}_{\bar{T}_{(j-1)}}\right)} \right)^{\mathbb{1}\{\bar{\Delta}_{(j)}=a\}} \frac{1}{\prod_{j=1}^{k-1} \tilde{S}^c\left(\bar{T}_{(j)} - \mid \bar{\mathcal{F}}_{\bar{T}_{(j-1)}}\right)}$$

$$\times \mathbb{1}\left\{\bar{\Delta}_{(k-1)} \in \{\ell, a\}, \bar{T}_{(k-1)} < \tau\right\} \left(\left(\bar{Z}_{k,\tau}^a(\tau) - \bar{Q}_{k-1,\tau}^g(\tau)\right) \right.$$

$$\left. + \int_{\bar{T}_{(k-1)}}^{\tau \wedge \bar{T}_{(k)}} \left(\bar{Q}_{k-1,\tau}^g(\tau) - \bar{Q}_{k-1,\tau}^g(u)\right) \frac{1}{\tilde{S}^c\left(u \mid \bar{\mathcal{F}}_{\bar{T}_{(k-1)}}\right) S\left(u - \mid \bar{\mathcal{F}}_{\bar{T}_{(k-1)}}\right)} \widetilde{M}^c(\mathrm{d}u) \right)$$

$$+ \bar{Q}_{0,\tau}^g(\tau) - \Psi_\tau^g(P).$$

- **Estimator**
  - One-step estimator based on the efficient influence function

- **Estimator**
  - One-step estimator based on the efficient influence function
- **Simulations**
  - Lower bias than LTMLE (van der Laan & Gruber (2012)) and good CI coverage.
  - Mean squared errors are however close to being the same.
- **Application**: Apply to real-world data with Danish registry data:
  - EIF provides confidence intervals comparable to bootstrap CIs.

- **Estimator**
  - One-step estimator based on the efficient influence function
- **Simulations**
  - Lower bias than LTMLE (van der Laan & Gruber (2012)) and good CI coverage.
  - Mean squared errors are however close to being the same.
- **Application**: Apply to real-world data with Danish registry data:
  - EIF provides confidence intervals comparable to bootstrap CIs.
- **Challenges**
  - Variance estimation is difficult (censoring martingale term)
  - Possible approaches:
    - Undersmooth estimation of censoring compensator
    - Machine learning (e.g., neural networks) methods for multivariate outcomes (to $\bar{Q}^g_{k,\tau}(u)$)

- **Estimator**
  - One-step estimator based on the efficient influence function
- **Simulations**
  - Lower bias than LTMLE (van der Laan & Gruber (2012)) and good CI coverage.
  - Mean squared errors are however close to being the same.
- **Application**: Apply to real-world data with Danish registry data:
  - EIF provides confidence intervals comparable to bootstrap CIs.
- **Challenges**
  - Variance estimation is difficult (censoring martingale term)
  - Possible approaches:
    - Undersmooth estimation of censoring compensator
    - Machine learning (e.g., neural networks) methods for multivariate outcomes (to $\bar{Q}^g_{k,\tau}(u)$)
- **Next steps**
  - Empirical process & remainder term conditions not yet addressed (ongoing work)
  - Consider TMLE instead of one-step $\Rightarrow$ ensures estimates in $[0, 1]$
  - Apply flexible, data-adaptive estimators for nuisance parameters
  - Clarify causal interpretation of target parameter (identifiability)

Bibliography

Adams, R., Saria, S., & Rosenblum, M. (2020). The impact of time series length and discretization on longitudinal causal estimation methods. *Arxiv Preprint Arxiv:2011.15099*.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, *21*(1), C1–C68. https://doi.org/10.1111/ectj.12097

Ferreira Guerra, S., Schnitzer, M. E., Forget, A., & Blais, L. (2020). Impact of discretization of the timeline for longitudinal causal inference methods. *Statistics in Medicine*, *39*(27), 4069–4085. https://doi.org/https://doi.org/10.1002/sim.8710

Gill, R. D., & Johansen, S. (1990). A survey of product-integration with a view toward application in survival analysis. *The Annals of Statistics*, 1501–1555.

Kant, W. M., & Krijthe, J. H. (2025). Irregular measurement times in estimating time-varying treatment effects: Categorizing biases and comparing adjustment methods. *Arxiv Preprint Arxiv:2501.11449*.

Last, G., & Brandt, A. (1995). *Marked Point Processes on the Real Line: The Dynamical Approach*. Springer. https://link.springer.com/book/9780387945477

Ohlendorff, J. S., Munch, A., & Gerds, T. A. (2025). *A Novel Sequential Regression Approach for Efficient Continuous-Time Causal Inference.*

Ryalen, P. C., Stensrud, M. J., & Røysland, K. (2019). The additive hazard estimator is consistent for continuous-time marginal structural models. *Lifetime Data Analysis*, *25*, 611–638.

Rytgaard, H. C., Gerds, T. A., & van der Laan, M. J. (2022). Continuous-Time Targeted Minimum Loss-Based Estimation of Intervention-Specific Mean Outcomes. *The Annals of Stat Istics*, *50*(5), 2469–2491. https://doi.org/10.1214/21-AOS2114

Sofrygin, O., Zhu, Z., Schmittdiel, J. A., Adams, A. S., Grant, R. W., van der Laan, M. J., & Neugebauer, R. (2019). Targeted learning with daily EHR data. *Statistics in Medicine*, *38*(16), 3073–3090.

Sun, J., & Crawford, F. W. (2023). The role of discretization scales in causal inference with continuous-time treatment. *Arxiv Preprint Arxiv:2306.08840.*

van der Laan, M. J., & Gruber, S. (2012). *The International Journal of Biostatistics*, *8*(1). https://doi.org/doi:10.1515/1557-4679.1370

van der Laan, M. J., & Rubin, D. (2006). Targeted Maximum Likelihood Learning. *The International Journal of Biostatistics*, *2*(1). https://doi.org/10.2202/1557-4679.1043