Survival analysis II

Final Report

Name: Pan, Chi-Hung

Download the data from Bioconductor curated OvarianData package.

Reproduce table2

Dataset	Median follow-up (month)	Sample size	The number of obs. Event (event rate%)		
			Relapse $\delta_{ij}=1$	Death $\delta^*_{ij}=1$	Censor $\delta^*_{ij}=0$
GSE17260		$N_1=110$	76 (69%)	46 (42%)	64 (58%)
GSE30161		$N_2=58$	48 (83%)	36 (62%)	22 (38%)
GSE9891		$N_3=278$	185 (68%)	113 (41%)	165 (59%)
TCGA		$N_4=557$	266 (48%)	290(52%)	267 (48%)
Total		$\sum_{i=1}^{4} N_4=1003$	575 (57%)	485 (48%)	518(52%)

Using joint.Cox package. This package produce estimators of β_1 and β_2. Let $g(x)=\exp(x)$, then $g'(x)=\exp(x)$. By invariance property, we estimate $\exp(\beta_1)$ and $\exp(\beta_2)$ by $g(\beta_2)$ and $g(\beta_1)$. By the delta method, we obtaine the SE of $g(\beta_i)$.

Reproduce table5

	Proposed method: Estimate (95% CI)	Method of Rondeau et al.: Estimate (95% CI)
RR for relapse (TTP): $\exp(\beta_1)$	1.22 (1.129-1.312)	1.24 (1.135-1.343)
RR for death (OB): $\exp(\beta_2)$	1.18 (1.079-1.281)	1.17 (1.062-1.280)
Heterogeneity: $\eta=Var_\eta(u_i)$	0.033	0.028
Copula parameter: θ	2.35	0.00 (fixed)
RR for death after relapse: $\theta+1$	3.35	1.00 (fixed)
Kendall's tau: $\theta/(\theta+2)$	0.54	0.00 (fixed)
Maximum penalized log-likelihood	-8604.093	8744.023

Joint frailty-copula model

Defined

$$\begin{cases} r_{ij}(t \mid u_i) = u_i r_0 \exp(\beta_1' Z_{ij}) \\ \lambda_{ij}(t \mid u_i) = u_i^{\alpha} \lambda_0 \exp(\beta_2' Z_{ij}) \end{cases} \tag{1}$$

$$\Pr(X_{ij} > x, D_{ij} > y \mid u_i) = C_\theta[\exp\{-R_{ij}(x \mid u_i)\}, \exp\{-\Lambda_{ij}(y \mid u_i)\}], \tag{2}$$

where C_θ is a copula with unknown parameter θ, and

$$R_{ij}(x \mid u_i) = \int_0^x r_{ij}(v \mid u_i) dv = \int_0^x u_i r_0(v) \exp(\beta_1' Z_{ij}) dv = u_i R_0(x) \exp(\beta_1' Z_{ij}) = u_i R_{ij}(x),$$

$$\Lambda_{ij}(y \mid u_i) = \int_0^x \lambda_{ij}(v \mid u_i) dv = \int_0^x u_i^{\alpha} \lambda_0(v) \exp(\beta_2' Z_{ij}) dv = u_i^{\alpha} \Lambda_0(y) \exp(\beta_2' Z_{ij}) = u_i^{\alpha} \Lambda_{ij}(y),$$

where

$$R_0(x) = \int_0^x r_0(v) dv,$$

$$\Lambda_0(y) = \int_0^x \lambda_0(v) dv.$$

By previous, the distribution is

$$\begin{aligned} \overline{G}(x, y \mid u) = \Pr(X_{ij} > x, D_{ij} > y \mid u_i) &= C_\theta[\exp\{-R_{ij}(x \mid u_i)\}, \exp\{-\Lambda_{ij}(y \mid u_i)\}] \\ &= D_\theta[R_{ij}(x \mid u_i), \Lambda_{ij}(y \mid u_i)] \\ &= D_\theta[u_i R_{ij}(x), u_i^{\alpha} \Lambda_{ij}(y)] \end{aligned}$$

$$\begin{aligned} -\frac{\partial \overline{G}(x, y \mid u)}{\partial x} &= -\frac{\partial D_\theta[u_i R_{ij}(x), u_i^{\alpha} \Lambda_{ij}(y)]}{\partial x} = D_\theta^{[1,0]}[u_i R_{ij}(x), u_i^{\alpha} \Lambda_{ij}(y)] u_i \frac{\partial R_{ij}(x)}{\partial x} \\ &= D_\theta^{[1,0]}[u_i R_{ij}(x), u_i^{\alpha} \Lambda_{ij}(y)] u_i r_{ij}(x), \end{aligned}$$

where $D_\theta^{[1,0]} = -\partial D_\theta(s,t) / \partial s$.

