A note on the potential outcomes framework in continuous time

Johan Sebastian Ohlendorff
johan.ohlendorff@sund.ku.dk
University of Copenhagen

Abstract

In this brief note, we consider the target parameters of Ryalen (2024) and compare it with the target parameter given in Rytgaard et al. (2022), corresponding to their marked point process settings. It is shown that the resulting target parameters are the same if and only if the probability of being treated given that you go to the doctor at time t is equal to 1 for Lebesgue-almost all t, provided that the transition hazards for dying are strictly positive for almost all t.

1 Introduction

We consider a multi-state model with at most one visitation time for the treatment (that is at most one point where treatment may change), no time-varying covariates, and no baseline covariates. In the initial state (0) everyone starts as treated. We consider the setting with no censoring. The multi-state model is shown in Figure 1. We observe the counting processes $N_t = (N_t^{01}, N_t^{02}, N_t^{03}, N_t^{13}, N_t^{23})$ on the canonical filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P\right)$, where $\mathcal{F}_t = \sigma(N_s \mid s \leq t)$. This means that we can represent the observed data as $O = \left(T_{(1)}, D_{(1)}, T_{(2)}\right)$, where $T_{(1)}$ is the first event time, $D_{(1)} \in \{01, 02, 03\}$ is the first event type, $A(T_1) = \mathbb{1}\{D_1 \neq 02\}$ is the treatment at the first event time, and $T_{(2)}$ is the second event time, possibly ∞. We will assume that the distribution of the jump times are continuous and that there are no jumps in common between the counting processes. By a well-known result for marked point processes (Proposition 3.1 of Jacod (1975)), we know there exist functions h^{ij}, such that the compensators Λ^{ij} of the counting processes N^{ij} with respect to $P - \mathcal{F}_t$ are given by

$$\Lambda^{0j}(dt) = \mathbb{1}\left\{t \leq T_{(1)}\right\} h^{0j}(t) dt, \quad j = 1, 2, 3$$

$$\Lambda^{i3}(dt) = \mathbb{1}\left\{T_{(1)} < t \leq T_{(2)}\right\} h^{i3}\left(T_{(1)}, t\right) dt, \quad i = 2, 3$$

We let $S_0(t) = \prod_{s \in (0,t]} \left(1 - \sum_j h^{0j}(s) ds\right)$ and $S_1(t \mid d, s) = \prod_{u \in (s,t]} \left(1 - \sum_i h^{i3}(s, u) \mathbb{1}\{d = i\} du\right)$ be the survival functions for the first and second event times, respectively. Furthermore, denote by $P_{0j}(t) = \int_0^t S_{s-} h^{0j}(s) ds$ the probability of having an of type j at time by time t and $P_{i3}(t \mid d, s) = \int_s^t S_1(w - \mid d, s) h^{i3}(s, w) dw$ be the probability of having a terminal at time t given that the first event was of type d at time s.

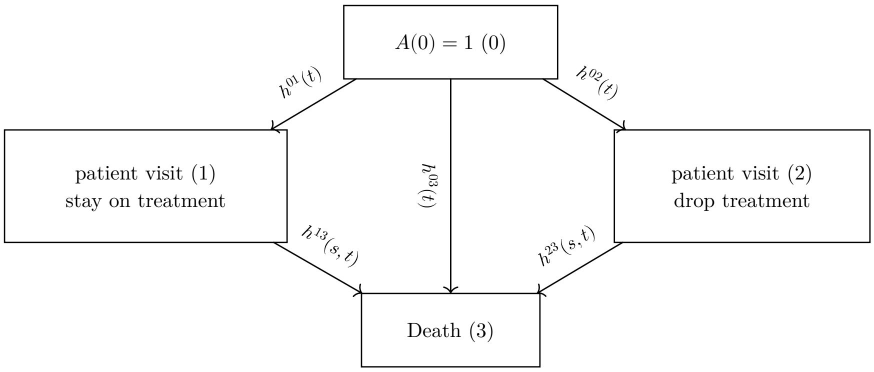

Figure 1: A multi-state model allowing one visitation time for the treatment with the possible treatment values 0/1.

2 The potential outcomes framework

To follow along Ryalen (2024), we restrict the observations to the interval $[0, \tau]$ for $\tau > 0$. We first need to define the intervention of interest, defining the counting processes that we would have like to have observed under the intervention. We can intervene on two components of N (N^{02}, N^{01}), defining the "interventional" processes as

$$N_t^{g,0} = 0$$
$$N_t^{g,1} = N_t^{01} + N_t^{02}$$

This treatment regime defines that the doctor always treats the patient at the visitation time and does not prevent the patient from visiting the doctor if they drop out of the treatment. This thus dictates that an individual that transitioned from 0 to 2 should instead transition to 1. We define $T^{a,g}$ as the first time where the observed and the interventional process deviate.

