1) A causal interpretation in terms of potential outcomes of target parameter in Rytgaard et al. (2022)

We consider a setting similar to the one of Ryalen (2024) and Rytgaard et al. (2022). As in Rytgaard et al. (2022), we consider some measure P on a probability space $(\Omega, \mathcal{F}, P)$. We consider a setting in which we observe a multivariate random measure $N = \left(N^y, N^a, N^\ell\right)$ which is defined on $(\Omega, \mathcal{F})$, where any two of the components do not jump at the same time. These processes are observed in $[0, T]$ for some $T > 0$. Here, N^y denotes an outcome process of interest Y (e.g., death), random measure N^a on $[0, T] \times \mathcal{A}$ for treatment A, where $\mathcal{A}$ is a measurable space;x N^ℓ denotes a random measure for covariates L on $[0, T] \times \mathcal{L}$, where $\mathcal{A}$ and $\mathcal{L}$ are measurable spaces; for instance finite subsets of $\mathbb{R}$ and $\mathbb{R}^d$. Such a random measure gives rise to a filtration $\left(\mathcal{F}_t\right)_{t \geq 0}$, where

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

Further, we make the assumption of no explosion of N.

We concern ourselves with the hypothetical question if the treatment process N^a had been intervened upon such that treatment was given according to some treatment regime g^*. We will work with an intervention that specifies the treatment decisions but does not change timing of treatment visits. What this means precisely will be made clear below. We are interested in the outcome process Y under this intervention, which we denote by $\tilde{Y}$. Importantly, the intervention is defined as a static/dynamic intervention

$$N^{g^*}(\mathrm{d}t \times \mathrm{d}x) = \pi_t^*(\mathrm{d}x) N^a(\mathrm{d}t \times \mathcal{A})$$

where $\pi_t^*(\mathrm{d}x)$ is some kernel that specifies the treatment decision deterministically at time t in the sense that there are functions $g_k^* : \mathbb{H}_{k-1} \times [0, T] \to \mathcal{A}$ such that

$$\pi_t^*(\mathrm{d}x) = \sum_k \mathbb{1}\big\{T_{(k-1)} < t \leq T_{(k)}\big\} \delta_{g_k^*\left(\mathcal{F}_{T_{(k-1)}}, T_{(k)}\right)}(\mathrm{d}x),$$

i.e., g_k^* is $\mathcal{F}_{T_{(k-1)}} \otimes \mathcal{B}([0, T])$-measurable This means that, critically, $N^{g^*}(\mathrm{d}t \times \mathrm{d}x)$ is also a random measure. Note that N^{g^*} has the compensator

$$\mathscr{L}(N)(\mathrm{d}t \times \mathrm{d}x) = \pi_t^*(\mathrm{d}x) \underbrace{\Lambda^a(\mathrm{d}t \times \mathcal{A})}_{=:\Lambda^a(\mathrm{d}t)},$$

where $\Lambda^a(\mathrm{d}t)$ is the P-$\mathcal{F}_t$-compensator of $N^a(\mathrm{d}t \times \mathcal{A})$. However, N^a generally has the compensator $\Lambda^a(\mathrm{d}t \times \mathrm{d}x) = \pi_t(\mathrm{d}x) \Lambda^a(\mathrm{d}t)$. Now define the time to deviation from the treatment regime as

$$\tau^{g^*} = \inf\{t \geq 0 \mid N^a((0, t] \times \{x\}) \neq N^{g^*}((0, t] \times \{x\}) \text{ for some } x \in \mathcal{A}\},$$

when $\mathcal{A} = \{a_1, ..., a_k\}$ consists of a finite set of treatment options.

