The Right-Censored Bivariate Semi-Nonparametric Mixed Proportional Hazard Model and its Implementation in EasyReg

Herman J. Bierens

March 8, 2007

Abstract

This note contains the setup of the model in Bierens and Carvalho (2007), and its implementation in EasyReg module SNPSURVIVAL2.

1 Introduction

Consider two durations, T_1 and T_2. Conditional on a vector X of covariates, and a common unobserved heterogeneity variable V, the durations T_1 and T_2 are assumed to be independent, i.e.

$$P\left[T_1 \leq t_1, T_2 \leq t_2 | X, V\right] = P\left[T_1 \leq t_1 | X, V\right] P\left[T_2 \leq t_2 | X, V\right] \tag{1}$$

We only observe $T = \min\left(T_1, T_2\right)$ and a discrete variable D which is 1 if $T_2 > T_1$ and 2 if $T_2 < T_1$:

$$D = 1 + I\left(T_2 < T_1\right),$$

where $I(.)$ is the indicator function. Also, there may be right-censoring, which will be indicated by a dummy variable C. Thus, given that T is only observed over the period $[0, \overline{T}]$, $C = I\left(T > \overline{T}\right)$. The observed T will be equal to $\overline{T}$ if $C = 1$.

For each duration T_i we assume a mixed proportional hazard model. Thus, the distribution function and corresponding survival function of T_i conditional on X and V take the form

$$F_i(t|X,V) = P[T_i \leq t|X,V] = 1 - \exp\left(-V \exp\left(\beta_i' X\right) \Lambda_i\left(t|\alpha_i\right)\right) \quad (2)$$
$$S_i(t|X,V) = 1 - F_i(t|X,V) = \exp\left(-V \exp\left(\beta_i' X\right) \Lambda_i\left(t|\alpha_i\right)\right) \quad (3)$$

$i = 1, 2$, where $\Lambda_i\left(t|\alpha_i\right) = \int_0^t \lambda_i\left(\tau|\alpha_i\right) d\tau$ is the integrated baseline hazard of T_i depending on parameter vector α_i, with $\lambda_i\left(t|\alpha_i\right)$ the baseline hazard.

The role of the common unobserve heterogeneity variable V is to make the two durations T_1 and T_2 dependent, conditional on the vector X of covariates alone, by integrating V out in (1).

2 Distributions of the dependent variables conditional on covariates and unobserved heterogeneity

For notational convenience the dependence of $\Lambda_i\left(t|\alpha_i\right)$ and $\lambda_i\left(t|\alpha_i\right)$ on parameters will be suppressed for the time being: $\Lambda_i(t) = \Lambda_i\left(t|\alpha_i\right)$ and $\lambda_2(t) = \lambda_i\left(t|\alpha_i\right)$.

Observe that the conditional density of T_i given X and V takes the form

$$\begin{aligned} f_i(t|X,V) &= \partial F_i(t|X,V)/\partial t \quad (4) \\ &= V \exp\left(-V \exp\left(\beta_i' X\right) \Lambda_i(t)\right) \exp\left(\beta_i' X\right) \lambda_i(t), \end{aligned}$$

Then conditional on X and V,

$$\begin{aligned} & P[T > t, D = 1, C = 0|X,V] = P\left[t < T_1 \leq \overline{T}, T_2 > T_1|X,V\right] \quad (5) \\ & = E\left[I\left(t < T_1 \leq \overline{T}, T_2 > T_1\right) X, V\right] \\ & = \int_0^\infty \int_0^\infty I\left(t < t_1 \leq \overline{T}, t_2 > t_1\right) f_1\left(t_1|X,V\right) f_2\left(t_2|X,V\right) dt_1 dt_2 \\ & = \int_t^{\overline{T}} f_1\left(t_1|X,V\right) \left(\int_{t_1}^\infty f_2\left(t_2|X,V\right) dt_2\right) dt_1 \\ & = F_1\left(\overline{T}|X,V\right) - F_1(t|X,V) - \int_t^{\overline{T}} F_2(\tau|X,V) f_1(\tau|X,V) d\tau \end{aligned}$$

$$= S_1 (t|X, V) - S_1 \left(\overline{T}|X, V\right) - \int_t^{\overline{T}} F_2 (\tau|X, V) f_1 (\tau|X, V) d\tau$$
$$= \int_t^{\overline{T}} S_2 (\tau|X, V) f_1 (\tau|X, V) d\tau$$
$$= \int_t^{\overline{T}} V \exp (-V (\exp (\beta_1' X) \Lambda_1 (\tau) + \exp (\beta_2' X) \Lambda_2 (\tau)))$$
$$\times \exp (\beta_1' X) \lambda_1 (\tau) d\tau$$

and similarly,

$$P [T > t, D = 2, C = 0|X, V] = \int_t^{\overline{T}} S_1 (\tau|X, V) f_2 (\tau|X, V) d\tau \quad (6)$$
$$= \int_t^{\overline{T}} V \exp (-V (\exp (\beta_1' X) \Lambda_1 (\tau) + \exp (\beta_2' X) \Lambda_2 (\tau)))$$
$$\times \exp (\beta_2' X) \lambda_2 (\tau) d\tau$$