Define also the single "intervention" process

$$N_t^{g^*,0} = N_t^{g,0} = 0$$

where the interventional component is N^{02}. This dictates that an individual that transitioned from 0 to 2 should not transition to anything at that point. This intuitively thus means that a patient is prevented from visiting the doctor if they drop out of the treatment. The key issue in Ryalen (2024) is that we will not be able to differentiate between target parameters for g and g^*. The reason is that the likelihood under the intervention only depends on the stopping time T^a and the problem that the stopping time T^a is the same under g and g^*.

To see this, let T^{a,g^*} be the first time where the observed and the interventional process (according to g^*) deviate. We have

$$T^{a,g} = \inf_{t>0}\left\{N_t^{g,0} \neq N_t^{01}\right\} \wedge \inf_{t>0}\left\{N_t^{g,1} \neq N_t^{02}\right\} = \inf_{t>0}\left\{N_t^{g,0} \neq 0\right\} = \inf_{t>0}\left\{N_t^{g^*,0} \neq 0\right\} = T^{a,g^*}$$

Applying Theorem 2.1, we find that the target parameters are the same because the weights W_t are the same under g and g^*. Also note that $\mathbb{1}\{T^a \leq t\} = N^{02}(t)$.

We now define the target parameter of interest in Ryalen (2024). The outcome of interest is death at time t, i.e.,

$$Y_t = N_t^{13} + N_t^{03} + N_t^{23} = \mathbb{1}\{T_1 \leq t, D_1 = y\} + \mathbb{1}\{T_2 \leq t\}$$

and we want to estimate $\mathbb{E}_P\left[\tilde{Y}_t\right]$ where $\tilde{Y}_t$ denotes the outcome at time t, had the treatment regime (staying on treatment), possibly contrary to fact, been followed.

Theorem 2.1 (Theorem 1 of Ryalen (2024)): We suppose that there exists a potential outcome process $\left(\tilde{Y}_t\right)_{t\in[0,\tau]}$ such that

1. Consistency: $\tilde{Y}_t \mathbb{1}\left\{T^A > t\right\} = Y_t \mathbb{1}\left\{T^A > t\right\}$ for all $t > 0$ P-a.s.
2. Exchangeability: The $P - \mathcal{F}_t$ compensators Λ^{01}, Λ^{02} are also compensators for $\mathcal{G}_t = \mathcal{F}_t \vee \sigma\left(\tilde{Y}_s, \tau \geq s \geq 0\right)$. Here $\tilde{Y}_s$ is added at baseline, so that $\mathcal{G}_0 = \sigma\left(\tilde{Y}_s, \tau \geq s \geq 0\right)$.
3. Positivity: $W_t = \frac{\mathbb{1}\left\{T^A > t\right\}}{\exp(-\Lambda^{02}(t))} = \frac{1-\mathbb{1}\left\{T_{(1)} \leq t, D_{(1)} = a, A_{(1)} = 0\right\}}{\exp\left(-\int_0^t \mathbb{1}\left\{s \leq T_{(1)}\right\} h^a(s) \pi_s(0) ds\right)}$[1] is a uniformly integrable martingale or equivalently that $R^{\text{Pål}}$ given by $dR^{\text{Pål}} = W_\tau dP$ is a probability measure.

   Then the estimand of interest $\Psi_t^{\text{Ryalen}} : \mathcal{M} \to \mathbb{R}_+$ is identifiable by
   $$\Psi_t^{\text{Ryalen}}(P) := \mathbb{E}_P\left[\tilde{Y}_t\right] = \mathbb{E}_P[Y_t W_t] = \mathbb{E}_{R^{\text{Pål}}}[Y_t]$$

From this, we can derive an alternate representation of the target parameter. We have that