Definition 1.1: A multivariate random measure $\tilde{N} = \left(\tilde{N}^y, \tilde{N}^a, \tilde{N}^\ell\right)$ is a **counterfactual random measure** under the intervention g^* if it satisfies the following conditions. Let $\tilde{\mathcal{F}}_t := \sigma\big(\tilde{N}^y(\mathrm{d}s), \tilde{N}^a(\mathrm{d}s \times \{x\}), \tilde{N}^\ell(\mathrm{d}s \times \{y\}) \mid s \in (0, t], x \in \mathcal{A}, y \in \mathcal{L}\big)$. Let $\mathscr{L}$ denote the P-$\mathcal{F}_t$-canonical compensator of N^{g^*} and let Λ denote the P-$\mathcal{F}_t$-compensator of N.
1. $\tilde{N}^a$ has compensator $\mathscr{L}\left(\tilde{N}\right)$ with respect to $\tilde{\mathcal{F}}_t$.
2. $\tilde{N}^x$, has the same compensator $\Lambda\left(\tilde{N}\right)$ with respect to $\tilde{\mathcal{F}}_t$ for $x \in \{y, \ell\}$.

The main outcome of interest is the counterfactual outcome process $\tilde{Y} := \widetilde{N}^y$; and we wish to identify $\mathbb{E}_P\big[\tilde{Y}_t\big]$.

Let $\big(T_{(k)}, \Delta_{(k)}, A\big(T_{(k)}\big)\big)$ 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). As a shorthand, we let $N^{a,x}(\mathrm{d}t) := N^a(\mathrm{d}t \times \{x\})$ for $x \in \mathcal{A}$.

> **Theorem 1.1**: If *all* of the following conditions hold:
> - **Consistency**: $\tilde{Y}\mathbb{1}\{\tau^{g^*} > \cdot\} = Y\mathbb{1}\{\tau^{g^*} > \cdot\}$ $P-$a.s.
> - **Exchangeability**: Define $\mathcal{H}_t := \mathcal{F}_t \vee \sigma\big(\tilde{Y}\big)$. The P-$\mathcal{F}_t$ compensator for N^a is also the P-$\mathcal{H}_t$ compensator.
> - **Positivity**:
>
> $$W(t) := \prod_{j=1}^{N_t}\left(\prod_{i=1}^{k}\left(\frac{\pi^*_{T_{(j)}}\big(\{a_i\}; \mathcal{F}_{T_{(j-1)}}\big)}{\pi_{T_{(j)}}\big(\{a_i\}; \mathcal{F}_{T_{(j-1)}}\big)}\right)^{\mathbb{1}\{A(T_{(k)})=a_i\}}\right)^{\mathbb{1}\{\Delta_{(j)}=a\}} \tag{1}$$
>
> is uniformly integrable.
>
> Furthermore, assume that $K(t) = \int_0^t \sum_{j=1}^{k}\left(\frac{\pi^*_s(\{a_j\})}{\pi_s(\{a_j\})} - 1\right) M^{a,a_j}(\mathrm{d}s)$ is a P-$\mathcal{F}_t$-martingale and that K is a process of **locally integrable variation**, meaning that $\mathbb{E}_P\big[\int_0^t |dK(s)|\big] < \infty$ for all $t > 0$.
>
> Then,
>
> $$\mathbb{E}_P\big[\tilde{Y}_t\big] = \mathbb{E}_P[Y_t W(t)]$$
>
> 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

$$W(t) = \mathcal{E}\left(\sum_{j=1}^{k}\left(\frac{\pi^*_s(\{a_j\})}{\pi_s(\{a_j\})} - 1\right)N^{a,a_j}(\mathrm{d}s)\right)_t$$

$$\overset{(*)}{=} \mathcal{E}\left(\sum_{j=1}^{k}\left(\frac{\pi^*_s(\{a_j\})}{\pi_s(\{a_j\})} - 1\right)N^{a,a_j}(\mathrm{d}s) - \sum_{j=1}^{k}(\pi^*_s(\{a_j\}) - \pi_s(\{a_j\}))\Lambda^a(\mathrm{d}s)\right)_t$$