Moreover,

$$P [C = 1|X, V] = P \left[T_1 > \overline{T}|X, V\right] P \left[T_2 > \overline{T}|X, V\right] \quad (7)$$
$$= S_1 \left(\overline{T}|X, V\right) S_2 \left(\overline{T}|X, V\right)$$
$$= \exp \left(-V \left(\exp (\beta_1' X) \Lambda_1 \left(\overline{T}\right) + \exp (\beta_2' X) \Lambda_2 \left(\overline{T}\right)\right)\right)$$

3 Integrating the unobserved heterogeneity out

Let $G(v)$ be the distribution function of V, and let

$$H (u) = \int_0^\infty u^v dG(v) \quad (8)$$
$$h(u) = \int_0^\infty v u^{v-1} dG(v) \quad (9)$$

Then it follows from (7) and (8) that

$$P [C = 1|X] = H \left(\exp \left(- \exp (\beta_1' X) \Lambda_1 \left(\overline{T}\right) - \exp (\beta_2' X) \Lambda_2 \left(\overline{T}\right)\right)\right), \quad (10)$$

and from (5), (6) and (9) that

$$
\begin{aligned}
&P\left[T>t, D=1, C=0 | X\right] \qquad (11)\\
&=\int_t^{\overline{T}} h\left(\exp\left(-\left(\exp\left(\beta_1' X\right)\Lambda_1\left(\tau\right)+\exp\left(\beta_2' X\right)\Lambda_2\left(\tau\right)\right)\right)\right)\\
&\times\exp\left(-\left(\exp\left(\beta_1' X\right)\Lambda_1\left(\tau\right)+\exp\left(\beta_2' X\right)\Lambda_2\left(\tau\right)\right)\right)\exp\left(\beta_1' X\right)\lambda_1\left(\tau\right)d\tau,
\end{aligned}
$$

and

$$
\begin{aligned}
&P\left[T>t, D=2, C=0 | X\right] \qquad (12)\\
&=\int_t^{\overline{T}} h\left(\exp\left(-\left(\exp\left(\beta_1' X\right)\Lambda_1\left(\tau\right)+\exp\left(\beta_2' X\right)\Lambda_2\left(\tau\right)\right)\right)\right)\\
&\times\exp\left(-\left(\exp\left(\beta_1' X\right)\Lambda_1\left(\tau\right)+\exp\left(\beta_2' X\right)\Lambda_2\left(\tau\right)\right)\right)\exp\left(\beta_2' X\right)\lambda_2\left(\tau\right)d\tau.
\end{aligned}
$$

Therefore, the conditional density of T given $D = d(= 1, 2)$ and $C = 0$ is

$$
\begin{aligned}
&f\left(t|X, D=d, C=0\right)\\
&=h\left(\exp\left(-\left(\exp\left(\beta_1' X\right)\Lambda_1\left(t\right)+\exp\left(\beta_2' X\right)\Lambda_2\left(t\right)\right)\right)\right)\\
&\times\exp\left(-\left(\exp\left(\beta_1' X\right)\Lambda_1\left(t\right)+\exp\left(\beta_2' X\right)\Lambda_2\left(t\right)\right)\right)\\
&\times\exp\left(\beta_d' X\right)\lambda_d\left(t\right)/P\left[D=d, C=0|X\right] \text{ if } t\leq\overline{T}\\
&=0 \text{ elsewhere}
\end{aligned}
$$

4 The log-likelihood function

Given i.i.d. observations $\{T_j, D_j, C_j\}_{j=1}^N$ on (T, D, C), the log-likelihood function is

$$
\begin{aligned}
&\ln\left(L_N\left(\alpha_1,\alpha_2,\beta_1,\beta_2,h\right)\right)\\
&=\sum_{j=1}^N\left(1-C_j\right)\left(2-D_j\right)\ln\left(f\left(T_j|X_j, D_j=1, C=0\right)P\left[D_j=1, C_j=0|X\right]\right)\\
&+\sum_{j=1}^N\left(1-C_j\right)\left(D_j-1\right)\ln\left(f\left(T_j|X_j, D_j=2, C=0\right)P\left[D_j=2, C_j=0|X\right]\right)\\
&+\sum_{j=1}^N C_j\ln\left(P\left[C_j=1|X_j\right]\right)
\end{aligned}
$$

$$= \sum_{j=1}^{N} C_j \ln \left(H \left(\exp \left(- \left(\exp \left(\beta_1' X_j\right) \Lambda_1 \left(T_j | \alpha_1\right) + \exp \left(\beta_2' X_j\right) \Lambda_2 \left(T_j | \alpha_2\right)\right)\right)\right)\right)$$