$$\begin{aligned}
\Psi_t^{\text{Ryalen}}(P) &= \mathbb{E}_P\left[\mathbb{1}\left\{T_{(1)} \leq t\right\} Y_t W_t\right] + \mathbb{E}_P\left[\mathbb{1}\left\{T_{(2)} \leq t\right\} Y_t W_t\right] \\
&= \mathbb{E}_P\left[\mathbb{1}\left\{T_{(1)} \leq t\right\} Y_t \frac{1-\mathbb{1}\{T^a > t\}}{\exp\left(-\int_0^{T_{(1)}} h^{02}(s) ds\right)}\right] \\
&\qquad + \mathbb{E}_P\left[\mathbb{1}\left\{T_{(2)} \leq t\right\} Y_t \frac{1-\mathbb{1}\{T^a > t\}}{\exp\left(-\int_0^{T_{(1)}} h^{02}(s) ds\right)}\right] \\
&= \mathbb{E}_P\left[\mathbb{1}\left\{T_{(1)} \leq t, D_{(1)} = 03\right\} \frac{1}{\exp\left(-\int_0^{T_{(1)}} h^{02}(s) ds\right)}\right] \\
&\qquad + \mathbb{E}_P\left[\mathbb{1}\left\{T_{(2)} \leq t, D_{(1)} = 01\right\} \frac{1}{\exp\left(-\int_0^{T_{(1)}} h^{02}(s) ds\right)}\right] \qquad (1) \\
&= \int_0^t \frac{1}{\exp\left(-\int_0^t h^{02}(s) ds\right)} P_{03}(dt) \\
&\qquad + \int_0^t \frac{1}{\exp\left(-\int_0^t h^{02}(s) ds\right)} P_{13}(t \mid 01, s) P_{01}(ds) \\
&= \int_0^t \exp\left(-\sum_{j \neq 2} \int_0^s h^{0j}(u) du\right) h^{03}(s) ds \\
&\qquad + \int_0^t \exp\left(-\sum_{j \neq 2} \int_0^s h^{0j}(u) du\right) P_{13}(t \mid 01, s) h^{01}(s) ds
\end{aligned}$$

2.1 The target parameter in Rytgaard et al. (2022)

To discuss Rytgaard et al. (2022), additionally define

[1] In the notation of Ryalen (2024), $\tau^A = T^a$, $\mathbb{N}_t = \mathbb{1}\left\{T^A \leq t\right\} = N_t^{02}$ and Λ_t^{02} is the compensator of this process.

$$\Lambda^a(t) = (h^{01}(t) + h^{02}(t))\mathbb{1}\left\{T_{(1)} \leq t\right\}$$

$$\pi_t(1) = \frac{h^{01}(t)}{h^{01}(t) + h^{02}(t)}$$

Here, we can interpret $\Lambda^a(t)$ as the cumulative intensity of the visitation times (i.e., $N^a_t = N^{01}_t + N^{02}_t$) and $\pi_t(1)$ as the probability of being treated given that you go to the doctor at time t. Furthermore, let $N^d_t = N^{03}_t + N^{13}_t + N^{23}_t$ be the counting process for the event of interest. Then, its compensator is given by

$$\Lambda^d(dt) = \mathbb{1}\left\{t \leq T_{(1)}\right\}h^{03}(t)dt$$
$$+\mathbb{1}\left\{T_{(1)} < t \leq T_{(2)}\right\}\left(\mathbb{1}\left\{D_{(1)} = 01\right\}h^{13}\left(T_{(1)}, t\right) + \mathbb{1}\left\{D_{(1)} = 02\right\}h^{23}\left(T_{(1)}, t\right)\right)dt$$

Furthermore, let $A(t) = \mathbb{1}\left\{T_{(1)} > t\right\} + \mathbb{1}\left\{T_{(1)} \leq t, D_{(1)} \neq 02\right\}$ be the treatment process at time t. Notationwise, we also define $\Delta N(t) = N_t - N_{t-}$ for a cadlag process N. Rytgaard et al. (2022) give their likelihood as

$$\begin{aligned}
dP(O) &= \prod_{t\in(0,\tau]}\left(d\Lambda^a(t)(\pi_t(1))^{\mathbb{1}\{A(t)=1\}}(1-\pi_t(1))^{\mathbb{1}\{A(t)=0\}}\right)^{\Delta N^a(t)}(1-d\Lambda^a(t))^{1-\Delta N^a(t)} \\
&\quad\times \prod_{t\in(0,\tau]}\left(d\Lambda^d(t)\right)^{\Delta N^d(t)}\left(1-d\Lambda^d(t)\right)^{1-\Delta N^d(t)} \\
&= \prod_{t\in(0,\tau]}\left((\pi_t(1))^{\mathbb{1}\{A(t)=1\}}(1-\pi_t(1))^{\mathbb{1}\{A(t)=0\}}\right)^{\Delta N^a(t)} \\
&\quad\times \prod_{t\in(0,\tau]}(d\Lambda^a(t))^{\Delta N^a(t)}(1-d\Lambda^a(t))^{1-\Delta N^a(t)}\left(d\Lambda^d(t)\right)^{\Delta N^d(t)}\left(1-d\Lambda^d(t)\right)^{1-\Delta N^d(t)} \\
&= \prod_{t\in(0,\tau]} dG_t dQ_t
\end{aligned}$$

where

$$dG_t = \left((\pi_t(1))^{\mathbb{1}\{A(t)=1\}}(1-\pi_t(1))^{\mathbb{1}\{A(t)=0\}}\right)^{\Delta N^a(t)}$$

$$dQ_t = (d\Lambda^a(t))^{\Delta N^a(t)}(1-d\Lambda^a(t))^{1-\Delta N^a(t)}\left(d\Lambda^d(t)\right)^{\Delta N^d(t)}\left(1-d\Lambda^d(t)\right)^{1-\Delta N^d(dt)}$$

