THE COX PARTIAL LIKELIHOOD
AS A LIMIT FROM A PIECEWISE EXPONENTIAL MODEL

Suppose we have survival data with maximum follow-up time τ^*. Let us partition the interval $[0, \tau^*]$ into K subintervals $(\tau_{k-1}, \tau_k], k = 1, \ldots, K$, with $\tau_k = kh, k = 0, \ldots, K$, where $h = \tau^*/K$. Consider the following version of the Cox model with a piecewise constant hazard:

$$\lambda(t|\mathbf{z}) = \lambda_0(t) e^{\boldsymbol{\beta}^T \mathbf{z}},$$

with

$$\lambda_0(t) = \lambda_{0k}, \quad t \in (\tau_{k-1}, \tau_k], \quad k = 1, \ldots, K.$$

We can write

$$\lambda_0(t) = \sum_{k=1}^K \lambda_{0k} \xi_k(t), \quad \xi_k(t) = I(t \in (\tau_{k-1}, \tau_k]).$$

We have $\Lambda(t|\mathbf{z}) = \Lambda_0(t) e^{\boldsymbol{\beta}^T \mathbf{z}}$ with $\Lambda_0(t) = \int_0^t \lambda_0(u) du$. We can express $\Lambda_0(t)$ as

$$\Lambda_0(t) = \sum_{k=1}^K \lambda_{0k} \Delta_k(t),$$

where

$$\Delta_k(t) = \int_0^t \xi_k(u) du = \begin{cases} 0 & \text{if } t \le \tau_{k-1} \\ t - \tau_{k-1} & \text{if } t \in (\tau_{k-1}, \tau_k] \\ h & \text{if } t > \tau_{k-1} \end{cases} .$$

We have now a parametric survival model where the parameters are $\boldsymbol{\beta}$ and $\lambda_{01}, \ldots, \lambda_{0k}$. As usual, denote by X_i the follow-up time on subject i and by δ_i the event status (1=event, 0=censoring). The log-likelihood is then given by

$$\begin{aligned} \ell &= \sum_{i=1}^n \delta_i \log \lambda(X_i|\mathbf{Z}_i) - \sum_{i=1}^n \Lambda(X_i|\mathbf{Z}_i) \\ &= \sum_{i=1}^n \delta_i \log \left[e^{\boldsymbol{\beta}^T \mathbf{Z}_i} \sum_{k=1}^K \lambda_{0k} \xi_k(X_i) \right] - \sum_{i=1}^n \left[e^{\boldsymbol{\beta}^T \mathbf{Z}_i} \sum_{k=1}^K \lambda_{0k} \Delta_k(X_i) \right] \\ &= \sum_{i=1}^n \delta_i \boldsymbol{\beta}^T \mathbf{Z}_i + \sum_{i=1}^n \delta_i \log \left[\sum_{k=1}^K \lambda_{0k} \xi_k(X_i) \right] - \sum_{i=1}^n e^{\boldsymbol{\beta}^T \mathbf{Z}_i} \left[\sum_{k=1}^K \lambda_{0k} \Delta_k(X_i) \right]. \end{aligned}$$

Differentiating with respect to λ_{0q}, we obtain

$$\begin{aligned} \frac{\partial \ell}{\partial \lambda_{0q}} &= \sum_{i=1}^n \frac{\delta_i \xi_q(X_i)}{\sum_{k=1}^K \lambda_{0k} \xi_k(X_i)} - \sum_{i=1}^n e^{\boldsymbol{\beta}^T \mathbf{Z}_i} \Delta_q(X_i) \\ &= \lambda_{0q}^{-1} \sum_{i=1}^n \delta_i \xi_q(X_i) - \sum_{i=1}^n e^{\boldsymbol{\beta}^T \mathbf{Z}_i} \Delta_q(X_i). \end{aligned}$$

The last step follows from the fact that $\xi_q(t) = 1$ implies $\xi_k(t) = 0$ for $k \neq q$. Setting the derivative equal to zero yields

$$\hat{\lambda}_{0q} = \frac{\sum_{i=1}^n \delta_i \xi_q(X_i)}{\sum_{i=1}^n e^{\boldsymbol{\beta}^T \mathbf{Z}_i} \Delta_q(X_i)}.$$