Similarly,

$$-\frac{\partial \overline{G}(x, y \mid u)}{\partial y} = -\frac{\partial D_\theta[u_i R_{ij}(x), u_i^{\alpha} \Lambda_{ij}(y)]}{\partial y} = D_\theta^{[0,1]}[u_i R_{ij}(x), u_i^{\alpha} \Lambda_{ij}(y)] u_i^{\alpha} \lambda_{ij}(y),$$

where $D_\theta^{[0,1]} = -\partial D_\theta(s,t) / \partial t$.

$$\frac{\partial^2 \overline{G}(x, y \mid u)}{\partial x \partial y} = D_\theta^{[1,1]}[u_i R_{ij}(x), u_i^{\alpha} \Lambda_{ij}(y)] u_i^{\alpha+1} r_{ij}(x) \lambda_{ij}(y),$$

The likelihood of a single data $(T_{ij}, T^*_{ij}, \delta_{ij}, \delta^*_{ij})$ given u_i under the joint frailty-copula model

$$L_j(\boldsymbol{\theta} \mid u_i) = f(T_{ij}, T^*_{ij} \mid u_i)^{\delta_{ij}\delta^*_{ij}} \left[-\frac{\partial \overline{G}(v, T^*_{ij} \mid u_i)}{\partial v} \bigg|_{v=T_{ij}} \right]^{\delta_{ij}(1-\delta^*_{ij})}$$

$$\times \left[-\frac{\partial \overline{G}(T_{ij}, w \mid u_i)}{\partial w} \bigg|_{w=T^*_{ij}} \right]^{\delta^*_{ij}(1-\delta_{ij})} \overline{G}(T^*_{ij}, T^*_{ij} \mid u_i)^{(1-\delta^*_{ij})(1-\delta_{ij})}$$

$$= \{ D^{[1,1]}_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})] u_i^{\alpha+1} r_{ij}(T_{ij}) \lambda_{ij}(T^*_{ij}) \}^{\delta_{ij}\delta^*_{ij}} \{ D^{[1,0]}_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})] u_i r_{ij}(T_{ij}) \}^{\delta_{ij}(1-\delta^*_{ij})}$$

$$\times \{ D^{[0,1]}_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})] u_i^\alpha \lambda_{ij}(T^*_{ij}) \}^{\delta^*_{ij}(1-\delta_{ij})} \{ D_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})] \}^{(1-\delta^*_{ij})(1-\delta_{ij})},$$

where $\boldsymbol{\theta} = (\alpha, \eta, \theta, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, r_0, \lambda_0)$.

Simplify the likelihood function as follows

$$L_j(\boldsymbol{\theta} \mid u_i)$$

$$= \{ D^{[1,1]}_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})] u_i^{\alpha+1} r_{ij}(T_{ij}) \lambda_{ij}(T^*_{ij}) \}^{\delta_{ij}\delta^*_{ij}} \{ D^{[1,0]}_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})] u_i r_{ij}(T_{ij}) \}^{\delta_{ij}(1-\delta^*_{ij})}$$

$$\times \{ D^{[0,1]}_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})] u_i^\alpha \lambda_{ij}(T^*_{ij}) \}^{\delta^*_{ij}(1-\delta_{ij})} \{ D_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})] \}^{(1-\delta^*_{ij})(1-\delta_{ij})}$$

$$= \left\{ \frac{D^{[1,1]}_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})] D_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})]}{D^{[1,0]}_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})] D^{[0,1]}_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})]} \right\}^{\delta_{ij}\delta^*_{ij}} \left\{ \frac{D^{[1,0]}_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})]}{D_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})]} \right\}^{\delta_{ij}}$$

$$\times \left\{ \frac{D^{[0,1]}_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})]}{D_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})]} \right\}^{\delta^*_{ij}} D_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})]$$

$$\times \{ u_i^{\alpha+1} r_{ij}(T_{ij}) \}^{\delta_{ij}\delta^*_{ij}} \{ u_i r_{ij}(T_{ij}) \}^{\delta_{ij}(1-\delta^*_{ij})} \{ u_i^\alpha \lambda_{ij}(T^*_{ij}) \}^{\delta^*_{ij}(1-\delta_{ij})}$$

