Sequential Regressions for Efficient Continuous-Time Causal Inference

Johan Sebastian Ohlendorff

2025-09-01

- 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 Rytgaard et al. (2022) or one-step estimation).
  - Efficiency theory.

- We observe a càdlàg, jump process for the treatment $(A(t))_{t \in [0, \tau_{\text{end}}]} \in \{0, 1\}$ and a covariate process $(L(t))_{t \in [0, \tau_{\text{end}}]}$, such that $L(t)$ almost surely takes values some finite subset of $\mathbb{R}^d$.

- We observe a càdlàg, jump process for the treatment $(A(t))_{t \in [0, \tau_{\text{end}}]} \in \{0, 1\}$ and a covariate process $(L(t))_{t \in [0, \tau_{\text{end}}]}$, such that $L(t)$ almost surely takes values some finite subset of $\mathbb{R}^d$.
- Assume that we are observe the counting processes N^x, $x = a, \ell, y$ (treatment, covariate, death, censoring) up to a right-censoring time C which is distinct from all event times with probability 1. Terminal event time is denoted by T^e.

- We observe a càdlàg, jump process for the treatment $(A(t))_{t \in [0, \tau_{\text{end}}]} \in \{0, 1\}$ and a covariate process $(L(t))_{t \in [0, \tau_{\text{end}}]}$, such that $L(t)$ almost surely takes values some finite subset of $\mathbb{R}^d$.
- Assume that we are observe the counting processes $N^x, x = a, \ell, y$ (treatment, covariate, death, censoring) up to a right-censoring time C which is distinct from all event times with probability 1. Terminal event time is denoted by T^e.
- Assume that $\Delta A(t) \neq 0$ only if $\Delta N^a(t) \neq 0$ and $\Delta L(t) \neq 0$ only if $\Delta N^\ell(t) \neq 0$ or $\Delta N^a(t) \neq 0$.

- We observe a càdlàg, jump process for the treatment $(A(t))_{t \in [0, \tau_{\text{end}}]} \in \{0, 1\}$ and a covariate process $(L(t))_{t \in [0, \tau_{\text{end}}]}$, such that $L(t)$ almost surely takes values some finite subset of $\mathbb{R}^d$.
- Assume that we are observe the counting processes N^x, $x = a, \ell, y$ (treatment, covariate, death, censoring) up to a right-censoring time C which is distinct from all event times with probability 1. Terminal event time is denoted by T^e.
- Assume that $\Delta A(t) \neq 0$ only if $\Delta N^a(t) \neq 0$ and $\Delta L(t) \neq 0$ only if $\Delta N^\ell(t) \neq 0$ or $\Delta N^a(t) \neq 0$. It then seems reasonable to assume that $\Delta N^a \Delta N^\ell \equiv 0$ and that, in fact, every counting process does not jump at the same time as any other counting process.

- We observe a càdlàg, jump process for the treatment $(A(t))_{t \in [0, \tau_{\text{end}}]} \in \{0, 1\}$ and a covariate process $(L(t))_{t \in [0, \tau_{\text{end}}]}$, such that $L(t)$ almost surely takes values some finite subset of $\mathbb{R}^d$.
- Assume that we are observe the counting processes N^x, $x = a, \ell, y$ (treatment, covariate, death, censoring) up to a right-censoring time C which is distinct from all event times with probability 1. Terminal event time is denoted by T^e.
- Assume that $\Delta A(t) \neq 0$ only if $\Delta N^a(t) \neq 0$ and $\Delta L(t) \neq 0$ only if $\Delta N^\ell(t) \neq 0$ or $\Delta N^a(t) \neq 0$. It then seems reasonable to assume that $\Delta N^a \Delta N^\ell \equiv 0$ and that, in fact, every counting process does not jump at the same time as any other counting process.
- The doctor may decide treatment based at times at which $\Delta N^a(t) \neq 0$. The intervention in which we are interested attempts to specify what this decision should be (or the probability of being treated), but does not naturally intervene on when the doctor decides to do so.