$$= \mathcal{E}\left(\sum_{j=1}^{k}\left(\frac{\pi^*_s(\{a_j\})}{\pi_s(\{a_j\})} - 1\right)N^{a,a_j}(\mathrm{d}s) - \sum_{j=1}^{k}\left(\frac{\pi^*_s(\{a_j\})}{\pi_s(\{a_j\})} - 1\right)\Lambda^{a,a_j}(\mathrm{d}s)\right)_t$$

$$= \mathcal{E}\left(\sum_{j=1}^{k}\left(\frac{\pi^*_s(\{a_j\})}{\pi_s(\{a_j\})} - 1\right)M^{a,a_j}(\mathrm{d}s)\right)_t.$$

In $(*)$, we use that $\sum_{j=1}^{k}\pi_s(\{a_j\}) = \sum_{j=1}^{k}\pi^*_s(\{a_j\}) = 1$.

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

$$W(t) = 1 + \int_0^t W(s-) \sum_{j=1}^k \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1 \right) M^{a,a_j}(\mathrm{d}s). \tag{2}$$

We have that

$$\zeta_t := \int_0^t W(s-) \sum_{j=1}^k \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1 \right) M^{a,a_j}(\mathrm{d}s)$$

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 by Theorem 2.1.42 of Last & Brandt (1995). Hence, $\left(\tilde{Y}_t W(s) \right)_{s \in [0,T]}$ is a P-$\mathcal{H}_t$-martingale. This implies that

$$\mathbb{E}_P[Y_t W(t)] \overset{(**)}{=} \mathbb{E}_P\left[\tilde{Y}_t W(t) \right] = \mathbb{E}_P\left[\tilde{Y}_t W(0) \right] = \mathbb{E}_P\left[\tilde{Y}_t \right]$$

where in $(**)$ we used consistency by noting that $W(t) \neq 0$ if and only if $\tau^{g^*} > t$. $\qquad\square$

It is also natural to ask oneself: how does our conditions relate to the ones of Ryalen (2024)? The condition of consistency is the same. However, the exchangeability condition and the positivity condition are different in general. We present slightly strengthened versions of the conditions as these are easier to compare. Let $\mathbb{N}_t^a = \mathbb{1}\{\tau^{g^*} \leq t\}$ and let $\mathbb{L}_t$ denote its P-$\mathcal{F}_t$-compensator.

- **Exchangeability**: Define $\mathcal{H}_t := \mathcal{F}_t \vee \sigma\left(\tilde{Y} \right)$. The P-$\mathcal{F}_t$ compensator for $\mathbb{N}^a$ is also the P-$\mathcal{H}_t$ compensator.

- **Positivity**:

$$\widetilde{W}(t) := \frac{(\mathcal{E}(-\mathbb{N}^a))_t}{(\mathcal{E}(-\mathbb{L}^a))_t} = \mathcal{E}\left(\widetilde{K} \right)_t$$

is uniformly integrable, where $\widetilde{K}_t = \int_0^t \frac{1}{1-\Delta\mathbb{L}_s^a}(\mathbb{N}^a(\mathrm{d}s) - \mathbb{L}^a(\mathrm{d}s))$. Furthermore, $\widetilde{K}$ is a process of **locally integrable variation** and a P-$\mathcal{F}_t$-martingale.

1.a) Comparison with the exchangeability condition of Ryalen (2024)

Theorem 1.1.1: Let $\mathbb{N}_t^a = \mathbb{1}\{\tau^{g^*} \leq t\}$. The exchangeability condition of Theorem 1.1 implies the one of Ryalen (2024), e.g., $\mathbb{L}_t$ is both the P-$\mathcal{F}_t$ compensator and the P-$\mathcal{H}_t$ compensator of $\mathbb{N}_t^a$.