Let $dG^*_t = \left((1)^{\mathbb{1}\{A(t)=1\}}(0)^{\mathbb{1}\{A(t)=0\}}\right)^{\Delta N^a(t)} = \left((0)^{\mathbb{1}\{A(t)=0\}}\right)^{\Delta N^a(t)}$, corresponding to staying on treatment. Then define the interventional density as

$$dP_{Q,G^*}(O) = \prod_{t\in(0,\tau]} dG^*_t dQ_t$$

and their target estimand $\Psi^{\text{Rytgaard}}_t : \mathcal{M} \to \mathbb{R}_+$ as

$$\Psi^{\text{Rytgaard}}_\tau(P) = \mathbb{E}_{P_{Q,G^*}}[N^d_\tau] = \int_{\mathcal{O}} y \prod_{t\in(0,\tau]} dG^*_t dQ_t \tag{2}$$

We first need to define the integral in Equation 2. To get a fully rigorous result, consider Proposition 1 in Ryalen (2024) and Theorem 8.1.2 in Last & Brandt (1995).

First note that we have

$$\prod_{t\in(0,\tau]} dG^*_t dQ_t = \prod_{t\in\left(0,t_{(1)}\right]} dG^*_t dQ_t \prod_{t\in\left(t_{(1)},\tau\right]} dG^*_t dQ_t \mathbb{1}\{t_1 < t_2\}$$

Let $Y_\tau = \mathbb{1}\left\{T_{(1)} \leq \tau, D_{(1)} = 03\right\} + \mathbb{1}\left\{T_{(2)} \leq \tau\right\} := Y^{(1)}_\tau + Y^{(2)}_\tau$ be death at time τ. Then, note that

$$y^{(1)}_\tau(t_1, d_1) \prod_{t\in\left(0,t_{(1)}\right]} dG^*_t dQ_t \prod_{t\in\left(t_{(1)},\tau\right]} dG^*_t dQ_t \mathbb{1}\{t_1 < t_2\} = y^{(1)}_\tau \prod_{t\in\left(0,t_{(1)}\right]} dG^*_t dQ_t$$

The second product integral evaluates to 1 because death at event 1 implies that all intensities are 0 after the first event.

We find

$$
\begin{aligned}
\prod_{t\in(0,t_{(1)}]} dG_t^* dQ_t &= \left((0)^{\mathbb{1}\{d_1=02\}}\right)^{\mathbb{1}\{d_1\in\{01,02\}\}} (d\Lambda^a(t_1))^{\mathbb{1}\{d_1\in\{01,02\}\}} \left(d\Lambda^d(t_1)\right)^{\mathbb{1}\{d_1=03\}} \\
&\quad \times \prod_{t\in(0,t_{(1)})} \left(1-d\Lambda^d(t)\right)(1-d\Lambda^a(t)) \\
&= (d\Lambda^a(t_1))^{\mathbb{1}\{d_1\in\{01\}\}} (0dt_1)^{\mathbb{1}\{d_1\in\{02\}\}} \left(d\Lambda^d(t_1)\right)^{\mathbb{1}\{d_1=03\}} S_0(t_1-) \\
&= ((h^{01}(t_1)+h^{02}(t_1))dt_1)^{\mathbb{1}\{d_1\in\{01\}\}} (0dt_1)^{\mathbb{1}\{d_1\in\{02\}\}} (h^{03}(t_1)dt_1)^{\mathbb{1}\{d_1=03\}} S_0(t_1-) \\
&= S_0(t_1-)\mathbb{1}\{d_1=01\}(h^{01}(t_1)+h^{02}(t_1))dt_1 \\
&\quad + S_0(t_1-)\mathbb{1}\{d_1=03\}h^{03}(t_1)dt_1
\end{aligned}
$$

(compare with Equation (11) in Ryalen (2024)). In the second equality, we used that the counting processes do not jump at the same time with probability one to get $S(t_1) = \prod_{t\in(0,t_{(1)}]} (1 - d\Lambda^d(t))(1-d\Lambda^a(t))$. But multiplying with $y_\tau^{(1)}$, we find

$$
y_\tau^{(1)} \prod_{t\in(0,t_{(1)}]} dG_t^* dQ_t = y_\tau^{(1)} S(t_1-)\mathbb{1}\{d_1=03\}h^{03}(t_1)dt_1
$$