- We observe a càdlàg, jump process for the treatment $(A(t))_{t \in [0, \tau_{\text{end}}]} \in \{0, 1\}$ and a covariate process $(L(t))_{t \in [0, \tau_{\text{end}}]}$, such that $L(t)$ almost surely takes values some finite subset of $\mathbb{R}^d$.
- Assume that we are observe the counting processes N^x, $x = a, \ell, y$ (treatment, covariate, death, censoring) up to a right-censoring time C which is distinct from all event times with probability 1. Terminal event time is denoted by T^e.
- Assume that $\Delta A(t) \neq 0$ only if $\Delta N^a(t) \neq 0$ and $\Delta L(t) \neq 0$ only if $\Delta N^\ell(t) \neq 0$ or $\Delta N^a(t) \neq 0$. It then seems reasonable to assume that $\Delta N^a \Delta N^\ell \equiv 0$ and that, in fact, every counting process does not jump at the same time as any other counting process.
- The doctor may decide treatment based at times at which $\Delta N^a(t) \neq 0$. The intervention in which we are interested attempts to specify what this decision should be (or the probability of being treated), but does not naturally intervene on when the doctor decides to do so.
- Each individual has at most K events in $[0, \tau_{\text{end}}]$, i.e., $\sum_{x=a,y,c,\ell} N^x(\tau_{\text{end}}) \leq K$ almost surely.

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

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

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

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

- $\mathcal{F}_t$ is the natural filtration for the processes without censoring.
- $\mathcal{F}_t^\beta$ is the natural filtration for the processes including censoring.

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

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

$$\begin{aligned}\mathcal{F}_t^{\tilde{\beta}} = \sigma\big(\big(&A(s \wedge C), L(s \wedge C), N^a(s \wedge C), N^\ell(s \wedge C), \\ &N^y(s \wedge C), N^c(s \wedge T^e)\big) : s \leq t\big)\end{aligned}$$

- $\mathcal{F}_t$ is the natural filtration for the processes without censoring.
- $\mathcal{F}_t^\beta$ is the natural filtration for the processes including censoring.
- $\mathcal{F}_t^{\tilde{\beta}}$ is the observed filtration, i.e., the natural filtration stopped by death and censoring.

- Data format:

$$O = \left(T_{(K)}, \Delta_{(K)}, A\left(T_{(K-1)}\right), L\left(T_{(K-1)}\right), T_{(K-1)}, \Delta_{(K-1)}, ..., A(0), L(0)\right)$$

- Data format:

$$O = \left(T_{(K)}, \Delta_{(K)}, A\left(T_{(K-1)}\right), L\left(T_{(K-1)}\right), T_{(K-1)}, \Delta_{(K-1)}, \ldots, A(0), L(0)\right)$$

- Let $N^a_t(\cdot)$ denote the random measure associated with N^a and $A(\cdot)$,

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

- Data format:

$$O = \left(T_{(K)}, \Delta_{(K)}, A\left(T_{(K-1)}\right), L\left(T_{(K-1)}\right), T_{(K-1)}, \Delta_{(K-1)}, \ldots, A(0), L(0)\right)$$

- Let $N_t^a(\cdot)$ denote the random measure associated with N^a and $A(\cdot)$,

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

- Interested on "intervening" on the compensator of $N^a(\cdot)$. Generally, this can be written as $\Lambda_t^a(A) = \pi_t(A)\Lambda^a(t)$.

- Data format:

$$O = \left(T_{(K)}, \Delta_{(K)}, A\left(T_{(K-1)}\right), L\left(T_{(K-1)}\right), T_{(K-1)}, \Delta_{(K-1)}, ..., A(0), L(0)\right)$$

- Let $N^a_t(\cdot)$ denote the random measure associated with N^a and $A(\cdot)$,

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

- Interested on "intervening" on the compensator of $N^a(\cdot)$. Generally, this can be written as $\Lambda^a_t(A) = \pi_t(A)\Lambda^a(t)$.
- The intervention defines a probability measure P^{G^*}, where $N^a(\cdot)$ has compensator $\pi^*_t(A)\Lambda^a(t)$ for specified $\pi^*_t(A)$.

