A counterfactual interpretation of target parameter in Rytgaard et al. (2022)

Let us consider a setting similar to the one of Ryalen (2024). It has been discussed in discrete time (Gill & Robins (2001)) that the g-formula is unique; however, as we shall see the g-formula in continuous time may not necessarily be uniquely defined. Specifically this may relate to conditional distributions in this setting not being uniquely defined. We will work with an intervention that specifies the treatment decisions but not the timing of treatment visits. We consider death as the outcome of interest and are interested in the probability of death, had had we followed the regime of always treating. To simplify, we work without right-censoring, no covariates, and compliance to treatment at time 0. Let $(\Omega, \mathcal{F}, P)$ be a probability space. and consider (N^y, N^a), where

- N^y is a counting process on $[0, T]$ for death.
- N^a is a random measure for treatment on $[0, T] \times \{1, 0\}$, where 1 denotes treatment and 0 no treatment.

We consider the filtration generated by (N^y, N^a) and denote it by $(\mathcal{F}_t)_{t \geq 0}$, i.e.,

$$\mathcal{F}_t := \sigma(N^y(\mathrm{d}s), N^a(\mathrm{d}s \times \{x\}) \mid s \in (0, t], x \in \{0, 1\}).$$

Further, we assume that

- N^y and $N^a(\{(0, t] \times \{1, 0\}\})$ do not jump at the same time.
- $M^y = N^y - \Lambda^y$ denotes their P-$\mathcal{F}_t$ (local) martingale, where Λ^y is the P-$\mathcal{F}_t$-compensator of N^y.
- $M^a(\mathrm{d}t \times \{x\}) = N^a(\mathrm{d}t \times \{x\}) - (\pi_t)^{\mathbb{1}\{x=1\}}(1 - \pi_t)^{\mathbb{1}\{x=0\}} \Lambda^a(\mathrm{d}t)$ is the P-$\mathcal{F}_t$ (local) martingale for $x \in \{1, 0\}$, where π_t is the $\mathcal{F}_t$-predictable probability of treatment at time t (mark probability) and $\Lambda^a(\mathrm{d}t)$ is the total P-$\mathcal{F}_t$-compensator of $N^a(\mathrm{d}t \times \mathrm{d}x)$.

For this treatment regime, we see that

$$\tau^A = \inf\{t \geq 0 \mid N^a((0, t] \times \{0\}) > 0\}.$$

We are interested in the counterfactual mean outcome $\mathbb{E}_P\left[\tilde{Y}_t\right]$, where $\left(\tilde{Y}_t\right)_{t \geq 0}$ is the counterfactual outcome process of $Y := N^y$ under the intervention that sets treatment to 1 at all visitation times. This process is assumed to satisfy the definition of counterfactual outcome processes of Ryalen (2024) with their Example 4. Note the different exchangeability condition compared to Ryalen (2024), as Ryalen (2024) expresses exchangeability through the counting process $\mathbb{1}\{\tau^A \leq \cdot\}$. Let $\left(T_{(k)}, \Delta_{(k)}, A\left(T_{(k)}\right)\right)$ denote the ordered event times, event types, and treatment decisions at event k. Note that Equation 1 is the same likelihood ratio as in Rytgaard et al. (2022). We also impose the assumption that $N_t := N^y_t + N^a(\{(0, t] \times \{1, 0\}\})$ does not explode; we also assume that we work with a version of the compensator such that $\Lambda(\{t\} \times \{y, a\} \times \{1, 0\}) < \infty$ for all $t > 0$. We may generally also work with a compensator Λ that fulfills conditions (10.1.11)-(10.1.13) of Last & Brandt (1995). Let $\pi^*_{T_{(k)}}\left(\mathcal{F}_{T_{(k-1)}}\right)$ denote the interventional probability, which in this case we take to be 1. In this case,

$$\pi_t = \sum_k \mathbb{1}\left\{T_{(k-1)} < t < T_{(k)}\right\} \pi_{T_{(k)}}\left(\mathcal{F}_{T_{(k-1)}}\right)$$

$$\pi^*_t = \sum_k \mathbb{1}\left\{T_{(k-1)} < t < T_{(k)}\right\} \pi^*_{T_{(k)}}\left(\mathcal{F}_{T_{(k-1)}}\right) = 1.$$

Let $N^{ax}(\mathrm{d}t) := N^a(\mathrm{d}t \times \{x\})$ for $x \in \{1, 0\}$.

NOTES:

- Does the exchangeability condition simplify in the case of $\mathcal{n}^a$ predictable in P-$\mathcal{F}_t$ as specified in Ryalen (2024); as noted in their article the two likelihood ratios turn out two be the same in the case of orthogonal martingales.