$$= \{ \Theta_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})] \}^{\delta_{ij}\delta^*_{ij}} \{ \psi_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})] \}^{\delta_{ij}} \{ \psi^*_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})] \}^{\delta^*_{ij}}$$

$$\times D_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T^*_{ij})] \{ u_i^{\delta_{ij}+\alpha\delta^*_{ij}} r_{ij}(T_{ij})^{\delta_{ij}} \lambda_{ij}(T^*_{ij})^{\delta^*_{ij}} \},$$

where $\Theta_\theta(s,t) = \dfrac{D^{[1,1]}_\theta(s,t) D_\theta(s,t)}{D^{[1,0]}_\theta(s,t) D^{[0,1]}_\theta(s,t)}$, $\psi_\theta(s,t) = \dfrac{D^{[1,0]}_\theta(s,t)}{D_\theta(s,t)}$, $\psi^*_\theta(s,t) = \dfrac{D^{[0,1]}_\theta(s,t)}{D_\theta(s,t)}$.

Because u_i is unknown, take expectation with respect to u_i.

Let the density of u_i is $f_\eta(u_i)$.

Then the likelihood function based on the data $\{ (T_{ij}, T_{ij}^*, \delta_{ij}, \delta_{ij}^*) \}, i = 1, 2, ..., G$ and $j = 1, 2, ..., N_i$ is

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{G} \int_0^\infty \prod_{j=1}^{N_i} L_j(\boldsymbol{\theta} \,|\, u_i) f(u_i) du_i$$

$$= \prod_{i=1}^{G} \int_0^\infty \prod_{j=1}^{N_i} r_{ij}(T_{ij})^{\delta_{ij}} \lambda_{ij}(T_{ij}^*)^{\delta_{ij}^*} u_i^{\delta_{ij}+\alpha\delta_{ij}^*} \{ \Theta_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}\delta_{ij}^*} \{ \psi_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}}$$
$$\times \{ \psi_\theta^*[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}^*} D_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] f(u_i)\, du_i$$

The log-likelihood function is [Equation(4)]

$$l(\boldsymbol{\theta}) = \log L(\boldsymbol{\theta})$$

$$= \log \prod_{i=1}^{G} \int_0^\infty \prod_{j=1}^{N_i} r_{ij}(T_{ij})^{\delta_{ij}} \lambda_{ij}(T_{ij}^*)^{\delta_{ij}^*} u_i^{\delta_{ij}+\alpha\delta_{ij}^*} \{ \Theta_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}\delta_{ij}^*}$$
$$\times \{ \psi_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}} \{ \psi_\theta^*[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}^*}$$
$$\times D_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] f(u_i)\, du_i$$

$$= \log \Bigg[\prod_{i=1}^{G} \prod_{j=1}^{N_i} \{ r_{ij}(T_{ij})^{\delta_{ij}} \lambda_{ij}(T_{ij}^*)^{\delta_{ij}^*} \} \prod_{i=1}^{G} \Bigg\{ \int_0^\infty u_i^{\sum_j \delta_{ij}+\alpha\delta_{ij}^*} \prod_{j=1}^{N_i} (\{ \Theta_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}\delta_{ij}^*}$$
$$\times \{ \psi_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}} \{ \psi_\theta^*[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}^*}$$
$$\times D_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] f(u_i)) \Bigg\} du_i \Bigg]$$

$$= \sum_{i=1}^{G} \sum_{j=1}^{N_i} \{ \delta_{ij} \log r_{ij}(T_{ij}) + \delta_{ij}^* \log \lambda_{ij}(T_{ij}^*) \}$$

$$+ \log \prod_{i=1}^{G} \int_0^\infty u_i^{\sum_j \delta_{ij}+\alpha\delta_{ij}^*} \prod_{j=1}^{N_i} (\{ \Theta_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}\delta_{ij}^*}$$
$$\times \{ \psi_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}} \{ \psi_\theta^*[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}^*}$$
$$\times D_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] f(u_i)) du_i$$

$$= \sum_{i=1}^{G} \sum_{j=1}^{N_i} \{ \delta_{ij} \log r_{ij}(T_{ij}) + \delta_{ij}^* \log \lambda_{ij}(T_{ij}^*) \}$$