- Data format:

$$O = \left(T_{(K)}, \Delta_{(K)}, A\left(T_{(K-1)}\right), L\left(T_{(K-1)}\right), T_{(K-1)}, \Delta_{(K-1)}, ..., A(0), L(0)\right)$$

- Let $N_t^a(\cdot)$ denote the random measure associated with N^a and $A(\cdot)$,

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

- Interested on "intervening" on the compensator of $N^a(\cdot)$. Generally, this can be written as $\Lambda_t^a(A) = \pi_t(A)\Lambda^a(t)$.
- The intervention defines a probability measure P^{G^*}, where $N^a(\cdot)$ has compensator $\pi_t^*(A)\Lambda^a(t)$ for specified $\pi_t^*(A)$.
- Focus on the case $\pi_t^*(\{x\}) \equiv \mathbb{1}\{x = 1\}$.

4.1 Target parameter (continued)

- We are then interested (are we?) in

$$\Psi_\tau(P) = \mathbb{E}_P\left[\frac{dP^{G^*}}{dP}(\tau) N^y(\tau)\right] = \mathbb{E}_{P^{G^*}}[N^y(\tau)]$$

4.1 Target parameter (continued)

- We are then interested (are we?) in

$$\Psi_\tau(P) = \mathbb{E}_P\left[\frac{dP^{G^*}}{dP}(\tau) N^y(\tau)\right] = \mathbb{E}_{P^{G^*}}[N^y(\tau)]$$

- With $W^g(t) = \frac{dP^{G^*}}{dP}(t)$, Rytgaard et al. (2022) claims that the following is the EIF:

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

4.2 Sequential regressions

- Here $\widetilde{M}^x(t) = \widetilde{N}^x(t) - \Lambda^x(t)$ is the martingale for $\widetilde{N}^x(t) = N^x(t \wedge C)$.
- The above EIF suggests an estimation procedure based on sequential regressions.

4.2 Sequential regressions

- Here $\widetilde{M}^x(t) = \widetilde{N}^x(t) - \Lambda^x(t)$ is the martingale for $\widetilde{N}^x(t) = N^x(t \wedge C)$.
- The above EIF suggests an estimation procedure based on sequential regressions.
- It is unclear how to estimate $\mathbb{E}_{P^{G^*}}\left[N_y(\tau) \mid \Delta N^x(t), \mathcal{F}_{t-}\right]$.

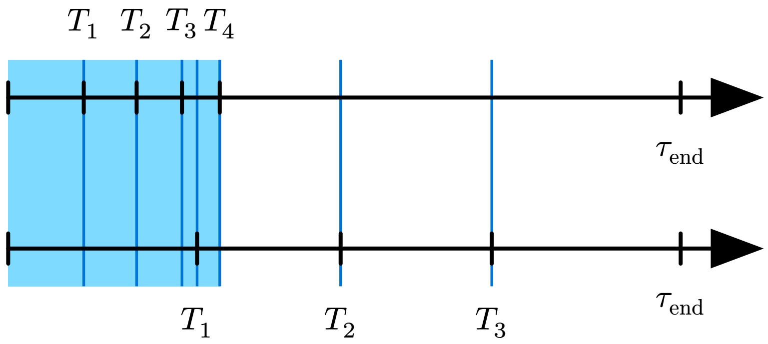
4.2 Sequential regressions

- Here $\widetilde{M}^x(t) = \widetilde{N}^x(t) - \Lambda^x(t)$ is the martingale for $\widetilde{N}^x(t) = N^x(t \wedge C)$.
- The above EIF suggests an estimation procedure based on sequential regressions.
- It is unclear how to estimate $\mathbb{E}_{P^{G^*}}\left[N_y(\tau) \mid \Delta N^x(t), \mathcal{F}_{t-}\right]$.
- Sequential regression not clear how to implement.
- Rytgaard et al. (2022) iterative procedure requires 1000s of iterative steps.
  - Assume that $n = 1000$; if all registrations in the sample are unique and each person has 10 events on average, then we need to fit 10,000 regressions.
- Hard to work with $\mathcal{F}_{t-}$.

