3 Monotone Censored Data

3.1 Data Structure and Model

Let \( \{X(t) : t \in \mathbb{R}_{\geq 0}\} \) be a multivariate stochastic process indexed by time \( t \). Let \( T \) denote an endpoint of this stochastic process, and define \( X(t) = X(\min(t,T)) \). Let \( R(t) = I(T \leq t) \) be one of the components of \( X(t) \). We define the full data as \( X = \bar{X}(T) = (X(s) : s \leq T) \), where \( T \) is thus a function \( X \).

Suppose that we observe the full data process \( X(\cdot) \) up to the minimum of a univariate censoring variable \( C \) and \( T \) so that for the observed data we have:

\[
Y = (\bar{T} = \min(T,C), \Delta = I(T \leq \bar{T}) = I(C \geq T), \bar{X}(\bar{T})).
\]

We will define \( C = \infty \) if \( C > T \) so that this data structure can be represented as

\[
Y = (C, \bar{X}(C)).
\]

In the next section, we provide several important examples of this monotone censored data structure.

Let \( \mathcal{M}^F \) be a specified full data model for the distribution \( F_X \) of \( X \), and let \( \mu = \mu(F_X) \in \mathbb{R}^k \) be the full data parameter of interest. Let \( G(\cdot | X) \) be the conditional distribution of \( C \), given \( X \), and it is assumed that \( G \) satisfies CAR (i.e., \( G \in \mathcal{G}(\text{CAR}) \)). Given working models \( \mathcal{M}^{F,w} \subset \mathcal{M}^F \) and \( G \subset \mathcal{G}(\text{CAR}) \), we define the observed data model \( \mathcal{M} = \{P_{F_X,G} : F_X \in \mathcal{M}^{F,w}\} \cup \{P_{F_X,G} : G \in \mathcal{G}\} \). We also define the observed data model \( \mathcal{M}(\mathcal{G}) = \{P_{F_X,G} : F_X \in \mathcal{M}^F, G \in \mathcal{G}\} \). Candidates for the censoring model \( G \) are given below.

Let \( g(c | X) \) be the conditional density of \( C \), given \( X \), either w.r.t. a Lebesgue density or counting measure, and let \( \lambda_C(c | X) \) be the corresponding conditional hazard. Define \( A(t) \equiv I(C \leq t) \). The conditional distribution \( G \) satisfies CAR if

\[
E(dA(t) | X, \bar{A}(t-)) = E(dA(t) | \bar{A}(t-), \bar{X}(\min(t,C))).
\]

In other words, the intensity of \( A(t) \) w.r.t. the unobserved history \( (X, \bar{A}(t-)) \) should equal the intensity of \( A(t) \) w.r.t. the observed history \( (\bar{A}(t-), \bar{X}(\min(t,C))) \). Equivalently, \( G \) satisfies CAR if for \( c < T \)

\[
\lambda_C(c | X) = m(c, \bar{X}(c)) \text{ for some measurable function } m.
\]

If \( C \) is continuous, then a practical and useful submodel \( \mathcal{G} \subset \mathcal{G}(\text{CAR}) \) is the multiplicative intensity model w.r.t. the Lebesgue measure

\[
E(dA(t) | \bar{A}(t-), \bar{X}(\min(t,C))) = I(\bar{T} > t) \lambda_0(t) \exp \left( \alpha_0^T W(t) \right),
\]

where \( \alpha_0 \) is a \( k \)-dimensional vector of coefficients, \( W(t) \) is a \( k \)-dimensional time-dependent vector that is a function of \( \bar{X}(t) \), and \( \lambda_0 \) is an unspecified baseline hazard. Note that

\[
\lambda_C(t | X, T > t) \equiv \lambda_0(t) \exp(\alpha_0^T W(t))
\]

denotes the Cox proportional hazards model for the conditional hazard \( \lambda_C \).

If we knew that the censoring was independent of the survival time and the history, then, for \( t < T \), this would reduce to

\[
\lambda_C(t | X) = \lambda_0(t).
\]

If \( C \) is discrete, then a natural model \( \mathcal{G} \subset \mathcal{G}(\text{CAR}) \) is

\[
E(dA(t) | \bar{A}(t-), \bar{X}(\min(t,C))) = I(\bar{T} \geq t) \frac{1}{1 + \exp(-\{h_0(c) + \alpha_0^T W(t)\})},
\]

where \( h_0 \) could be left unspecified. This corresponds with assuming a logistic regression model for the conditional censoring hazard \( \lambda_C(t | X) = P(C = t | X, C \geq t) \): for \( t < T \)

\[
\log \left( \frac{\lambda_C(t | X)}{1 - \lambda_C(t | X)} \right) = h_0(t) + \alpha_0^T W(t).
\]

If the support of \( C \) gets finer and finer so that \( P(C = t | X, C \geq t) \) approximates zero, then this model with \( h_0 \) unspecified converges to the Cox proportional hazards model with \( \lambda_0 = \exp(h_0) \) and regression coefficients \( \alpha_0 \) (see e.g., Kalbfleisch and Prentice, 1980).

Whatever CAR model for \( \lambda_C(t | X) \) is used, the \( G \) part of the density of \( P_{F_X,G} \) in terms of \( \lambda_C(t | X) \) is given by the partial likelihood of \( A(t) = I(C \leq t) \) w.r.t. history \( F(t) \equiv (\bar{A}(t-), \bar{X}(\min(t,C))) \) as defined in Andersen, Borgan, Gill and Keiding (1993) for the continuous case. Let