  Suppose that $\mathcal{n}^a$ is predictable so that $N^{a1}(\mathrm{d}t)$ is predictable in that case the first exchangeability condition is trivial; Pål's condition only grants exchangeability for $N^{a1}(t \wedge \tau^A)$ is predictable; I think that this is the sufficient for the argument to go through.

- Positivity holds for example if π_t is bounded away from 0 and 1 and N_t has bounded number of jumps in the study period.

Theorem 0.1: If *all* of the following conditions hold:

- **Consistency**: $\tilde{Y}_\cdot \mathbb{1}\{\tau^a > \cdot\} = Y_\cdot \mathbb{1}\{\tau^a > \cdot\} \quad P - \text{a.s.}$
- **Exchangeability**: Define $\mathcal{H}_t := \mathcal{F}_t \vee \sigma(\tilde{Y})$. The P-$\mathcal{F}_t$ compensator for N^a is also the P-$\mathcal{H}_t$ compensator.
- **Positivity**:

$$W(t) := \prod_{j=1}^{N_t} \left(\left(\frac{\pi^*_{T_{(j)}}\left(\mathcal{F}_{T_{(j-1)}}\right)}{\pi_{T_{(j)}}\left(\mathcal{F}_{T_{(j-1)}}\right)} \right)^{\mathbb{1}\{A(T_{(k)})=1\}} \left(\frac{1-\pi^*_{T_{(j)}}\left(\mathcal{F}_{T_{(j-1)}}\right)}{1-\pi_{T_{(j)}}\left(\mathcal{F}_{T_{(j-1)}}\right)} \right)^{\mathbb{1}\{A(T_{(k)})=0\}} \right)^{\mathbb{1}\{\Delta_{(j)}=a\}} \tag{1}$$

  is uniformly integrable.

  Furthermore, assume that $K(s) = \left(\frac{\pi^*_s}{\pi_s} - 1\right) M^{a1}(\mathrm{d}s) + \left(\frac{1-\pi^*_s}{1-\pi_s} - 1\right) M^{a0}(\mathrm{d}s)$ is a P-$\mathcal{F}_t$-martingale and that K is a process of **locally integrable variation**, meaning that $\mathbb{E}_P\left[\int_0^t |dK(s)|\right] < \infty$ for all $t > 0$.

Then,

$$\mathbb{E}_P\left[\tilde{Y}_t\right] = \mathbb{E}_P[Y_t W(t)]$$

and $W(t) = \mathcal{E}(K)_t$ is a uniformly integrable P-$\mathcal{F}_t$-martingale, where $\mathcal{E}$ denotes the Doléans-Dade exponential (Protter (2005)).

Proof: We shall use that the likelihood ratio solves a specific stochastic differential equation. To this end, note that

$$\begin{aligned}
&W(t) \\
&= \prod_{s \le t} \left(1 + \left(\frac{\pi^*_s}{\pi_s} - 1\right) N^{a1}(\mathrm{d}s) + \left(\frac{1-\pi^*_s}{1-\pi_s} - 1\right) N^{a0}(\mathrm{d}s)\right) \\
&= \prod_{s \le t} \left(1 + \left(\frac{\pi^*_s}{\pi_s} - 1\right) N^{a1}(\mathrm{d}s) + \left(\frac{1-\pi^*_s}{1-\pi_s} - 1\right) N^{a0}(\mathrm{d}s) - (\pi^*_s - \pi_s)\Lambda^a(ds) - (\pi_s - \pi^*_s)\Lambda^a(ds)\right) \\
&= \prod_{s \le t} \left(1 + \left(\frac{\pi^*_s}{\pi_s} - 1\right) N^{a1}(\mathrm{d}s) + \left(\frac{1-\pi^*_s}{1-\pi_s} - 1\right) N^{a0}(\mathrm{d}s) - \frac{\pi^*_s - \pi_s}{\pi_s}\Lambda^{a1}(ds) - \frac{\pi_s - \pi^*_s}{1-\pi_s}\Lambda^{a0}(ds)\right) \\
&= \prod_{s \le t} \left(1 + \left(\frac{\pi^*_s}{\pi_s} - 1\right) M^{a1}(\mathrm{d}s) + \left(\frac{1-\pi^*_s}{1-\pi_s} - 1\right) M^{a0}(\mathrm{d}s)\right).
\end{aligned}$$

