

# Influence function calculation for Brier score for event time data

Johan Sebastian Ohlendorff & Thomas Alexander Gerds

2022-06-07

## IF Calculation

To describe the situation with competing risks (and also survival) we introduce a random variable  $D \in \{1, 2\}$  which indicates the cause (i.e., type of the event) observed at time  $T$  such that  $D = 1$  means that the event of interest occurred, and  $D = 2$  that a competing risk occurred. As in the survival setting we let  $Q$  denote the joint probability measure of the uncensored data,  $(T, D, X) \sim Q$ , and  $P$  the joint probability measure of the right censored data  $Z = (\tilde{T}, \Delta, X) \sim P$  now with  $\Delta = D1_{\{T \leq C\}}$  taking values in the set  $\{0, 1, 2\}$ . We are interested in the following definition of the time-dependent discrimination measure for cause 1. We can easily calculate the influence function for the Brier score, which can be written as:

$$\begin{aligned} & \int \{1_{\{t \leq \tau\}} - R(\tau | x)\}^2 Q(dt, dx) \\ &= \int 1_{\{t \leq \tau\}} - 21_{\{t \leq \tau\}}R(\tau | x) + R(\tau | x)^2 Q(dt, dx) \\ &= \int 1_{\{t \leq \tau\}}(1 - 2R(\tau | x))Q(dt, dx) + \int R(\tau | x)^2 Q(dx) \\ &= \int \frac{1_{\{t \leq \tau\}}}{G(t - |x)}(1 - 2R(\tau | x))P(dt, 1, dx) + \int R(\tau | x)^2 P(dx) \end{aligned}$$

We find

$$\begin{aligned} IC_{\text{Brier}}(\tilde{T}_i, \Delta_i, X_i; \tau) &= \partial_\varepsilon \int \frac{1_{\{t \leq \tau\}}}{G(t - |x)}(1 - 2R(\tau | x))P_\varepsilon(dt, 1, dx) + \partial_\varepsilon \int R(\tau | x)^2 P_\varepsilon(dx) \\ &= \int 1_{\{t \leq \tau\}}(1 - 2R(\tau | x)) \frac{d(\delta_{\{\tilde{T}_i, \Delta_i, X_i\}})(t, 1, x) + dP(t, 1, x)[f_i(t-, x) - 1]}{G(t - |x)} \\ &\quad + \int R(\tau | x)^2 (\delta_{\{\tilde{T}_i, \Delta_i, X_i\}} - P) \\ &= 1_{\{\tilde{T}_i \leq \tau, \Delta_i = 1\}}(1 - 2R(\tau | X_i)) \frac{1}{G(\tilde{T}_i - |X_i)} \\ &\quad + \int 1_{\{t \leq \tau\}}(1 - 2R(\tau | x))f_i(t-, x) \frac{dP(t, 1, x)}{G(t - |x)} + R(\tau | X_i)^2 \\ &\quad - \int R(\tau | x)^2 dP(x) - \int 1_{\{t \leq \tau\}}(1 - 2R(\tau | x)) \frac{dP(t, 1, x)}{G(t - |x)} \end{aligned}$$

The last term that is subtracted is the Brier score. Then using that

$$f_i(t, z) = \frac{1_{\{\tilde{T}_i \leq t, \Delta_i = 0\}} \delta_{Z_i}(z)}{G(\tilde{T}_i | Z_i) S(\tilde{T}_i | Z_i)} - \int_0^{\tilde{T}_i \wedge t} \frac{\delta_{Z_i}(z) dP(s, 0 | z)}{G(s | Z_i)^2 S(s | Z_i)^2}$$

we see that by splitting the double integral into two parts that and that  $\int f(z) \delta_{Z_i}(z) dz = f(Z_i)$

$$\int 1_{\{t \leq \tau\}}(1 - 2R(\tau | x))f_i(t-, x) \frac{dP(t, 1, x)}{G(t - |x)} = (1 - 2R(\tau | X_i)) \int 1_{\{t \leq \tau\}} \tilde{f}_i(t-) \frac{dP(t, 1)}{G(t - |X_i)}$$

where  $\tilde{f}_i(t) = \frac{\mathbb{1}_{\{\tilde{T}_i \leq t, \Delta_i = 0\}}}{G(\tilde{T}_i|Z_i)S(\tilde{T}_i|Z_i)} - \int_0^{\tilde{T}_i \wedge t} \frac{dP(s, 0|Z_i)}{G(s|Z_i)^2 S(s|Z_i)^2}$ . Rewriting  $\tilde{f}_i(t)$  a bit, we get

$$\tilde{f}_i(t) = \frac{\mathbb{1}_{\{\tilde{T}_i \leq t, \Delta_i = 0\}}}{G(\tilde{T}_i|Z_i)S(\tilde{T}_i|Z_i)} + \int_0^{\tilde{T}_i \wedge t} \frac{S(s|Z_i)G(ds|Z_i)}{G(s|Z_i)^2 S(s|Z_i)^2} = \frac{\mathbb{1}_{\{\tilde{T}_i \leq t, \Delta_i = 0\}}}{G(\tilde{T}_i|Z_i)S(\tilde{T}_i|Z_i)} + \int_0^{\tilde{T}_i \wedge t} \frac{1}{G(s|Z_i)S(s|Z_i)} \frac{G(ds|Z_i)}{G(s|Z_i)}$$

As for the estimation, we have

$$\begin{aligned} IC_{\text{Brier}}(\tilde{T}_i, \Delta_i, X_i; \tau) &= \mathbb{1}_{\{\tilde{T}_i \leq \tau, \Delta_i = 1\}} (1 - 2R(\tau | X_i)) \frac{1}{\hat{G}(\tilde{T}_i - |X_i)} \\ &\quad + \int \mathbb{1}_{\{t \leq \tau\}} (1 - 2R(\tau | x)) \hat{f}_i(t-, x) \frac{dIP_n(t, 1, x)}{\hat{G}(t - |x)} + R(\tau | X_i)^2 \\ &\quad - \int R(\tau | x)^2 dIP_n(x) - \int \mathbb{1}_{\{t \leq \tau\}} (1 - 2R(\tau | x)) \frac{dIP_n(t, 1, x)}{\hat{G}(t - |x)} \end{aligned}$$

Then using that