Proof: First, let

$$N^{a,*}(\mathrm{d}t) = \sum_k \sum_j \mathbb{1}\left\{ T_{(k-1)} < t \leq T_{(k)} \right\} \mathbb{1}\left\{ j \neq g_k^*\left(\mathcal{F}_{T_{(k-1)}}, t \right) \right\} N^{a,a_j}(\mathrm{d}t),$$

which is the counting measure for the number of deviations from the treatment regime g^* and consider

$$\mathbb{1}\left(t \leq T_{(K)} \right) N^{a,*}(\mathrm{d}t) = \sum_{k=1}^K \sum_j \mathbb{1}\left\{ T_{(k-1)} < t \leq T_{(k)} \right\} \mathbb{1}\left\{ j \neq g_k^*\left(\mathcal{F}_{T_{(k-1)}}, t \right) \right\} N^{a,a_j}(\mathrm{d}t),$$

for some K. Note that $\left(\mathbb{1}\{\tau^{g^*} \leq t\} N^{a,*}(\mathrm{d}t)\right)^p = \mathbb{1}\{\tau^{g^*} \leq t\} \Lambda^{a,*}(\mathrm{d}t)$, where p denotes the predictable projection (Last & Brandt (1995)), by definition of the predictable projection. On the other hand,

$$
\begin{aligned}
&\left(\sum_{k=1}^{K} \sum_{j} \mathbb{1}\left\{T_{(k-1)} < t \leq T_{(k)}\right\} \mathbb{1}\left\{j \neq g_k^*\left(\mathcal{F}_{T_{(k-1)}}, t\right)\right\} N^{a,a_j}(\mathrm{d}t)\right)^p \\
&= \sum_{k=1}^{K} \sum_{j} \left(\mathbb{1}\left\{T_{(k-1)} < t \leq T_{(k)}\right\} \mathbb{1}\left\{j \neq g_k^*\left(\mathcal{F}_{T_{(k-1)}}, t\right)\right\} N^{a,a_j}(\mathrm{d}t)\right)^p \\
&\stackrel{(*)}{=} \sum_{k=1}^{K} \sum_{j} \mathbb{1}\left\{T_{(k-1)} < t \leq T_{(k)}\right\} \mathbb{1}\left\{j \neq g_k^*\left(\mathcal{F}_{T_{(k-1)}}, t\right)\right\} (N^{a,a_j}(\mathrm{d}t))^p \\
&= \sum_{k=1}^{K} \sum_{j} \mathbb{1}\left\{T_{(k-1)} < t \leq T_{(k)}\right\} \mathbb{1}\left\{j \neq g_k^*\left(\mathcal{F}_{T_{(k-1)}}, t\right)\right\} \Lambda^{a,a_j}(\mathrm{d}t),
\end{aligned}
\tag{3}
$$

where we use in $(*)$ that $\mathbb{1}\left\{T_{(k-1)} < t \leq T_{(k)}\right\} \mathbb{1}\left\{j \neq g_k^*\left(\mathcal{F}_{T_{(k-1)}}, t\right)\right\}$ is predictable (Theorem 2.2.22 of Last & Brandt (1995)). Letting $K \to \infty$ and using the assumption of no explosion shows that the P-$\mathcal{F}_t$-compensator of $N^{a,*}(\mathrm{d}t)$ is given by the limit as $K \to \infty$ of the right-hand side of Equation 3. Since $\mathbb{N}_t^a = N^{a,*}\left((0, t \wedge \tau^{g^*}]\right)$, the corollary on p. 10 of Protter (2005), implies that the P-$\mathcal{F}_t$-compensator of $\mathbb{N}_t^a$ is

$$
\mathbb{L}_t^a = \sum_{k} \sum_{j} \mathbb{1}\left\{T_{(k-1)} < t \leq T_{(k)}\right\} \mathbb{1}\left\{j \neq g_k^*\left(\mathcal{F}_{T_{(k-1)}}, t\right)\right\} \Lambda^{a,a_j}\left(t \wedge \tau^{g^*}\right).
$$

by using the limit of the right-hand side of Equation 3. If exchangeability holds (given in Theorem 1.1), then the same argument applies with $\mathcal{H}_t$ instead of $\mathcal{F}_t$. This is the desired result. $\qquad \square$

1.b) Comparisons of the positivity assumptions in Ryalen (2024)

One may ask oneself if positivity holds in Ryalen (2024); under what assumptions does positivity in Theorem 1.1 hold? In general, however, it would appear that the two positivity conditions are different and neither implies the other.