Thus, by properties of the product integral (e.g., Theorem II.6.1 of Andersen et al. (1993)),

$$W(t) = 1 + \int_0^t W(s-)\left(\frac{\pi_s^*}{\pi_s} - 1\right) M^{a1}(\mathrm{d}s) + \int_0^t W(s-)\left(\frac{1-\pi_s^*}{1-\pi_s} - 1\right) M^{a0}(\mathrm{d}s). \quad (2)$$

We have that

$$\zeta_t := \int_0^t W(s-)\left(\frac{\pi_s^*}{\pi_s} - 1\right) M^{a1}(\mathrm{d}s) + \int_0^t W(s-)\left(\frac{1-\pi_s^*}{1-\pi_s} - 1\right) M^{a0}(\mathrm{d}s)$$

is a zero mean P-$\mathcal{H}_t$-martingale by positivity. From this, we see that $\int_0^t \tilde{Y}_t \zeta(\mathrm{d}s)$ is also a uniformly integrable P-$\mathcal{H}_t$-martingale, which implies by Theorem 2.1.42 of Last & Brandt (1995) $\int_0^t \tilde{Y}_t \zeta(\mathrm{d}s)$ is a zero mean P-$\mathcal{H}_t$-martingale. This implies that

$$\mathbb{E}_P[Y_t W(t)] \overset{*}{=} \mathbb{E}_P\left[\tilde{Y}_t W(t)\right] = \mathbb{E}_P\left[\tilde{Y}_t\right] + \mathbb{E}_P\left[\int_0^t \tilde{Y}_t \zeta(\mathrm{d}s)\right] = \mathbb{E}_P\left[\tilde{Y}_t\right],$$

where in $*$ we used consistency by noting that $W(t) \neq 0$ if and only if $\tau^a > t$. □

It is also natural to ask oneself: how does our exchangeability condition relate to the one of Ryalen (2024)? We present a result in this direction.

Theorem 0.2: Let $\mathbb{N}_t^a = \mathbb{1}\{\tau^A \leq t\}$. The exchangeability condition of Theorem 0.1 implies the one of Ryalen (2024), e.g., $\mathbb{L}_t := \Lambda_t^a$ is both the P-$\mathcal{F}_{t \wedge \tau^A}$ compensator and the P-$\mathcal{H}_{t \wedge \tau^A}$ compensator of $\mathbb{N}_t^a$.

Proof: Consider some localizing sequence S_n for M^{a0}. We note that $\mathbb{N}_t^a = N^a\left((0, t \wedge \tau^A], \{0\}\right)$. Apply optional sampling to $M^{a0}_{\cdot \wedge S_n}$ at $S := \tau^A \wedge S_n \wedge s$ and $T := t \wedge S_n \wedge \tau^A$ to see that

$$\mathbb{E}_P\left[M^{a0}_{t \wedge S_n \wedge \tau^A} \mid \mathcal{F}_{s \wedge \tau^A}\right] = M^{a0}_{s \wedge S_n \wedge \tau^A} \quad P - \text{a.s.}$$

If exchangeability for Λ^{a0} holds (given in Theorem 0.1), then the same argument applies with $\mathcal{H}_t$ instead of $\mathcal{F}_t$, so that

$$\mathbb{E}_P\left[M^{a0}_{t \wedge S_n \wedge \tau^A} \mid \mathcal{H}_{s \wedge \tau^A}\right] = M^{a0}_{s \wedge S_n \wedge \tau^A} \quad P - \text{a.s.}$$

This is the desired result. □

We can also ask ourselves: Is the exchangeability criterion in Theorem 0.1 close in interpretation to the statement of Rytgaard et al. (2022)? In Rytgaard et al. (2022), the statement is:

$$\left(\tilde{Y}_t\right)_{t \in [0,T]} \perp A(T_k^a) \mid \mathcal{F}_{T_k^a}^-, \qquad (')$$

for all k, where T_k^a are the ordered treatment event times, where $\mathcal{F}_T^-$ is defined on p. 62 of Last & Brandt (1995). This σ-algebra contains all the information that occurs strictly before time T.