$$+ \sum_{j=1}^{N} \left(1 - C_j\right) \ln \left(h \left(\exp \left(- \left(\exp \left(\beta_1' X_j\right) \Lambda_1 \left(T_j | \alpha_1\right) + \exp \left(\beta_2' X_j\right) \Lambda_2 \left(T_j | \alpha_2\right)\right)\right)\right)\right)$$

$$- \sum_{j=1}^{N} \left(1 - C_j\right) \left(\exp \left(\beta_1' X_j\right) \Lambda_1 \left(T_j | \alpha_1\right) + \exp \left(\beta_2' X_j\right) \Lambda_2 \left(T_j | \alpha_2\right)\right)$$

$$+ \sum_{j=1}^{N} \left(1 - C_j\right) \left((2 - D_j)\beta_1' X_j + (D_j - 1)\beta_2' X_j\right)$$

$$+ \sum_{j=1}^{N} \left(1 - C_j\right) \left((2 - D_j) \ln \left(\lambda_1 \left(T_j | \alpha_1\right)\right) + (D_j - 1) \ln \left(\lambda_2 \left(T_j | \alpha_2\right)\right)\right)$$

At this point the density $h(u)$ is treated as a parameter.

5 Identification

5.1 Introduction

For $t \leq \overline{T}$, let the true conditional probability $P\left[T \leq t, D = 1, C = 0 | X\right]$ be

$$\begin{aligned} & P\left[T \leq t, D = 1, C = 0 | X\right] \\ & = \int_0^t h_0 \left(\exp \left(- \left(\exp \left(\beta_{0,1}' X\right) \Lambda_{0,1}\left(\tau\right) + \exp \left(\beta_{0,2}' X\right) \Lambda_{0,2}\left(\tau\right)\right)\right)\right) \\ & \times \exp \left(- \left(\exp \left(\beta_{0,1}' X\right) \Lambda_{0,1}\left(\tau\right) + \exp \left(\beta_{0,2}' X\right) \Lambda_{0,2}\left(\tau\right)\right)\right) \\ & \times \exp \left(\beta_{0,1}' X\right) \lambda_{0,1}\left(\tau\right) d\tau \end{aligned} \tag{13}$$

Suppose there exist a density h on $[0, 1]$, parameter vectors β_1, β_2 and hazard functions $\lambda_1(t)$ and $\lambda_2(t)$ with corresponding integrated hazards $\Lambda_1(t)$ and $\Lambda_2(t)$ such that for all $t \leq \overline{T}$,

$$\begin{aligned} & P\left[T \leq t, D = 1, C = 0 | X\right] \\ & = \int_0^t h \left(\exp \left(- \left(\exp \left(\beta_1' X\right) \Lambda_1\left(\tau\right) + \exp \left(\beta_2' X\right) \Lambda_2\left(\tau\right)\right)\right)\right) \\ & \times \exp \left(- \left(\exp \left(\beta_1' X\right) \Lambda_1\left(\tau\right) + \exp \left(\beta_2' X\right) \Lambda_2\left(\tau\right)\right)\right) \\ & \times \exp \left(\beta_1' X\right) \lambda_1\left(\tau\right) d\tau \end{aligned} \tag{14}$$

as well. Taking the derivative to t, it then follows that for all $t \leq \overline{T}$,

$$\begin{aligned} & h_0\left(\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_{0,1}(t)+\exp\left(\beta_{0,2}'X\right)\Lambda_{0,2}(t)\right)\right)\right) \\ & \times\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_{0,1}(t)+\exp\left(\beta_{0,2}'X\right)\Lambda_{0,2}(t)\right)\right) \\ & \times\exp\left(\beta_{0,1}'X\right)\lambda_{0,1}(t) \\ & = h\left(\exp\left(-\left(\exp\left(\beta_1'X\right)\Lambda_1(t)+\exp\left(\beta_2'X\right)\Lambda_2(t)\right)\right)\right) \\ & \times\exp\left(-\left(\exp\left(\beta_1'X\right)\Lambda_1(t)+\exp\left(\beta_2'X\right)\Lambda_2(t)\right)\right) \\ & \times\exp\left(\beta_1'X\right)\lambda_1(t) \text{ a.s.} \end{aligned} \tag{15}$$