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

Theorem 1.2.1: φ is given by

$$\varphi_t = K_t - \mathbb{L}_t^a + \mathbb{N}_t^a - [K, \mathbb{L}^a]_t,$$

where $[\cdot, \cdot]$ denotes the quadratic covariation process (Protter (2005)), where

$$[K, \mathbb{L}^a]_t = \int_0^{t \wedge \tau^{g^*}} \sum_v \mathbb{1}\{T_{(v-1)} < s \leq T_{(v)}\} \sum_{i \neq g_v^*\left(\mathcal{F}_{T_{(v-1)}}, T_{(v)}\right)} \pi_s(\{a_j\}) \Delta \Lambda^a(s)$$

$$\times \sum_{j=1}^k \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1\right) N^{a, a_j}(\mathrm{d}s)$$

In the absolutely continuous case, $[K, \mathbb{L}^a]_t = 0$ as $\Delta \Lambda_t^a = 0$ for all $t > 0$. If, further, $\pi_t^*(\{a_j\}) = 1$ for some j, then

$$\varphi_t = \int_0^{t \wedge \tau^{g^*}} \left(\frac{1}{\pi_s} - 1\right) M^{a, a_j}(\mathrm{d}s).$$

Proof: To this end, let $v := \mathbb{1}\{W(t) > 0, \widetilde{W}(t) > 0\} = \mathbb{1}\{\tau^{g^*} > t\} = \mathcal{E}(-\mathbb{N}^a)$ and calculate

$$\mathcal{E}(\varphi)v = \frac{\mathcal{E}(K)\mathcal{E}(-\mathbb{L}^a)}{\mathcal{E}(-\mathbb{N}^a)} v$$

$$= \mathcal{E}(K)\mathcal{E}(-\mathbb{L}^a)v$$

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

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

where the last equality follows since $\mathbb{N}^a v \equiv 0$. Note that

$$[K, \mathbb{L}^a]_t = \int_0^t \Delta \mathbb{L}_s^a \sum_{j=1}^k \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1\right) N^{a, a_j}(\mathrm{d}s)$$

$$\stackrel{*}{=} \int_0^{t \wedge \tau^{g^*}} \sum_v \mathbb{1}\{T_{(v-1)} < s \leq T_{(v)}\} \sum_{i \neq g_v^*\left(\mathcal{F}_{T_{(v-1)}}, T_{(v)}\right)} \pi_s(\{a_j\}) \Delta \Lambda^a(s)$$

$$\times \sum_{j=1}^k \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1\right) N^{a, a_j}(\mathrm{d}s)$$

In the case where $\Delta \Lambda_s^a \equiv 0$ and $\pi_s^*(\{a_j\}) = 1$ for some j, then

$$v(K_t - \mathbb{L}_t^a + \mathbb{N}_t^a) = v \int_0^{t \wedge \tau^{g^*}} \left(\frac{1}{\pi_s} - 1\right) M^{a, a_j}(\mathrm{d}s)$$

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

A simple consequence of this is the following. Assume that $Q_{\mathrm{ryalen}} \ll P$ with $Q_{\mathrm{ryalen}} = \frac{\mathcal{E}(-\mathbb{N}^a)}{\mathcal{E}(-\mathbb{L}^a)} \bullet P$. and, say, $\mathcal{E}(\varphi)$ is a uniformly integrable Q_{ryalen}-$\mathcal{F}_t$-martingale, i.e., that $Q \ll Q_{\mathrm{ryalen}}$, then $Q_{\mathrm{ryalen}} \ll P$ implies that $Q \ll P$. This happens for example if $\mathcal{E}(\varphi)$ is uniformly bounded by a constant.