In this case, we can express our exchangeability condition via something that is very similar to this statement:

$$\left(\tilde{Y}_t\right)_{t \in [0,T]} \perp A\left(T_{(k)}\right) \mid \Delta_{(k)} = a, T_{(k)}, \mathcal{F}_{T_{(k-1)}}, \qquad (*)$$

for all k. The statements appear similar, but are generally not the same, since T is not generally $\mathcal{F}_T^-$ measurable. If S and ΔN_S^a are $\mathcal{F}_S^-$ measurable, then the two statements should be the same. However, if S is predictable, then $S \in \sigma(\mathcal{F}_S^-)$ (Theorem 2.2.19 of Last & Brandt (1995)), If N_t^a-predictable, it should also be the case that $\Delta N_S^a \in \sigma(\mathcal{F}_S^-)$? Yes, if N_t^a is predictable, then $N_t^a \in \sigma(\mathcal{F}_{t-})$ (Theorem 2.2.9 of Last & Brandt (1995)); therefore $\Delta N_S^a \in \sigma(\mathcal{F}_S^-)$. However, due to the classical fact that conditional independence and independence never imply each other, the two statements are not equivalent and are generally different.

To have exchangeability, we also need that the compensator for $N^a = N^a(\cdot \times \{0,1\})$ is the same under $\mathcal{F}_t$ and $\mathcal{H}_t$, i.e., that

- $\Lambda^a(\mathrm{d}t)$ is the P-$\mathcal{F}_t$-compensator and the P-$\mathcal{H}_t$-compensator of $N^a(\mathrm{d}t \times \{0,1\})$. (**)

A slight strengthening of (*) is that

- The Radon-Nikodym derivative of $\Lambda^{a1}(\mathrm{d}t)$ with respect to $\Lambda^a(\mathrm{d}t)$ is the same for $\mathcal{F}_t$ and $\mathcal{H}_t$. (*!)

This is because there is a version of π_t such that

$$\pi_t' = \sum_k \mathbb{1}\left\{T_{(k-1)} < t < T_{(k)}\right\}\pi_{T_{(k)}}\left(\mathcal{H}_{T_{(k-1)}}\right)$$

We then have the following result.

Theorem 0.3: The conditions (*!) and (**) hold if and only if the exchangeability condition of Theorem 0.1 holds.

Proof: To see that it is sufficient, note that (*) and (**) imply that

$$N^a(\mathrm{d}t \times \{x\}) - (\pi_t)^{\mathbb{1}\{x=1\}}(1-\pi_t)^{\mathbb{1}\{x=0\}}\Lambda^a(\mathrm{d}t) \tag{3}$$

However, it must also be a P-$\mathcal{F}_t$-local martingale; to see this, let S_n be a localizing sequence for Equation 3 and consider $0 \leq s < t$. Then,

$$\begin{aligned}\mathbb{E}_P\left[M^{ax}_{t\wedge S_n} \mid \mathcal{F}_s\right] &= \mathbb{E}_P\left[\mathbb{E}_P\left[M^{ax}_{t\wedge S_n} \mid \mathcal{H}_s\right] \mid \mathcal{F}_s\right] \\ &= \mathbb{E}_P\left[M^{ax}_{s\wedge S_n} \mid \mathcal{F}_s\right] \overset{!}{=} M^{ax}_{s\wedge S_n} \quad P-\text{a.s.}.\end{aligned}$$

In !, we used that $N^{ax}_{s\wedge S_n}$ is (trivially) $\mathcal{F}_s$ measurable; moreover, by (*), π_s is $\mathcal{F}_s$ measurable; finally, by (**), $\Lambda^a(\mathrm{d}t)$ is also the P-$\mathcal{F}_t$-compensator of $N^a(\mathrm{d}t \times \{0,1\})$; and hence it is predictable with respect to this filtration because it is $\mathcal{F}_{t-}$ measurable (Theorem 2.2.6 of Last & Brandt (1995)). Conversely, to see that it is necessary, we have directly (**); however this is precisely what we needed to show (*!). □

One may ask oneself if positivity holds in Ryalen (2024); under what assumptions does positivity in Theorem 0.1 hold?

Ryalen (2024) introduces the weight

$$\widetilde{W} = \frac{\mathcal{E}(-\mathbb{N}^a)}{\mathcal{E}(-\mathbb{L}^a)}$$

Similarly, we can introduce the weight

$$W = \mathcal{E}(K),$$

where

$$K_t = \int_0^t \left(\frac{\pi_s^*}{\pi_s} - 1\right) N^{a1}(\mathrm{d}s) + \left(\frac{1-\pi_s^*}{1-\pi_s} - 1\right) N^{a0}(\mathrm{d}s).$$