Similarly, if for all $t \leq \overline{T}$,

$$\begin{aligned} & P\left[T\leq t, D=2, C=0|X\right] \\ & =\int_0^t h_0\left(\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_{0,1}(\tau)+\exp\left(\beta_{0,2}'X\right)\Lambda_{0,2}(\tau)\right)\right)\right) \\ & \times\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_{0,1}(\tau)+\exp\left(\beta_{0,2}'X\right)\Lambda_{0,2}(\tau)\right)\right) \\ & \times\exp\left(\beta_{0,2}'X\right)\lambda_{0,2}(\tau)\,d\tau \end{aligned}$$

is equal to

$$\begin{aligned} & P\left[T\leq t, D=2, C=0|X\right] \\ & =\int_0^t h\left(\exp\left(-\left(\exp\left(\beta_1'X\right)\Lambda_1(\tau)+\exp\left(\beta_2'X\right)\Lambda_2(\tau)\right)\right)\right) \\ & \times\exp\left(-\left(\exp\left(\beta_1'X\right)\Lambda_1(\tau)+\exp\left(\beta_2'X\right)\Lambda_2(\tau)\right)\right) \\ & \times\exp\left(\beta_2'X\right)\lambda_2(\tau)\,d\tau \end{aligned}$$

then

$$\begin{aligned} & h_0\left(\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_{0,1}(t)+\exp\left(\beta_{0,2}'X\right)\Lambda_{0,2}(t)\right)\right)\right) \\ & \times\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_{0,1}(t)+\exp\left(\beta_{0,2}'X\right)\Lambda_{0,2}(t)\right)\right) \\ & \times\exp\left(\beta_{0,2}'X\right)\lambda_{0,2}(t) \\ & = h\left(\exp\left(-\left(\exp\left(\beta_1'X\right)\Lambda_1(t)+\exp\left(\beta_2'X\right)\Lambda_2(t)\right)\right)\right) \\ & \times\exp\left(-\left(\exp\left(\beta_1'X\right)\Lambda_1(t)+\exp\left(\beta_2'X\right)\Lambda_2(t)\right)\right) \\ & \times\exp\left(\beta_2'X\right)\lambda_2(t) \text{ a.s.} \end{aligned} \tag{16}$$

5.2 The Weibull case

Suppose that $h_0(1) = h(1) = 1$, which corresponds to $E[V] = 1$. Then, letting $t \downarrow 0$, it follows from (15) that

$$\exp\left((\beta_{0,1} - \beta_1)' X\right) \lim_{t \downarrow 0} \frac{\lambda_{0,1}(t)}{\lambda_1(t)} = 1 \text{ a.s.} \tag{17}$$

If $\lambda_{0,1}(t)$ and $\lambda_1(t)$ are Weibull baseline hazards, including scale factors, i.e.,

$$\lambda_{0,1}(t) = \alpha^*_{1,1}\alpha^*_{1,2} t^{\alpha^*_{1,2}-1}, \ \lambda_1(t) = \alpha_{1,1}\alpha_{1,2} t^{\alpha_{1,2}-1}, \tag{18}$$

where $\alpha^*_{1,1}$ and $\alpha_{1,1}$ are the scale factors involved, and all the parameters involved are positive valued, then

$$\begin{aligned} \lim_{t \downarrow 0} \frac{\lambda_{0,1}(t)}{\lambda_1(t)} &= \frac{\alpha^*_{1,1}}{\alpha_{1,1}} \times \frac{\alpha^*_{1,2}}{\alpha_{1,2}} \lim_{t \downarrow 0} t^{\alpha^*_{1,2}-\alpha_{1,2}} \\ &= \begin{cases} 0 & if \quad \alpha^*_{1,2} > \alpha_{1,2}, \\ \frac{\alpha^*_{1,1}}{\alpha_{1,1}} & if \quad \alpha^*_{1,2} = \alpha_{1,2}, \\ \infty & if \quad \alpha^*_{1,2} < \alpha_{1,2}. \end{cases} \end{aligned}$$

so that by (17), $\alpha^*_{1,2} = \alpha_{1,2}$ and

$$X'(\beta_{0,1} - \beta_1) = \ln\left(\frac{\alpha_{1,1}}{\alpha^*_{1,1}}\right) \text{ a.s.} \tag{19}$$

Because of the presence of the scale factors $\alpha^*_{1,1}$ and $\alpha_{1,1}$, we cannot allows a constant in X:

Assumption 1. *The vector X of covariates does* ***not*** *contain a constant.*

Moreover, assume that

Assumption 2. *The variance matrix Σ_x of X is well defined*[1] *and is non-singular.*

[1] A sufficient condition for this is that $E[X'X] < \infty$.

Then it follows from (19) that

$$\begin{aligned}\Sigma_x(\beta_{0,1}-\beta_1) &= E\left[(X-E[X])(X-E[X])\right]'(\beta_{0,1}-\beta_1)\\ &= 0\end{aligned}$$

hence $\beta_{0,1}=\beta_1$, and thus $\alpha_1=\alpha_{0,1}$.