$$+ \sum_{i=1}^{G} \log \int_0^\infty u_i^{\sum_j \delta_{ij}+\alpha\delta_{ij}^*} \prod_{j=1}^{N_i} \{ \Theta_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}\delta_{ij}^*} \{ \psi_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}}$$
$$\times \{ \psi_\theta^*[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}^*} D_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] f_\eta(u_i) \} du_i$$

$$
\begin{aligned}
l(\boldsymbol{\theta}) = \sum_{i=1}^{G} \Bigg[& \sum_{j=1}^{N_i} \delta_{ij} \log r_{ij}(T_{ij}) + \delta_{ij}^{*} \log \lambda_{ij}(T_{ij}^{*}) \\
& + \log \int_0^{\infty} u_i^{m_i + \alpha m_i^{*}} \prod_{j=1}^{N_i} \{ \Theta_\theta [u_i R_{ij}(T_{ij}), u_i^{\alpha} \Lambda_{ij}(T_{ij}^{*})] \}^{\delta_{ij}\delta_{ij}^{*}} \{ \psi_\theta [u_i R_{ij}(T_{ij}), u_i^{\alpha} \Lambda_{ij}(T_{ij}^{*})] \}^{\delta_{ij}} \\
& \times \{ \psi_\theta^{*} [u_i R_{ij}(T_{ij}), u_i^{\alpha} \Lambda_{ij}(T_{ij}^{*})] \}^{\delta_{ij}^{*}} D_\theta [u_i R_{ij}(T_{ij}), u_i^{\alpha} \Lambda_{ij}(T_{ij}^{*})] f_\eta(u_i) \} du_i \Bigg],
\end{aligned}
$$

where $m_i = \sum_j \delta_{ij}, m_i^{*} = \sum_j \delta_{ij}^{*}$.

In the left-truncation case, only available samples when $L_{ij} \le T_{ij}$. Hence, the observed data are $(L_{ij}, T_{ij}, T_{ij}^{*}, \delta_{ij}, \delta_{ij}^{*})$, subject to $L_{ij} \le T_{ij}$, for $j = 1, 2, .., N_i$ and $i = 1, 2, .., G$.

$$
\Pr(T_{ij} \ge L_{ij}, T_{ij}^{*} \ge L_{ij} \mid u_i) = D_\theta [u_i R_{ij}(L_{ij}), u_i^{\alpha} \Lambda_{ij}(L_{ij})].
$$

Similarly, u_i is unknown, take expectation with respect to u_i.

$$
\Pr(T_{ij} \ge L_{ij}, T_{ij}^{*} \ge L_{ij}) = \int_0^{\infty} \prod_{j=1}^{N_i} \{ D_\theta [u_i R_{ij}(L_{ij}), u_i^{\alpha} \Lambda_{ij}(L_{ij})] f(u_i) \} du_i
$$

The likelihood function needs to modify as follows:

$$
\begin{aligned}
L^{*}(\boldsymbol{\theta}) = \prod_{i=1}^{G} \Bigg\{ & \left(\int_0^{\infty} \prod_{j=1}^{N_i} \{ D_\theta [u_i R_{ij}(L_{ij}), u_i^{\alpha} \Lambda_{ij}(L_{ij})] f(u_i) \} du_i \right)^{-1} \\
& \times \int_0^{\infty} \prod_{j=1}^{N_i} r_{ij}(T_{ij})^{\delta_{ij}} \lambda_{ij}(T_{ij}^{*})^{\delta_{ij}^{*}} u_i^{\delta_{ij} + \alpha \delta_{ij}^{*}} \{ \Theta_\theta [u_i R_{ij}(T_{ij}), u_i^{\alpha} \Lambda_{ij}(T_{ij}^{*})] \}^{\delta_{ij}\delta_{ij}^{*}} \{ \psi_\theta [u_i R_{ij}(T_{ij}), u_i^{\alpha} \Lambda_{ij}(T_{ij}^{*})] \}^{\delta_{ij}} \\
& \times \{ \psi_\theta^{*} [u_i R_{ij}(T_{ij}), u_i^{\alpha} \Lambda_{ij}(T_{ij}^{*})] \}^{\delta_{ij}^{*}} D_\theta [u_i R_{ij}(T_{ij}), u_i^{\alpha} \Lambda_{ij}(T_{ij}^{*})] f(u_i) \, du_i \Bigg\}
\end{aligned}
$$