Therefore, we have

$$
\int y_\tau^{(1)} \prod_{t\in(0,t_{(1)}]} dG_t^* dQ_t = \int_0^\tau S(s-)h^{03}(s)ds
$$

Similarly, we may find

$$
\begin{aligned}
& y_\tau^{(2)} \prod_{t\in(0,t_{(1)}]} dG_t^* dQ_t \prod_{t\in(t_{(1)},t_{(2)}]} dG_t^* dQ_t \mathbb{1}\{t_1<t_2\} \\
&= y_\tau^{(2)} \mathbb{1}\{t_1<t_2\} S(t_1-)\mathbb{1}\{d_1=01\}(h^{01}(t_1)+h^{02}(t_1)) \\
&\quad \times S(t_2- \mid 01, t_1) h^{13}(t_1,t_2) dt_2 dt_1
\end{aligned}
$$

Thus the target estimand is

$$
\begin{aligned}
\Psi_\tau^{\text{Rytgaard}}(P) &= \int_0^\tau S_0(s-)h^{03}(s)ds \\
&\quad + \int_0^\tau S_0(s-)P_{13}(\tau \mid 01, s)(h^{01}(s)+h^{02}(s))ds
\end{aligned} \tag{3}
$$

2.2 Comparison of the approaches

We are now in a position, where we can readily compare the approaches in Rytgaard et al. (2022) and Ryalen (2024) by considering the difference between Equation 3 and Equation 1.

Suppose that $h^{02}(s) > 0$ and $h^{13}(s,w) > 0$ for Lebesgue almost all s, w. From this, we conclude that $\Psi_\tau^{\text{Rytgaard}}(P) = \Psi_\tau^{\text{Ryalen}}(P)$ if and only if $h^{02} \equiv 0$ a.e. if and only if $\pi_t(1) \equiv 1$ a.e. (with respect to the Lebesgue measure restricted to $[0,\tau]$). To see this, note that

$$
\begin{aligned}
\Psi_\tau^{\text{Ryalen}}(P) - \Psi_\tau^{\text{Rytgaard}}(P) = &\int_0^\tau \exp\left(-\sum_{j\neq 2}\int_0^s h^{0j}(u)du\right)\left(1-\exp\left(-\int_0^s h^{02}(u)du\right)\right)h^{03}(s)ds \\
&+ \int_0^\tau \exp\left(-\sum_{j\neq 2}\int_0^s h^{0j}(u)du\right)P_{13}(\tau \mid 01, s) \\
&\qquad\qquad \times \left(1-\exp\left(-\int_0^s h^{02}(u)du\right)\right)h^{01}(s)ds \\
&+ \int_0^\tau S_0(s-)P_{13}(\tau \mid 01, s)h^{02}(s)ds
\end{aligned} \tag{4}
$$

Since each term is non-negative, $\Psi_\tau^{\text{Rytgaard}}(P) = \Psi_\tau^{\text{Ryalen}}(P)$ implies that each term is equal to zero. Since each of the integrands are non-negative, we must have that the integrands are equal to zero (almost surely). By letting $m_{[0,\tau]}$ denote the Lebesgue measure on $[0,\tau]$, we have for the first term in Equation 4,

$$
\begin{aligned}
\exp\left(-\sum_{j\neq 2}\int_0^s h^{0j}(u)du\right)\left(1-\exp\left(-\int_0^s h^{02}(u)du\right)\right)h^{03}(s) = 0 \quad & m_{[0,\tau]} - \text{almost all } s \Leftrightarrow \\
\left(1-\exp\left(-\int_0^s h^{02}(u)du\right)\right)h^{03}(s) = 0 \quad & m_{[0,\tau]} - \text{almost all } s \Leftrightarrow \\
\left(1-\exp\left(-\int_0^s h^{02}(u)du\right)\right) = 0 \quad & m_{[0,\tau]} - \text{almost all } s \Leftrightarrow \\
h^{02}(s) = 0 \quad & m_{[0,\tau]} - \text{almost all } s
\end{aligned}
$$

with similar arguments for the second and third terms in Equation 4.

Bibliography

Jacod, J. (1975). Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales. *Zeitschrift Für Wahrscheinlichkeitstheorie Und Verwandte Gebiete*, *31*(3), 235–253.

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

Ryalen, P. (2024). *On the role of martingales in continuous-time causal inference.*

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