Next, suppose that

$$\lambda_1(t)=\frac{2\alpha_{1,1}t}{\alpha_{1,2}^2+t^2},\ \lambda_{0,1}(t)=\frac{2\alpha_{1,1}^*t}{\left(\alpha_{1,2}^*\right)^2+t^2}, \tag{20}$$

where again $\alpha_{1,1}^*$ and $\alpha_{1,1}$ are scale factors, and all the parameters are positive valued. These hazard functions are unimodal, with modes at $\alpha_{1,2}>0$ and $\alpha_{1,2}^*>0$, respectively. Then

$$\lim_{t\downarrow 0}\frac{\lambda_{0,1}(t)}{\lambda_1(t)}=\frac{\alpha_{1,1}^*}{\alpha_{1,1}}\times\frac{\alpha_{1,2}^2}{\left(\alpha_{1,2}^*\right)^2},$$

hence (19) now becomes

$$(\beta_{0,1}-\beta_1)'X=\ln\left(\left(\alpha_{1,2}^*\right)^2/\alpha_{1,2}^2\right)-\ln\left(\alpha_{1,1}/\alpha_{1,1}^*\right)\ \text{a.s.} \tag{21}$$

Under Assumptions 1-2 and the condition $h_0(1)=h(1)=1$, (21) still implies that $\beta_{0,1}=\beta_1$ but now only

$$\frac{\alpha_{1,2}^2}{\left(\alpha_{1,2}^*\right)^2}=\frac{\alpha_{1,1}}{\alpha_{1,1}^*}.$$

Thus, in the unimodal hazard case the condition $h_0(1)=h(1)=1$ does not guarantee identification of the parameters of the unimodal baseline hazard. It is easy to verify that the same problem occurs whenever

$$\lim_{t\downarrow 0}\lambda_{0,1}(t)/\lambda_1(t)\in(0,\infty)\backslash\{1\}$$

is possible.

On the other hand, the condition $h_0(1)=h(1)=1$ together with Assumptions 1-2 guarantee that in the case $D=1$, $\beta_{0,1}=\beta_1$, and similarly in the case $D=2$ that $\beta_{0,2}=\beta_2$. Therefore, we will maintain this condition:

Assumption 3. *The common unobserved heterogeneity distribution $G(v)$ satisfies $\int_0^\infty v dG(v) = 1$. This condition is implemented by confining the true density h_0 and the density h in the log-likelihood function to the space of density functions h on $(0,1]$ satisfying $h(1) = 1$.*

Thus, (15) now reads,

$$\begin{aligned} & h_0\left(\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_{0,1}(t)+\exp\left(\beta_{0,2}'X\right)\Lambda_{0,2}(t)\right)\right)\right) \\ & \times\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_{0,1}(t)+\exp\left(\beta_{0,2}'X\right)\Lambda_{0,2}(t)\right)\right) \\ & \times\lambda_{0,1}(t) \\ & = h\left(\exp\left(-\left(\exp\left(\beta_1'X\right)\Lambda_1(t)+\exp\left(\beta_2'X\right)\Lambda_2(t)\right)\right)\right) \\ & \times\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_1(t)+\exp\left(\beta_{0,2}'X\right)\Lambda_2(t)\right)\right) \\ & \times\lambda_1(t) \text{ a.s.} \end{aligned} \tag{22}$$

and (16) reads

$$\begin{aligned} & h_0\left(\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_{0,1}(t)+\exp\left(\beta_{0,2}'X\right)\Lambda_{0,2}(t)\right)\right)\right) \\ & \times\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_{0,1}(t)+\exp\left(\beta_{0,2}'X\right)\Lambda_{0,2}(t)\right)\right) \\ & \times\lambda_{0,2}(t) \\ & = h\left(\exp\left(-\left(\exp\left(\beta_1'X\right)\Lambda_1(t)+\exp\left(\beta_2'X\right)\Lambda_2(t)\right)\right)\right) \\ & \times\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_1(t)+\exp\left(\beta_{0,2}'X\right)\Lambda_2(t)\right)\right) \\ & \times\lambda_2(t) \text{ a.s.} \end{aligned} \tag{23}$$

5.3 Additional identification conditions

It follows from (22) and (23) that for all $t < \overline{T}$,

$$\frac{\lambda_2(t)}{\lambda_1(t)} = \frac{\lambda_{0,2}(t)}{\lambda_{0,1}(t)}. \tag{24}$$

To see what this result implies for the unimodal case, let similarly to (20),

$$\lambda_2(t) = \frac{2\alpha_{2,1}t}{\alpha_{2,2}^2 + t^2},\ \lambda_{0,2}(t) = \frac{2\alpha_{2,1}^* t}{\left(\alpha_{2,2}^*\right)^2 + t^2}, \tag{25}$$

and assume that $\alpha_{1,2}^* \neq \alpha_{2,2}^*$, so that $\lambda_{0,1}(t)$ and $\lambda_{0,2}(t)$ are not proportional. Then it follows easily from (24), (20) and (25) that $\alpha_{1,2} = \alpha_{1,2}^*$, which implies that $\lambda_1(t)$ and $\lambda_{0,1}(t)$ are proportional, and therefore $\lambda_2(t)$ and $\lambda_{0,2}(t)$ are proportional as well, with common proportionality factor $c > 0$, say:

$$\lambda_1(t) = c.\lambda_{0,1}(t), \quad \lambda_2(t) = c.\lambda_{0,2}(t). \tag{26}$$