Can we find a process φ such that $\mathcal{E}(K) = \frac{\mathcal{E}(-\mathbb{N}^a)}{\mathcal{E}(-\mathbb{L}^a)}\mathcal{E}(\varphi)$? To this end, note that

$$\begin{aligned}
\mathcal{E}(\varphi) &= \frac{\mathcal{E}(K)\mathcal{E}(-\mathbb{L}^a)}{\mathcal{E}(-\mathbb{N}^a)} \\
&= \frac{\mathcal{E}(-\mathbb{N}^a)\mathcal{E}(K)\mathcal{E}(-\mathbb{L}^a)}{\mathcal{E}(-\mathbb{N}^a)} \\
&= \mathcal{E}(K)\mathcal{E}(-\mathbb{L}^a) \\
&= \mathcal{E}(K - \mathbb{L}^a - [K, \mathbb{L}^a]).
\end{aligned}$$

where we use that

$$\mathcal{E}(K) = \mathcal{E}(K)\mathbb{1}\{\tau^a > \cdot\} = \mathcal{E}(K)\mathcal{E}(-\mathbb{N}^a)$$

taking $\frac{0}{0} = 1$. Note that

$$\begin{aligned}
[K, \mathbb{L}^a]_t &= \int_0^t \Delta\mathbb{L}_s^a \left(\frac{\pi_s^*}{\pi_s} - 1\right) dN_s^{a1} \\
&\quad + \int_0^t \Delta\mathbb{L}_s^a \left(\frac{1-\pi_s^*}{1-\pi_s} - 1\right) dN_s^{a0} \\
&\overset{*}{=} \int_0^{t\wedge\tau^a} \pi_s \Delta\Lambda^a(s) \left(\frac{1-\pi_s^*}{1-\pi_s} - 1\right) dN_s^{a0} \\
&\quad + \int_0^{t\wedge\tau^a} \pi_s \Delta\Lambda^a(s) \left(\frac{\pi_s^*}{\pi_s} - 1\right) dN_s^{a1}, \\
&= \int_0^{t\wedge\tau^a} \Delta\Lambda^a(s)(\pi_s - \pi_s^*)\frac{\pi_s}{1-\pi_s} dN_s^{a0} \\
&\quad + \int_0^{t\wedge\tau^a} \Delta\Lambda^a(s)(\pi_s^* - \pi_s) dN_s^{a1},
\end{aligned}$$

In the absolutely continuous case, $[K, \mathbb{L}^a]_t = 0$ as $\Delta\Lambda_t^a = 0$ for all $t > 0$. Also note that in $(*)$, we apply the corollary on p. 10 of Protter (2005), e.g., stopped martingales are martingales and $\mathbb{N}_t^a = N^{a0}(t \wedge \tau^a)$.

However, it is also the case that

$$\mathcal{E}(\varphi) = \mathcal{E}(K - \mathbb{L}^a + \mathbb{N}^a - [K, \mathbb{L}^a])$$

because $\mathbb{N}^a \equiv 0$ whenever $\mathcal{E}(K) \neq 0$ and $\frac{\mathcal{E}(-\mathbb{N}^a)}{\mathcal{E}(-\mathbb{L}^a)} \neq 0$. Also

$$\begin{aligned}
K_t - \mathbb{L}_t^a + \mathbb{N}_t^a &= \int_0^{t\wedge\tau^a} \left(\frac{\pi_s^*}{\pi_s} - 1\right) M^{a1}(\mathrm{d}s) + \left(\frac{1-\pi_s^*}{1-\pi_s} - 1\right) M^{a0}(\mathrm{d}s) \\
&\quad + \int_0^{t\wedge\tau^a} \frac{1}{1-(1-\pi_s^*)\Delta\Lambda_s^a}(1-\pi_s^*)(N^a(\mathrm{d}s) - \Lambda^a(\mathrm{d}s)) \\
&\quad + \int_0^{t\wedge\tau^a} \frac{1}{1-\pi_s^*\Delta\Lambda_s^a}\pi_s^*(N^a(\mathrm{d}s) - \Lambda^a(\mathrm{d}s))
\end{aligned}$$

The last $\mathbb{N}^a$ can be ignored. When $\pi^*_s = 1$ and Λ^a_s absolutely continuous, then

$$K_t - \mathbb{L}^a_t + \mathbb{N}^a_t = \int_0^{t \wedge \tau^a} \left(\frac{1}{\pi_s} - 1 \right) M^{a1}(\mathrm{d}s)$$

and $[K, \mathbb{L}^a]_t = 0$.

Bibliography

Andersen, P. K., Borgan, Ø., Gill, R. D., & Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer US. https://doi.org/10.1007/978-1-4612-4348-9

Gill, R. D., & Robins, J. M. (2001). Causal Inference for Complex Longitudinal Data: The Continuous Case. *The Annals of Statistics*, *29*(6), 1785–1811. https://doi.org/10.1214/aos/1015345962

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

Protter, P. E. (2005). *Stochastic differential equations*. Springer.

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

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