2) Comparison with the Rytgaard et al. (2022)

In Rytgaard et al. (2022), both an exchangeability condition and a positivity condition are presented, but no proof is given that these conditions imply that their target parameter is identified. Our proposal shows that under the conditions of Theorem 1.1, the g-formula given in Rytgaard et al. (2022) causally identifies the counterfactual mean outcome under the assumption that the other martingales are orthogonal to the treatment martingale. Lemma 1 of Ryalen (2024) then gives the desired target parameter. Note that this is slightly weaker than Rytgaard et al. (2022), as they implicitly require that *all* martingales are orthogonal due to their factorization of the likelihood (?).

Theorem 2.1 (g-formula): Let $\Lambda_P^a(\mathrm{d}t)$ denote the total P-$\mathcal{F}_t$-compensator of N^a. Furthermore, let Λ_P^x denote the P-$\mathcal{F}_t$-compensator of N^x for $x \in \{y, \ell\}$. Under positivity, then

1. The Q-$\mathcal{F}_t$ compensator of $N^a(\mathrm{d}t \times \mathrm{d}x)$ is $\pi_t^*(\mathrm{d}x)\Lambda_P^a(\mathrm{d}t)$.
2. The Q-$\mathcal{F}_t$ compensator of N^x is $\Lambda_P^x + \sum_j \left(\frac{\pi_s^*(a_j)}{\pi_s(a_j)} - 1\right)\mathrm{d}\langle M_P^x, M_P^{a,a_j}\rangle_s^P$ for $x \in \{y, \ell\}$.

Consequently, if $\langle M^x, \bar{M}^a\rangle_s^P \equiv 0$ for $x \in \{y, \ell\}$, where $\bar{M}^a(\mathrm{d}t) = N^a(\mathrm{d}t \times \mathcal{A}) - \Lambda_P^a(\mathrm{d}t)$, then the Q-$\mathcal{F}_t$ compensator of N^x is Λ_P^x for $x \in \{y, \ell\}$.

Proof: First note that for a local $\mathcal{F}_t$-martingale X in P, we have

$$\int_0^t \frac{1}{W_{s-}}\,\mathrm{d}\langle W, X\rangle_s^P = \langle K, X\rangle_t^P \tag{4}$$

since we have that $W_t = 1 + \int_0^t W_{s-}\,\mathrm{d}K_s$; whence

$$\langle W, X\rangle_t = \langle 1, X\rangle_t + \langle W_- \bullet K, X\rangle_t = W_{t-} \bullet \langle K, X\rangle_t$$

With $X = M^{a,x}$, we find

$$\langle K, M^{a,x} \rangle_t^P = \int_0^t \sum_j \left(\frac{\pi_s^*(a_j)}{\pi_s(a_j)} - 1 \right) \mathrm{d}\langle M_P^{a,a_j}, M_P^{a,x} \rangle_s^P$$

$$= \int_0^t \left(\frac{\pi_s^*(x)}{\pi_s(x)} - 1 \right) \mathrm{d}\langle M_P^{a,x} \rangle_s^P$$

$$+ \sum_{j \neq x} \int_0^t \left(\frac{\pi_s^*(a_j)}{\pi_s(a_j)} - 1 \right) \mathrm{d}\langle M_P^{a,a_j}, M_P^{a,x} \rangle_s^P$$

$$= \int_0^t \left(\frac{\pi_s^*(x)}{\pi_s(x)} - 1 \right) \pi_s(x) \Lambda_P^a(\mathrm{d}s)$$

$$- \int_0^t \left(\frac{\pi_s^*(x)}{\pi_s(x)} - 1 \right) \Delta(\pi(x)\Lambda_P^a)_s \pi_s(x) \Lambda_P^a(\mathrm{d}s) \qquad (5)$$