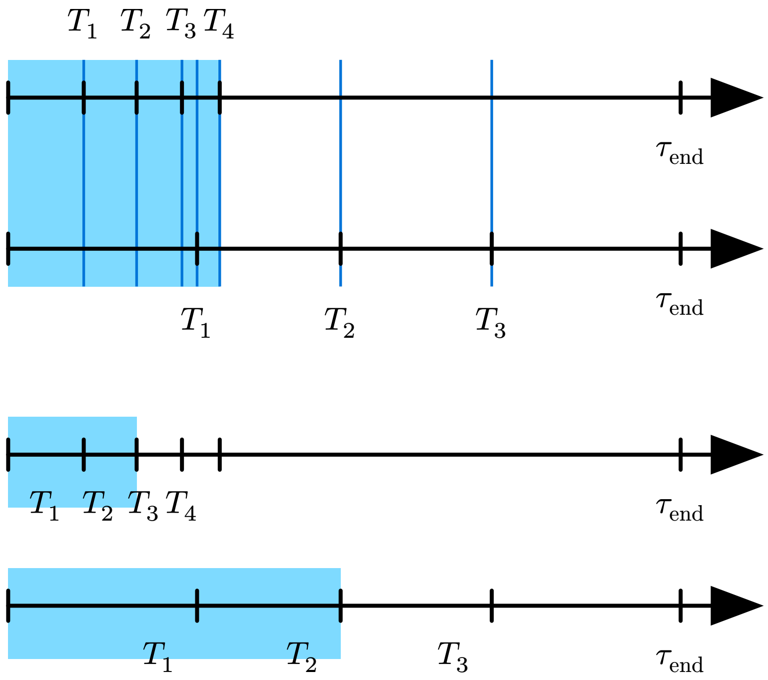
4.2 Sequential regressions

- Here $\widetilde{M}^x(t) = \widetilde{N}^x(t) - \Lambda^x(t)$ is the martingale for $\widetilde{N}^x(t) = N^x(t \wedge C)$.
- The above EIF suggests an estimation procedure based on sequential regressions.
- It is unclear how to estimate $\mathbb{E}_{PG^*}\left[N_y(\tau) \mid \Delta N^x(t), \mathcal{F}_{t-}\right]$.
- Sequential regression not clear how to implement.
- Rytgaard et al. (2022) iterative procedure requires 1000s of iterative steps.
  - Assume that $n = 1000$; if all registrations in the sample are unique and each person has 10 events on average, then we need to fit 10,000 regressions.
- Hard to work with $\mathcal{F}_{t-}$.
- My idea: Can we work with $\mathcal{F}_{T_{(k)}} = \sigma\left(A\left(T_{(j)}\right), L\left(T_{(j)}\right), T_{(j)}, \Delta_{(j)} : j \leq k\right) \vee \sigma((A(0), L(0)))$ instead and more generally $\mathcal{F}^{\tilde{\beta}}_{\bar{T}_{(k)}} = \sigma\left(A\left(\bar{T}_j\right), L\left(\bar{T}_j\right), \bar{T}_{(j)}, \bar{\Delta}_{(j)} : j \leq k\right) \vee \sigma((A(0), L(0)))$ and regress back on that information instead of $\mathcal{F}_{t-}$?