Let us now substitute this result back into the expression for the log-likelihood, obtaining a profile log-likelihood for $\boldsymbol{\beta}$. We get

$$\begin{aligned} \ell &= \sum_{i=1}^n \delta_i \boldsymbol{\beta}^T \mathbf{Z}_i + \sum_{i=1}^n \delta_i \log \left[\sum_{k=1}^K \left(\frac{\sum_{j=1}^n \delta_j \xi_k(X_j)}{\sum_{j=1}^n e^{\boldsymbol{\beta}^T \mathbf{Z}_j} \Delta_k(X_j)} \right) \xi_k(X_i) \right] \\ &\quad - \sum_{i=1}^n e^{\boldsymbol{\beta}^T \mathbf{Z}_i} \left[\sum_{k=1}^K \left(\frac{\sum_{j=1}^n \delta_j \xi_k(X_j)}{\sum_{j=1}^n e^{\boldsymbol{\beta}^T \mathbf{Z}_j} \Delta_k(X_j)} \right) \Delta_k(X_i) \right] \\ &= \sum_{i=1}^n \delta_i \boldsymbol{\beta}^T \mathbf{Z}_i + \sum_{i=1}^n \delta_i \left[\sum_{k=1}^K \xi_k(X_i) \log \left(\frac{\sum_{j=1}^n \delta_j \xi_k(X_j)}{\sum_{j=1}^n e^{\boldsymbol{\beta}^T \mathbf{Z}_j} \Delta_k(X_j)} \right) \right] \\ &\quad - \sum_{k=1}^K \left[\left(\frac{\sum_{j=1}^n \delta_j \xi_k(X_j)}{\sum_{j=1}^n e^{\boldsymbol{\beta}^T \mathbf{Z}_j} \Delta_k(X_j)} \right) \left(\sum_{i=1}^n e^{\boldsymbol{\beta}^T \mathbf{Z}_i} \Delta_k(X_i) \right) \right]. \end{aligned}$$

In the last step, the expression for the second term is obtained by using the fact that, for a given t, the value of $\xi_k(t)$ is equal to 1 for exactly one k and 0 for the others, so that the sum over k in the corresponding term in the preceding line includes only one nonzero summand. The expression for the third term is obtained by interchanging the order of summation. At this point, note that a cancellation takes place in the third term. We thus find that ℓ is given by

$$\begin{aligned} \ell &= C_1 + \sum_{i=1}^n \delta_i \boldsymbol{\beta}^T \mathbf{Z}_i - \sum_{i=1}^n \delta_i \sum_{k=1}^K \xi_k(X_i) \log \left[\sum_{j=1}^n e^{\boldsymbol{\beta}^T \mathbf{Z}_j} \Delta_k(X_j) \right] \\ &= C_2 + \sum_{i=1}^n \delta_i \boldsymbol{\beta}^T \mathbf{Z}_i - \sum_{i=1}^n \delta_i \log \left[\sum_{j=1}^n e^{\boldsymbol{\beta}^T \mathbf{Z}_j} h^{-1} \Delta_{k(i)}(X_j) \right], \end{aligned}$$

where C_1 and C_2 denote constants that do not depend on $\boldsymbol{\beta}$, and $k(i)$ denotes the index k for which $X_i \in (\tau_{j-1}, \tau_k]$. Now for small h, the probability that X_i and X_j will fall in the same interval is $O(h^2)$. Hence, recalling the definition of $\Delta_k(t)$, we find that for small h we have $h^{-1}\Delta_{k(i)}(X_j) \doteq I(X_j \geq X_i)$. We thus find that, as $h \to 0$, the function ℓ tends to the following (up to a constant that does not depend on $\boldsymbol{\beta}$):

$$\sum_{i=1}^n \delta_i \left\{ \boldsymbol{\beta}^T \mathbf{Z}_i - \log \left[\sum_{j=1}^n I(X_j \geq X_i) e^{\boldsymbol{\beta}^T \mathbf{Z}_j} \right] \right\}.$$

This last expression is precisely the log of the Cox partial likelihood.