$$- \sum_{j \neq x} \int_0^t \left(\frac{\pi_s^*(a_j)}{\pi_s(a_j)} - 1 \right) \Delta(\pi(x)\Lambda_P^a)_s \pi_s(a_j) \Lambda_P^a(\mathrm{d}s)$$

$$= \int_0^t (\pi_s^*(x) - \pi_s(x)) \Lambda_P^a(\mathrm{d}s)$$

$$- \sum_j \int_0^t (\pi_s^*(a_j) - \pi_s(a_j)) \Delta(\pi(x)\Lambda_P^a)_s \Lambda_P^a(\mathrm{d}s)$$

$$= \int_0^t (\pi_s^*(x) - \pi_s(x)) \Lambda_P^a(\mathrm{d}s).$$

Girsanov's theorem (Theorem 41, p. 136 of Protter (2005)) together with Equation 4 and Equation 5 gives that

$$N^a(\mathrm{d}t \times \mathrm{d}x) - \pi_t(\mathrm{d}x)\Lambda_P^a(\mathrm{d}t) - (\pi_t^*(\mathrm{d}x) - \pi_t(\mathrm{d}x))\Lambda_P^a(\mathrm{d}t) = N^a(\mathrm{d}t \times \mathrm{d}x) - \pi_t^*(\mathrm{d}x)\Lambda_P^a(\mathrm{d}t)$$

is a Q-$\mathscr{F}_t$-local martingale. The second statement follows similarly. $\qquad \square$

2.a) Comparison of the exchangeability condition of Rytgaard et al. (2022)

Additionally, consistency is not explicitly stated, so we cannot compare this condition. The stated positivity condition is the same as the one in Theorem 1.1, so left is to compare the exchangeability conditions.

We can also ask ourselves: Is the exchangeability criterion in Theorem 1.1 close in interpretation to the statement of Rytgaard et al. (2022)? In Rytgaard et al. (2022), the statement is:

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

for all k, where T_k^a are the ordered treatment event times, where $\mathscr{F}_T^-$ 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 one condition (in addition to another) for our exchangeability as something that is very similar (Theorem 2.1.1),

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

for all k. A slightly weaker statement than (*) is that

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

The statements $(')$ and $(*)$ appear similar, but are generally not the same, since S and ΔN_S^a are not generally $\mathcal{F}_T^-$ measurable for a stopping time S.

Let $S = T_k^a$ for some k. If S and ΔN_S^a are $\mathcal{F}_S^-$ measurable, then whenever $(')$ holds, we have

$$\left(\tilde{Y}_t\right)_{t\in[0,T]} \perp A(T_k^a) \mid T_k^a, \Delta N_{T_k^a}^a, \mathcal{F}_{T_k^a}^-$$

which implies $(*)$.

If $N^a(\cdot \times \{0,1\})$ is a predictable counting process with respect to $\mathcal{F}_t$, we will see that this indeed is the case. If this holds, S is predictable, and so $S \in \sigma(\mathcal{F}_S^-)$ (Theorem 2.2.19 of Last & Brandt (1995)). If N_t^a-predictable, then $N_t^a \in \sigma(\mathcal{F}_{t-})$ (Theorem 2.2.6 of Last & Brandt (1995)); therefore $\Delta N_S^a \in \sigma(\mathcal{F}_S^-)$.

To have exchangeability in the sense of Theorem 2.1.1, 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 $N^a(\mathrm{d}t \times \{0,1\})$. (**)

This is obviously the case if $N^a(\cdot \times \mathcal{A})$ is $\mathcal{F}_t$-predictable.

We then have the following result.

> **Theorem 2.1.1**: The conditions (*!) and (**) hold if and only if the exchangeability condition of Theorem 1.1 holds.

Proof: Obvious.