4.3 Illustration



4.3 Illustration



Theorem 5.1

Let $H_k = \left(L\left(T_{(k)}\right), T_{(k)}, \Delta_{(k)}, A\left(T_{(k-1)}\right), L\left(T_{(k-1)}\right), T_{(k-1)}, \Delta_{(k-1)}, \dots, A(0), L(0)\right)$ be the history up to and including the k'th event, but excluding the k'th treatment values for $k > 0$. For $k = 0$, let $H_0 = L(0)$. Let $\bar{Q}^g_{K,\tau} : (a_k, h_k) \mapsto 0$ and recursively define for $k = K - 1, \dots, 1$,

$$\begin{aligned} Z^a_{k+1,\tau}(u) = \mathbb{1}\Big\{T_{(k+1)} \leq u, T_{(k+1)} < \tau, \Delta_{(k+1)} = \ell\Big)\bar{Q}^g_{k+1,\tau}\left(\tau, A\left(T_{(k)}\right), H_{k+1}\right) \\ + \mathbb{1}\Big\{T_{(k+1)} \leq u, T_{(k+1)} < \tau, \Delta_{(k+1)} = a\Big)\bar{Q}^g_{k+1,\tau}\left(\tau, 1, H_{k+1}\right) \\ + \mathbb{1}\Big\{T_{(k+1)} \leq u, \Delta_{(k+1)} = y\Big), \end{aligned}$$

and

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

for $u \leq \tau$. Then,

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

Let

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

Let

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

- $\tilde{S}^c\left(t \mid \mathcal{F}^{\tilde{\beta}}_{\bar{T}_{(k-1)}}\right) = \prod_{s \in \left(\bar{T}_{(k-1)}, t\right]}\left(1 - d\tilde{\Lambda}^c_k\left(s \mid \mathcal{F}^{\tilde{\beta}}_{\bar{T}_{(k-1)}}\right)\right).$

Let

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

- $\tilde{S}^c\left(t \mid \mathcal{F}^{\tilde{\beta}}_{\bar{T}_{(k-1)}}\right) = \prod_{s \in (\bar{T}_{(k-1)}, t]} \left(1 - d\tilde{\Lambda}^c_k\left(s \mid \mathcal{F}^{\tilde{\beta}}_{\bar{T}_{(k-1)}}\right)\right)$.
- $\tilde{\Lambda}^c_k\left(t \mid \mathcal{F}^{\tilde{\beta}}_{\bar{T}_{(k-1)}}\right)$ denotes the hazard measure of $\left(\bar{T}_{(k)}, \mathbb{1}\left\{\bar{\Delta}_{(k)} = c\right\}\right)$ given $\mathcal{F}^{\tilde{\beta}}_{\bar{T}_{(k-1)}}$ and $\Lambda^x_k\left(t, \mathcal{F}_{T_{(k-1)}}\right)$ denotes the hazard measure of $\left(T_{(k)}, \mathbb{1}\left\{\Delta_{(k)} = x\right\}\right)$ given $\mathcal{F}_{T_{(k-1)}}$ for $x \in \{a, \ell, y, d\}$.

Assume that the compensator Λ^α of N^α with respect to the filtration $\mathcal{F}_t^\beta$ is also the compensator with respect to the filtration $\mathcal{F}_t$.

Assume that the compensator Λ^α of N^α with respect to the filtration $\mathcal{F}_t^\beta$ is also the compensator with respect to the filtration $\mathcal{F}_t$. If

1. $\Delta\tilde{\Lambda}_k^c\left(t \mid \mathcal{F}^{\tilde{\beta}}_{\bar{T}_{(k-1)}}\right) + \sum_x \Delta\Lambda_k^x\left(t, \mathcal{F}_{T_{(k-1)}}\right) = 1 \quad P-\text{a.s.} \Rightarrow \Delta\tilde{\Lambda}_{k+1}^c\left(t \mid \mathcal{F}_{T_{(k-1)}}\right) = 1 \quad P-\text{a.s.} \vee \sum_x \Delta\Lambda_k^x\left(t, \mathcal{F}_{T_{(k-1)}}\right) = 1 \quad P-\text{a.s.}$.
2. $\tilde{S}^c\left(t \mid \mathcal{F}^{\tilde{\beta}}_{\bar{T}_{(k-1)}}\right) > \eta$ for all $t \in (0, \tau]$ and $k \in \{1, ..., K\}$ P-a.s. for some $\eta > 0$.

