the full data model) so that we can enforce $D_h \in T^{F,\perp}_{nuis}$ or just using the bootstrap. We recommend the latter.

3.3.4 Projecting on the tangent space of the Cox proportional hazards model of the censoring mechanism

Let $H(P_{F_X,G}) \subset T_{CAR}(P_{F_X,G})$ be a subspace of $T_{CAR}(P_{F_X,G})$ and $(Q_1,G) \to IC_{nu}(\cdot \mid Q_1,G)$ be a mapping from $\mathcal{Q}_1 \times \mathcal{G}$ to pointwise-defined functions of Y so that $\{IC_{nu}(\cdot \mid Q_1,G) : Q_1 \in \mathcal{Q}_1\} \subset H(P_{F_X,G})$ for all $G \in \mathcal{G}$. As in Chapter 2, we define a more efficient choice of IC_0 as

$$IC_0(Y \mid Q_1,G,D) = \frac{D(X)\Delta(D)}{\bar{G}(V(D) \mid X)} - IC_{nu}(\cdot \mid Q_1,G,D),$$

where the true parameter value of Q_1 is $Q_1(F_X,G)$ defined by

$$IC_{nu}(\cdot \mid Q_1(F_X,G),G,D) = \Pi_{F_X,G}\left(\frac{D(X)\Delta(D)}{\bar{G}(V(D) \mid X)} \mid H(P_{F_X,G})\right).$$

The following lemma establishes this projection IC_{nu} for the case where $H(P_{F_X,G})$ is the tangent space of G under the Cox proportional hazards model.

Lemma 3.2 *Consider the partial likelihood of the counting process $A(t) = I(C \le t)$ w.r.t. observed history $\mathcal{F}(t) = (\bar{A}(t-), \bar{X}_A(\min(t-,C)))$ for a multiplicative intensity model $\alpha_A(t)dt \equiv E(dA(t) \mid \mathcal{F}(t)) = Y_A(t)\lambda_{0A}(t)\exp(\beta_A W(t))dt$,*

$$L(\beta_A,\Lambda_{0A}) = \prod_t \alpha_A(t)^{dA(t)}(1-\alpha_A(t)dt)^{1-dA(t)},$$

where $\prod$ denotes the product integral (Gill and Johansen, 1990; Andersen, Borgan, Gill, and Keiding, 1993). Let $dM_G(t) \equiv dA(t) - \alpha_A(t)dt$. The score for the regression parameter β_A is given by

$$S_{\beta_A} = \int W(t)dM_G(t).$$

The tangent space $T_{\Lambda_{0A}}$ is given by

$$T_{\Lambda_{0A}} = \overline{\left\{\int g(t)dM_G(t) : g\right\}}.$$

Thus the tangent space $T_{\Lambda_{0A},\beta_A}(P_{F_X,G})$ generated by the censoring mechanism parameters (Λ_{0A},β_A) is given by

$$T_{\Lambda_{0A},\beta_A}(P_{F_X,G}) = \langle \int W(t)dM_G(t)\rangle + \overline{\left\{\int g(t)dM_G(t) : g\right\}}.$$

We have

$$\Pi\left(\int H(t,\mathcal{F}(t))dM_G(t) \mid T_{\Lambda_{0A}}\right) = \int g(H)(t)dM_G(t) \quad (3.14)$$

$$\equiv \int \frac{E\{H(t,\mathcal{F}(t))Y_A(t)\exp(\beta_A W(t))\}}{E\{Y_A(t)\exp(\beta_A W(t))\}}dM_G(t).$$

Thus, the efficient score of β_A is given by

$$\begin{aligned} S^*_{\beta_A} &= S_{\beta_A} - \Pi(S_{\beta_A} \mid T_{\Lambda_{0A}}) \\ &= \int \left\{W(t) - \frac{E\{W(t)Y_A(t)\exp(\beta_A W(t))\}}{E\{Y_A(t)\exp(\beta_A W(t))\}}\right\} dM_G(t) \end{aligned}$$

and

$$T_{\Lambda_{0A},\beta_A}(P_{F_X,G}) = \langle S^*_{\beta_A}\rangle \oplus \overline{\left\{\int g(t)dM_G(t) : g\right\}}. \quad (3.15)$$

Consider now the special case where $A(t)$ is discrete on a fine grid $t_1 < \ldots < t_k$ so that the assumed multiplicative intensity model $\alpha_A(t_j) = E(dA(t_j) \mid \mathcal{F}(t_j)) = Y_A(t_j)\lambda_{0A}(t_j)\exp(\beta_A W(t_j))$, $j = 1,\ldots,k$, is appropriate: note that $\alpha_A(t_j)$ are now conditional probabilities. Given any function $V(Y) \in L^2(P_{F_X,G})$, we have $\Pi(V \mid T_{CAR}) = \int H_V(t,\mathcal{F}(t))dM_G(t)$, where $H_V(t,\mathcal{F}(t)) = H_{V,1}(t,\bar{X}(t)) - H_{V,2}(t,\bar{X}(t))$ with $H_{V,1} = E(V(Y) \mid C = t, \bar{X}(t))$ and $H_{V,2} = E(V(Y) \mid C > t, \bar{X}(t))$. Thus, given any function $V(Y)$, we have

$$\Pi(V \mid T_{\Lambda_{0A}}) = \int g(H_V)(t)dM_G(t).$$

Thus, given any function $V(Y)$, we have that $\Pi(V \mid T_{\Lambda_{0A},\beta_A})$ is given by

$$E(VS^{*\top}_{\beta_A})E(S^*_{\beta_A}S^{*\top}_{\beta_A})^{-1}S^*_{\beta_A} + \int \{g(H_{V,1}) - g(H_{V,2})\}(t)dM_G(t). \quad (3.16)$$

Consider the case where $Y_A(t) = I(C \ge t, T > t)$. We have for any $V = V(X,C)$ (thus, in particular, for $V = V(Y)$)

$$g(H_{V,1})(t) = \frac{E\left\{V(X,t)I(T \ge t)\bar{G}(t \mid X)\exp(\beta_A W(t))\frac{I(C \ge \min(t,T))}{\bar{G}(\min(t,T) \mid X)}\right\}}{E\{Y_A(t)\exp(\beta_A W(t))\}}$$

and

$$g(H_{V,2})(t) = \frac{E\{Q_{V,2}(\bar{X}(t))I(C \ge \min(t,T))/\bar{G}(\min(t,T) \mid X)\}}{E\{Y_A(t)\exp(\beta_A W(t))\}},$$

where $Q_{V,2}(\bar{X}(t)) = E\{V(X,C)I(T \ge t)\bar{G}(t \mid X)\exp(\beta_A W(t)) \mid C > t, \bar{X}(t)\}$ and $\bar{G}(t \mid X) \equiv P(C \ge t \mid X)$. (Note that we can set $I(C \ge \min(t,T))/\bar{G}(\min(t,T) \mid X)$ equal to 1 as well in both formulas, but the inclusion of this term shows how one can estimate it from the observed data).

In particular, if $U_G(D_h)(Y) = D_h(X)\Delta/\bar{G}(T \mid X)$, then

$$\Pi(U_G(D_h) \mid T_{\Lambda_{0A},\beta_A}) = E(U_G(D_h)S^{*\top}_{\beta_A})E(S^*_{\beta_A}S^{*\top}_{\beta_A})^{-1}S^*_{\beta_A}$$