$\square$

3) Score operator calculations

Consider some parametric submodel P_ε of P_0, $\varepsilon \in (-\delta, \delta)$ for some $\delta > 0$. These give rise to submodels of the compensator $\Lambda_\varepsilon^a(A)$. For any $\mathbb{W}^\varepsilon = \mathcal{E}(\mathbb{K}^\varepsilon)$, L'Hopitals rule gives that

$$\lim_{\varepsilon \to 0} \frac{\frac{\mathbb{W}^\varepsilon(t)}{\mathbb{W}^0(t)} - 1}{\varepsilon} = \lim_{\varepsilon^* \to 0} \frac{\frac{\partial}{\partial \varepsilon}\mathbb{W}^\varepsilon(t)}{\mathbb{W}^0(t)} \mid_{\varepsilon=\varepsilon^*}$$

$$\overset{\underset{\text{Thm. 8 of}}{\text{Gill \& Johansen (1990)}}}{=} \lim_{\varepsilon^* \to 0} \frac{\mathbb{W}^{\varepsilon^*}(t)}{\mathbb{W}^0(t)} \int_0^t \frac{1}{1 + \Delta\mathbb{K}_s^{\varepsilon^*}}\left(\frac{\partial}{\partial \varepsilon}\mathbb{K}^\varepsilon(\mathrm{d}s)\right) \mid_{\varepsilon=\varepsilon^*}$$

$$= \int_0^t \lim_{\varepsilon^* \to 0} \frac{1}{1 + \Delta\mathbb{K}_s^{\varepsilon^*}}\left(\frac{\partial}{\partial \varepsilon}\mathbb{K}^\varepsilon(\mathrm{d}s)\right) \mid_{\varepsilon=\varepsilon^*}$$

$$= \int_0^t \frac{1}{1 + \Delta\mathbb{K}_s^0} \lim_{\varepsilon^* \to 0}\left(\frac{\partial}{\partial \varepsilon}\mathbb{K}^\varepsilon(\mathrm{d}s)\right) \mid_{\varepsilon=\varepsilon^*}$$

For our specific case, i.e.,

$$\mathbb{K}_t^\varepsilon = \int_0^t \sum_{j=1}^k \left(\frac{\pi_s^*(\{a_j\})}{\pi_{s,\varepsilon}(\{a_j\})} - 1\right) M_\varepsilon^{a,a_j}(\mathrm{d}s) = \int_0^t \sum_{j=1}^k \left(\frac{\pi_s^*(\{a_j\})}{\pi_{s,\varepsilon}(\{a_j\})} - 1\right) N^{a,a_j}(\mathrm{d}s),$$

but also that

$$\partial\mathbb{K}_t^\varepsilon = (\partial \int_0^t \sum_{j=1}^k \left(\frac{\pi_s^*(\{a_j\})}{\pi_{s,\varepsilon}(\{a_j\})} - 1 \right) M_\varepsilon^{a,a_j}(\mathrm{d}s)$$

$$= \int_0^t (\partial \sum_{j=1}^k \left(\frac{\pi_s^*(\{a_j\})}{\pi_{s,\varepsilon}(\{a_j\})} - 1 \right) N^{a,a_j}(\mathrm{d}s)$$

$$- \int_0^t \sum_{j=1}^k \pi_s^*(\{a_j\})\partial\Lambda_\varepsilon^a(\mathrm{d}s) + \int_0^t \sum_{j=1}^k \partial(\Lambda_\varepsilon^{a,a_j}(\mathrm{d}s))$$

$$= - \int_0^t \sum_{j=1}^k \frac{\pi_s^*(\{a_j\})}{\left(\pi_{s,0}(\{a_j\})\right)^2} \pi_{s,0}(\{a_j\})\Gamma_s^{a,a_j,\pi} N^{a,a_j}(\mathrm{d}s)$$

First note that

$$\pi_{s,\varepsilon}(\{a_j\}) = \frac{\mathrm{d}\Lambda_{\varepsilon,s}^{a,a_j}}{\mathrm{d}\Lambda_{\varepsilon,s}^a}$$

This equals

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

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