Assume that the compensator Λ^α of N^α with respect to the filtration $\mathcal{F}_t^\beta$ is also the compensator with respect to the filtration $\mathcal{F}_t$. If

1. $\Delta\tilde{\Lambda}_k^c\left(t \mid \mathcal{F}^{\tilde{\beta}}_{\bar{T}_{(k-1)}}\right) + \sum_x \Delta\Lambda_k^x\left(t, \mathcal{F}_{T_{(k-1)}}\right) = 1 \quad P - \text{a.s.} \Rightarrow \Delta\tilde{\Lambda}_{k+1}^c\left(t \mid \mathcal{F}_{T_{(k-1)}}\right) = 1 \quad P - \text{a.s.} \vee \sum_x \Delta\Lambda_k^x\left(t, \mathcal{F}_{T_{(k-1)}}\right) = 1 \quad P - \text{a.s.}$.
2. $\tilde{S}^c\left(t \mid \mathcal{F}^{\tilde{\beta}}_{\bar{T}_{(k-1)}}\right) > \eta$ for all $t \in (0, \tau]$ and $k \in \{1, \dots, K\}$ P-a.s. for some $\eta > 0$.

Then with $h_k = (a_k, l_k, t_k, d_k, \dots, a_0, l_0)$,

$$\mathbb{1}\{d_1 \in \{a, \ell\}, \dots, d_k \in \{a, \ell\}\}\bar{Q}^g_{k,\tau}(u, a_k, h_k) = \mathbb{E}_P\left[\bar{Z}^a_{k+1,\tau}(u) \mid A\left(\bar{T}_k\right) = a_k, \bar{H}_k = h_k\right].$$

Hence $\Psi^g_\tau(P)$ is identifiable from the observed data.

7.1 Rewriting the efficient influence function

- $\widetilde{M}^c(t) = \widetilde{N}^c(t) - \tilde{\Lambda}^c(t)$. Here $\widetilde{N}^c(t) = \mathbb{1}\{C \leq t, T^e > t\} = \sum_{k=1}^{K} \mathbb{1}\left\{\bar{T}_{(k)} \leq t, \bar{\Delta}_{(k)} = c\right\}$ is the censoring counting process.

7.1 Rewriting the efficient influence function

- $\widetilde{M}^c(t) = \widetilde{N}^c(t) - \tilde{\Lambda}^c(t)$. Here $\widetilde{N}^c(t) = \mathbb{1}\{C \leq t, T^e > t\} = \sum_{k=1}^K \mathbb{1}\left\{\bar{T}_{(k)} \leq t, \bar{\Delta}_{(k)} = c\right\}$ is the censoring counting process.
- $S\left(t \mid \mathcal{F}_{T_{(k-1)}}\right) = \prod_{s \in (0,t]} \left(1 - \sum_{x=a,\ell,y,d} \Lambda_k^x\left(\mathrm{d}s \mid \mathcal{F}_{T_{(k-1)}}\right)\right)$.

7.1 Rewriting the efficient influence function

- $\widetilde{M}^c(t) = \widetilde{N}^c(t) - \tilde{\Lambda}^c(t)$. Here $\widetilde{N}^c(t) = \mathbb{1}\{C \leq t, T^e > t\} = \sum_{k=1}^K \mathbb{1}\left\{\bar{T}_{(k)} \leq t, \bar{\Delta}_{(k)} = c\right\}$ is the censoring counting process.
- $S\left(t \mid \mathcal{F}_{T_{(k-1)}}\right) = \prod_{s \in (0,t]} \left(1 - \sum_{x=a,\ell,y,d} \Lambda_k^x\left(\mathrm{d}s \mid \mathcal{F}_{T_{(k-1)}}\right)\right)$.
- Suppose that there is a universal constant $C^* > 0$ such that $\tilde{\Lambda}_k^c\left(\tau_{\text{end}} \mid \mathcal{F}^{\tilde{\beta}}_{\bar{T}_{(k-1)}}; P\right) \leq C^*$ for all $k = 1, \ldots, K$ and every $P \in \mathcal{M}$.