But then it follows from (22) and Assumption 3 that

$$\begin{aligned} c &= \lim_{t\downarrow 0} \frac{h_0\left(\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_{0,1}(t) + \exp\left(\beta_{0,2}'X\right)\Lambda_{0,2}(t)\right)\right)\right)}{h\left(\exp\left(-\left(\exp\left(\beta_1'X\right)\Lambda_1(t) + \exp\left(\beta_2'X\right)\Lambda_2(t)\right)\right)\right)} \\ &\quad \times \frac{\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_{0,1}(t) + \exp\left(\beta_{0,2}'X\right)\Lambda_{0,2}(t)\right)\right)}{\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_1(t) + \exp\left(\beta_{0,2}'X\right)\Lambda_2(t)\right)\right)} \\ &= 1, \end{aligned}$$

hence $\lambda_1(t) = \lambda_{0,1}(t)$, $\lambda_2(t) = \lambda_{0,2}(t)$.

If $\alpha_{1,2}^* = \alpha_{2,2}^*$, so that $\lambda_{0,1}(t)$ and $\lambda_{0,2}(t)$ are proportional: $.\lambda_{0,2}(t) = \kappa\lambda_{0,1}(t)$, say, then (24) implies that $\lambda_2(t) = \kappa\lambda_1(t)$ as well, but not necessarily that (26) holds. Therefore, proportionality of $\lambda_{0,1}(t)$ and $\lambda_{0,2}(t)$ has to be excluded, at least for t close to zero.

Assumption 4. *The true baseline hazards $\lambda_{0,1}(t)$ and $\lambda_{0,2}(t)$ are non-proportional, in the sense that there exists a small $\varepsilon > 0$ such that for any constant $\kappa > 0$ the set $\{t \in (0,\varepsilon) : \lambda_{0,2}(t) = \kappa.\lambda_{0,1}(t)\}$ has Lebesgue measure zero.*

In general (24) and Assumption 4 are necessary but not sufficient conditions for (26), because we can always choose a hazards function $\lambda_1(t)$ such that $\lambda_2(t)$ defined by

$$\lambda_2(t) \equiv \left(\frac{\lambda_1(t)}{\lambda_{0,1}(t)}\right)\lambda_{0,2}(t). \tag{27}$$

is a valid hazard function. The reason that (26) holds for Weibull hazards and our unimodal hazards is that the four hazard functions involved have the same functional forms, which is such that (27) implies that $\lambda_1(t)$ and $\lambda_{0,1}(t)$ are proportional. Therefore, we need to require that:

Assumption 5. *The baseline hazard functions in our model belong to a class of parametric hazard functions $\mathcal{L} = \{\lambda(t|\alpha), \alpha \in A\}$ such that for any pair $\lambda_{0,1}, \lambda_{0,2}$ of non-proportional*[2] *hazard functions in $\mathcal{L}$, (27) can only hold for a pair $\lambda_1, \lambda_2 \in \mathcal{L}$ if and only if $\lambda_1(t) \equiv c.\lambda_{0,1}(t)$ for some constant $c > 0$.*

Summarizing, we have shown that

Theorem 1. *If the baseline hazards are of the Weibull type, i.e., $\lambda_i(t|\alpha_i) = \alpha_{i,1}\alpha_{i,2}t^{\alpha_{i,2}-1}$ for $i = 1, 2$, with $\alpha_i = (\alpha_{i,1}, \alpha_{i,2})'$, $\alpha_{i,1} > 0$, $\alpha_{i,2} > 0$, then the parameters of our model are identified under Assumptions 1-3. For other types of baseline hazards the parameters are identified under Assumptions 1-5.*

6 Specifying the unobserved heterogeneity distribution

6.1 Flexible functional form

Under Assumptions 1-5 it follows now from (15) that for $t \leq \overline{T}$,

$$\begin{aligned} &h_0\left(\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_{0,1}(t) + \exp\left(\beta_{0,2}'X\right)\Lambda_{0,2}(t)\right)\right)\right) \\ &= h\left(\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_{0,1}(t) + \exp\left(\beta_{0,2}'X\right)\Lambda_{0,2}(t)\right)\right)\right) \end{aligned}$$

a.s. By a similar argument it can be shown that

$$\begin{aligned} &H_0\left(\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_{0,1}\left(\overline{T}\right) + \exp\left(\beta_{0,2}'X\right)\Lambda_{0,2}\left(\overline{T}\right)\right)\right)\right) \\ &= H\left(\exp\left(-\left(\exp\left(\beta_{0,1}'X\right)\Lambda_{0,1}\left(\overline{T}\right) + \exp\left(\beta_{0,2}'X\right)\Lambda_{0,2}\left(\overline{T}\right)\right)\right)\right) \end{aligned}$$

a.s. Therefore, in the case of right-censoring it may not be true that $h(u) = h_0(u)$ a.e. on $(0, 1]$, except if for all $x > 0$,

$$P\left[\exp\left(\beta_{0,1}'X\right)\Lambda_{0,1}\left(\overline{T}\right) + \exp\left(\beta_{0,2}'X\right)\Lambda_{0,2}\left(\overline{T}\right) > x\right] > 0.$$

[2] As defined in Assumption 4.

