

If we define the process $X(\cdot) = (I(T_1 > \cdot), \dots, I(T_k > \cdot), \bar{L}(\cdot))$, then we can also represent the full data structure as $X = \bar{X}(T)$ and the observed data structure Y as

$$Y = (\tilde{T} = C \wedge T, \Delta = I(\tilde{T} = T), \bar{X}(\tilde{T})).$$

We refer to the latter general data structure as “univariately right-censored multivariate data”. Note that CAR allows the hazard of censoring at c to be a function of $\bar{X}(c)$ and thus of the observed part (up until time c) of (T_1, T_2, \dots) and of time-dependent covariates $\bar{L}(c)$.

Possible parameters of interest are the multivariate failure time distribution and the distribution of waiting times if $T_1 < \dots < T_k$. Note that such parameters can be defined as $\mu = EB$, where $B = b(X)$ is a function of X . If, for given $\bar{t} \in \mathbb{R}^k$, one takes $B = I(\tilde{T} > \bar{t})$, then $\mu = S(\bar{t})$. Likewise, if, for given $t > 0$, $B = I(T_2 - T_1 > t)$, then $\mu = P(T_2 - T_1 > t)$. In addition, regression parameters in a generalized linear regression model of each survival time on baseline covariates and multiplicative intensity model involving the intensities of T_j w.r.t. a history including only a subset of the observed past are of great interest as well.

We will now review previous proposals for estimation of multivariate survival functions in the nonparametric full data model. All previous proposals based on CAR models have imposed the stronger assumption of independent censoring. Because the nonparametric maximum likelihood and self-consistency principles (Efron, 1967; Turnbull, 1976) do not lead to a consistent estimator for continuous survival data, most proposed estimators are explicit representations of the bivariate survival function in terms of distribution functions of the data (see Campbell and Földes, 1982; Tsai, Leurgans and Crowley, 1986; Dabrowska, 1988 and 1989; Burke, 1988; the so-called Volterra estimator of P.J. Bickel in Dabrowska, 1988; Prentice and Cai, 1992a and 1992b). These explicit estimators are generally inefficient, but because they are explicit, their influence curves can be explicitly calculated so asymptotic confidence intervals are easy to compute (see Gill, 1992; Gill, van der Laan and Wellner, 1995). In van der Laan (1996a), a modified NPMLE of the bivariate survival function, which requires a choice of a partition used to reduce the data, is proposed that is shown to be asymptotically efficient. The methods above allow but do not require that all failure times are censored by a common variable C . In contrast, Wang and Wells (1998) and Lin, Sun, and Ying (1999) assume that $T_1 < \dots < T_k$ with probability 1 and, as in the present example, the failure times are all right-censored by the same censoring variable C . The Lin, Sun, and Ying (1999) estimator is an inverse of probability of censoring weighted (IPCW) estimator as proposed by Robins and Rotnitzky (1992) and defined for general CAR-censored data models in Gill, van der Laan, and Robins (1997). None of the explicit estimators are efficient. Further, none of the estimators incorporate data on prognostic covariates such as $L(t)$. As a consequence, all are inconsistent under informative right-

censoring (i.e., when $\lambda_C(t | X)$ does actually depend on $\bar{X}(t)$). In van der Laan, Hubbard and Robins (2002), locally efficient estimators of the multivariate failure time distribution and waiting time distributions based on the general multivariate failure time data structure (3.5) are provided that are analogous to the methods presented in this chapter.

Finally, we remark that multiplicative intensity models for multivariate counting processes provide estimators of intensities for these data structures, but these methods are not appropriate (see Section 3.1) for estimating more marginal parameters. The literature on frailty models provides an extension of the multiplicative intensity model methodology for multivariate survival times by assuming that the time variables are independent given an unobserved time-independent frailty and the past (see e.g., Clayton and Cuzick, 1985; Hougaard, 1986; Oakes, 1989; Clayton, 1991; Klein, 1992; Costigan and Klein, 1993). In particular, this extension is implemented in S-plus as part of the Coxph function.

3.3 Inverse Probability Censoring Weighted (IPCW) Estimators

Let $\{D_h : h \in \mathcal{H}^F\}$ be a set of full data estimating functions $(\mu, \rho, X) \rightarrow D_h(X | \mu, \rho)$ for μ with nuisance parameter ρ indexed by $h \in \mathcal{H}^F$. Let $\mathcal{D} = \{D_h(\cdot | \mu, \rho) : h \in \mathcal{H}^F, \mu, \rho\}$ be the corresponding set of full data structure functions.

For $D(X) \in \mathcal{D}$, we define $\Delta(D) = I(D(X) \text{ is observed})$. There exists a real-valued random variable $V(D) \leq T$ so that $I(D(X) \text{ is observed}) = I(C \geq V(D))$. For $D \in \mathcal{D}$, define

$$IC_0(Y | G, D) = \frac{D(X)\Delta(D)}{P_G(\Delta(D) = 1 | X)} = \frac{D(X)\Delta(D)}{\bar{G}(V(D) | X)}, \quad (3.6)$$

where $\bar{G}(t | X) = P(C \geq t | X)$.

In model $\mathcal{M}(G)$, we can obtain an initial estimator of μ by solving

$$0 = \sum_{i=1}^n IC_0(Y_i | G_n, D_{h_n}(\cdot | \mu, \rho_n)),$$

where G_n is an estimator of G , and $h_n \in \mathcal{H}^F$ is a possibly data-dependent index specified by the user.

3.3.1 Identifiability condition

Let $\mathcal{D}(\rho_1, G)$ be defined (as in (2.13)) by

$$\begin{aligned} \mathcal{D}(\rho_1, G) &\equiv \{D \in \mathcal{D} : E_G(IC_0(Y | G, D) | X) = D(X) \text{ } F_X\text{-a.e.}\} \\ &= \{D \in \mathcal{D} : \bar{G}(V(D) | X) > 0 \text{ } F_X\text{-a.e.}\}. \end{aligned} \quad (3.7)$$