7.2 Rewriting the efficient influence function

The Gateaux derivative is then given by

$$
\begin{aligned}
\varphi_\tau^*(P) = {} & \frac{\mathbb{1}\{A(0)=1\}}{\pi_0(L(0))} \sum_{k=1}^{K} \prod_{j=1}^{k-1} \left(\frac{\mathbb{1}\{A(\bar{T}_j)=1\}}{\pi_j\left(\bar{T}_{(j)}, L(\bar{T}_j), \mathcal{F}^{\tilde{\beta}}_{\bar{T}_{(j-1)}}\right)} \right)^{\mathbb{1}\{\bar{\Delta}_{(j)}=a\}} \frac{1}{\prod_{j=1}^{k-1} \tilde{S}^c\left(\bar{T}_{(j)} - \mid \mathcal{F}^{\tilde{\beta}}_{\bar{T}_{(j-1)}}\right)} \\
& \times \mathbb{1}\left\{\bar{\Delta}_{(k-1)} \in \{\ell, a\}, \bar{T}_{(k-1)} < \tau\right\} \left(\left(\bar{Z}^a_{k,\tau}(\tau) - \bar{Q}^g_{k-1,\tau}(\tau)\right) \right. \\
& \left. + \int_{\bar{T}_{(k-1)}}^{\tau \wedge \bar{T}_{(k)}} \left(\bar{Q}^g_{k-1,\tau}(\tau) - \bar{Q}^g_{k-1,\tau}(u)\right) \frac{1}{\tilde{S}^c\left(u \mid \mathcal{F}^{\tilde{\beta}}_{\bar{T}_{(k-1)}}\right) S\left(u - \mid \mathcal{F}^{\tilde{\beta}}_{\bar{T}_{(k-1)}}\right)} \widetilde{M}^c(\mathrm{d}u) \right) \\
& + \bar{Q}^g_{0,\tau}(\tau) - \Psi^g_\tau(P),
\end{aligned}
$$

7.3 Practical considerations

- We consider a one-step estimator based on the EIF.

7.3 Practical considerations

- We consider a one-step estimator based on the EIF.
- Simulation studies demonstrate favorable performance of the proposed procedure – lower bias than discrete-time procedures and good coverage of confidence intervals.
- However, variance estimation is challenging due to the censoring martingale term.

7.4 Perspectives

- Estimating the martingale term
  - Undersmoothing of the estimation of the censoring compensator to avoid estimation altogether.
  - Using a machine learning methods that can handle multivariate outcomes.

7.4 Perspectives

- Estimating the martingale term
  - Undersmoothing of the estimation of the censoring compensator to avoid estimation altogether.
  - Using a machine learning methods that can handle multivariate outcomes.
- Using a TMLE approach instead of one-step approach $\Rightarrow$ better because we want estimates in $[0, 1]$.

7.4 Perspectives

- Estimating the martingale term
  - Undersmoothing of the estimation of the censoring compensator to avoid estimation altogether.
  - Using a machine learning methods that can handle multivariate outcomes.
- Using a TMLE approach instead of one-step approach $\Rightarrow$ better because we want estimates in $[0, 1]$.
- Identifiability, i.e., does the target parameter have a causal interpretation?

7.4 Perspectives

- Estimating the martingale term
  - Undersmoothing of the estimation of the censoring compensator to avoid estimation altogether.
  - Using a machine learning methods that can handle multivariate outcomes.
- Using a TMLE approach instead of one-step approach $\Rightarrow$ better because we want estimates in $[0, 1]$.
- Identifiability, i.e., does the target parameter have a causal interpretation?
- Similar ideas for other target parameters, e.g., recurrent events, restricted mean survival time, etc.

Bibliography

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