This is not too serious a problem, though, because we will model h_0 in a flexible way, but involving only a finite number of parameters, similar to the approach in Bierens (2007).

In particular, we will assume that $h_0(u)$ belongs to the space of density functions of the type

$$h_q(u) = h_q(u|\delta) = \frac{\left(1 + \sum_{k=1}^{q} \delta_k \rho_k(u)\right)^2}{1 + \sum_{k=1}^{q} \delta_k^2}, \ \delta = (\delta_1, ..., \delta_q)', \tag{28}$$

for a given **fixed** natural number q, where the $\rho_k(u)$'s are orthonormal Legendre polynomials on the unit interval:

$$\int_0^1 \rho_k(u)\rho_m(u)du = \begin{cases} 0 & if \quad k \neq m, \\ 1 & if \quad k = m. \end{cases} \tag{29}$$

Thus, it will be assumed that $h_0(u) \equiv h_q(u|\delta_0)$ for some unique $\delta_0 \in \mathbb{R}^q$.

The Legendre polynomials can easily be generated recursively by

$$\rho_n(u) = \frac{\sqrt{2n-1}\sqrt{2n+1}}{n}(2u-1)\rho_{n-1}(u) - \frac{(n-1)\sqrt{2n+1}}{n\sqrt{2n-3}}\rho_{n-2}(u), \tag{30}$$

for $n \geq 2$, starting from

$$\rho_0(u) = 1, \ \rho_1(u) = \sqrt{3}\,(2u-1)\,. \tag{31}$$

The condition $h_q(1|\delta) = 1$ can be imposed by restricting δ_1 to be

$$\begin{aligned} \delta_1 &= \frac{1}{2}\sqrt{2\left(1 + \sum_{k=2}^{q} \delta_k^2\right) + \left(1 + \sum_{k=2}^{q} \delta_k \rho_k(1)\right)^2} \\ &\quad -\frac{\sqrt{3}}{2}\left(1 + \sum_{k=2}^{q} \delta_k \rho_k(1)\right) \\ &= \frac{1}{2}\sqrt{2\left(1 + \sum_{k=2}^{q} \delta_k^2\right) + \left(1 + \sum_{k=2}^{q} \delta_k \sqrt{2k+1}\right)^2} \\ &\quad -\frac{\sqrt{3}}{2}\left(1 + \sum_{k=2}^{q} \delta_k \sqrt{2k+1}\right), \end{aligned} \tag{32}$$

where the latter equality follows from the fact that $\rho_k(1) = \sqrt{2k+1}$. See Bierens (2007) for further details and the motivation for using densities of the type (28).

6.2 The Gamma distribution

A popular specification of the unobserved heterogeneity distribution $G(v)$ is the Gamma(δ,τ) distribution, because the Laplace transform of the Gamma(δ,τ) distribution has a closed form:

$$\mathcal{L}(s) = E[\exp(-s.V)] = \int_0^\infty \exp(-s.v)dG(v) = (1+\tau.s)^{-\delta},$$

which is related to the distribution function $H(u) = \int_0^\infty u^v dG(v)$ by the equality

$$\mathcal{L}(s) = H(\exp(-s)) = (1+\tau.s)^{-\delta}.$$

Hence

$$\begin{aligned} H(u) &= (1+\tau.\ln(1/u))^{-\delta} \\ h(u) &= H'(u) = \frac{\tau.\delta}{u}\left(\frac{1}{1+\tau.\ln(1/u)}\right)^{\delta+1} = \frac{\tau.\delta}{u}H(u)^{(\delta+1)/\delta}. \end{aligned}$$

Because of the presence of scale factors in the Weibull baseline hazards, the parameter τ has to be fixed to a constant, or made dependent on δ. To facilitate the comparison with the previous flexible specification, it will be assumed that $\tau = 1/\delta$, so that in this case $h(1) = 1$ as well. Thus

$$H(u|\delta) = \left(1+\delta^{-1}.\ln(1/u)\right)^{-\delta},\ h(u|\delta) = \frac{1}{u}H(u|\delta)^{(\delta+1)/\delta}. \tag{33}$$

7 Implementation in EasyReg

In EasyReg module SNPSURVIVAL2 the following options for the baseline and integrated hazards of T_1 and T_2 are available. Because of Assumption 5, each option applies to **both** T_1 and T_2, although with different parameter values.

7.1 Weibull hazard

$$\lambda(t|\alpha) = \alpha_1\alpha_2 t^{\alpha_2-1}, \tag{34}$$

$$\alpha_1 > 0,\ \alpha_2 > 0,\ \alpha = (\alpha_1,\alpha_2)'.$$

Integrated hazard:

$$\Lambda(t|\alpha) = \int_0^t \lambda(\tau|\alpha)d\tau = \alpha_1 t^{\alpha_2}.$$

7.2 Generalized Weibull hazard

If in the Weibull case $\alpha_2 < 1$ then $\lambda(0|\alpha) = \infty$, whereas if $\alpha_2 > 1$ then $\lambda(0|\alpha) = 0$. This may be too restrictive. The following generalized Weibull hazard specification satisfies $0 < \lambda(0|\alpha) < \infty$:

$$\begin{aligned}
&\lambda(t|\alpha) = \alpha_1\alpha_2\left(\alpha_3 + t\right)^{\alpha_2 - 1}, \\
&\alpha_1 > 0,\ \alpha_2 > 0,\ \alpha_3 > 0,\ \alpha = (\alpha_1, \alpha_2, \alpha_3)' \\
&\text{Integrated hazard:} \\
&\Lambda(t|\alpha) = \int_0^t \lambda(\tau|\alpha)d\tau = \alpha_1\left((\alpha_3 + t)^{\alpha_2} - \alpha_3^{\alpha_2}\right).
\end{aligned}$$

7.3 Unimodal hazard

$$\begin{aligned}
&\lambda(t|\alpha) = \frac{2\alpha_1 t}{\alpha_2^2 + t^2},\ \alpha_2 = \arg\max_{t\geq 0}\lambda(t|\alpha), \\
&\alpha_1 > 0,\ \alpha_2 > 0,\ \alpha = (\alpha_1, \alpha_2)'. \\
&\text{Integrated hazard:} \\
&\Lambda(t|\alpha) = \int_0^t \lambda(\tau|\alpha)d\tau = \alpha_1.\ln\left(\frac{\alpha_2^2 + t^2}{\alpha_2^2}\right).
\end{aligned}$$

7.4 Generalized unimodal hazard

In the unimodal hazard case, $\lambda(0|\alpha) = 0$. Again, this may be too restrictive. The following generalized unimodal hazard specification allows $\lambda(0|\alpha) > 0$:

$$\begin{aligned}
&\lambda(t|\alpha) = \frac{2\alpha_1\left(\alpha_3 + t\right)}{\left(\alpha_2 + \alpha_3\right)^2 + \left(\alpha_3 + t\right)^2},\ \alpha_2 = \arg\max_{t\geq 0}\lambda(t|\alpha), \\
&\alpha_1 > 0,\ \alpha_2 > 0,\ \alpha_3 > 0,\ \alpha = (\alpha_1, \alpha_2, \alpha_3)'. \\
&\text{Integrated hazard:} \\
&\Lambda(t|\alpha) = \int_0^t \lambda(\tau|\alpha)d\tau = \alpha_1.\ln\left(\frac{\left(\alpha_2 + \alpha_3\right)^2 + \left(\alpha_3 + t\right)^2}{\left(\alpha_2 + \alpha_3\right)^2 + \alpha_3^2}\right).
\end{aligned}$$

In all four cases the parameter α_1 acts as a scale factor. Therefore, $\ln(\alpha_1)$ acts as a constant term. Consequently, the vector X of covariates should **not** contain a constant.

7.5 Other options

EasyReg will ask you to select $T = \min(T_1, T_2)$, the dummy variable $D = 1 + I(T_2 < T_1)$, the right-censoring dummy variable C, and the covariates X. If the selected dummy variable for D does not take the values 1 or 2, EasyReg will transform this variable such that it does.

Moreover, for the density $h(u) = \int_0^\infty vu^{v-1}dG(v)$ you can choose either (28) or (33)

8 Three-step ML estimation

EasyReg estimates the parameter vectors α_1, α_2, β_1, β_2 and δ in three steps. In first instance the α'a are fixed to 1, and δ is set equal to a zero vector

The (quasi-)maximum likelihood estimators $\widetilde{\beta}_{i,0}$ of β_i for $i = 1, 2$ in the first step will be used as starting values in the second step, together with the initial values of the α's, keeping $\delta = 0$. This step yields (quasi-)maximum likelihood estimators $\widetilde{\alpha}_i$ of α_i and $\widetilde{\beta}_i$ of β_i for $i = 1, 2$.

Again, these estimates are merely used as starting values in the final step, where $h(u)$ and $H(u)$ are approximated by $h_q(u|\delta)$ and $H_q(u|\delta)$, respectively.

If you check "Batch mode" these three rounds are conducted automatically, where in each round the iteration is automatically restarted until the log-likelihood does not change anymore. This option is recommended for big jobs.

References

Bierens, H. J. (2007), "Semi-Nonparametric Interval Censored Mixed Proportional Hazard Models: Identification and Consistency Results", forthcoming in *Econometric Theory*

Bierens, H. J., and J. R. Carvalho (2007), "Semi-Nonparametric Competing Risks Analysis of Recidivism", forthcoming in the *Journal of Applied Econometrics*.