The log-likelihood function as follows:

$$
\begin{aligned}
l^*(\boldsymbol{\theta}) &= \log L^*(\boldsymbol{\theta}) \\
&= \log \prod_{i=1}^{G} \left\{ \left(\int_0^\infty \prod_{j=1}^{N_i} \{ D_\theta[u_i R_{ij}(L_{ij}), u_i^\alpha \Lambda_{ij}(L_{ij})] f(u_i) \} \, du_i \right)^{-1} \right. \\
&\times \int_0^\infty \prod_{j=1}^{N_i} r_{ij}(T_{ij})^{\delta_{ij}} \lambda_{ij}(T_{ij}^*)^{\delta_{ij}^*} u_i^{\delta_{ij}+\alpha\delta_{ij}^*} \{ \Theta_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}\delta_{ij}^*} \{ \psi_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}} \\
&\left. \times \{ \psi_\theta^*[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}^*} D_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] f(u_i) \, du_i \right\} \\
&= -\sum_{i=1}^{G} \log \left(\int_0^\infty \prod_{j=1}^{N_i} \{ D_\theta[u_i R_{ij}(L_{ij}), u_i^\alpha \Lambda_{ij}(L_{ij})] f(u_i) \} \, du_i \right) \\
&+ \sum_{i=1}^{G} \log \int_0^\infty \prod_{j=1}^{N_i} r_{ij}(T_{ij})^{\delta_{ij}} \lambda_{ij}(T_{ij}^*)^{\delta_{ij}^*} u_i^{\delta_{ij}+\alpha\delta_{ij}^*} \{ \Theta_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}\delta_{ij}^*} \{ \psi_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}} \\
&\left. \times \{ \psi_\theta^*[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}^*} D_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] f(u_i) \, du_i \right\} \\
&= -\sum_{i=1}^{G} \log \left(\int_0^\infty \prod_{j=1}^{N_i} \{ D_\theta[u_i R_{ij}(L_{ij}), u_i^\alpha \Lambda_{ij}(L_{ij})] f(u_i) \} \, du_i \right) + l(\boldsymbol{\theta}) \\
&= \sum_{i=1}^{G} \left[\sum_{j=1}^{N_i} \delta_{ij} \log r_{ij}(T_{ij}) + \delta_{ij}^* \log \lambda_{ij}(T_{ij}^*) \right. \\
&\qquad + \log \int_0^\infty u_i^{m_i+\alpha m_i^*} \prod_{j=1}^{N_i} \{ \Theta_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}\delta_{ij}^*} \{ \psi_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}} \\
&\qquad\qquad \times \{ \psi_\theta^*[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] \}^{\delta_{ij}^*} D_\theta[u_i R_{ij}(T_{ij}), u_i^\alpha \Lambda_{ij}(T_{ij}^*)] f_\eta(u_i) \} du_i \\
&\qquad\qquad \left. - \log \left\{ \prod_{j=1}^{N_i} \int_0^\infty D_\theta[u_i R_{ij}(L_{ij}), u_i^\alpha \Lambda_{ij}(L_{ij})] f_\eta(u_i) du_i \right\} \right].
\end{aligned}
$$

Code

```
####Reproduce    table2
data(GSE9891_eset)
GSE9891_eset
data(GSE17260_eset)
GSE17260_eset
data(GSE30161_eset)
GSE30161_eset
data(TCGA_eset)
TCGA_eset

TTT = (is.na(pData(GSE17260_eset)$days_to_death)==1 &
is.na(pData(GSE17260_eset)$days_to_tumor_recurrence)==1)
TTT=1-TTT
sum(pData(GSE17260_eset)$recurrence_status=="recurrence" & TTT==1,na.rm =
T)
sum(pData(GSE17260_eset)$vital_status=="deceased"& TTT==1)
sum(pData(GSE17260_eset)$vital_status=="living"& TTT==1)
median(pData(GSE17260_eset)$days_to_death[TTT]/30)

TTT = (is.na(pData(GSE30161_eset)$days_to_death)==1 &
is.na(pData(GSE30161_eset)$days_to_tumor_recurrence)==1)
TTT=1-TTT
sum(pData(GSE30161_eset)$recurrence_status=="recurrence"& TTT==1,na.rm = T)
sum(pData(GSE30161_eset)$vital_status=="deceased"& TTT==1)
sum(pData(GSE30161_eset)$vital_status=="living"& TTT==1)

TTT = (is.na(pData(GSE9891_eset)$days_to_death)==1 &
is.na(pData(GSE9891_eset)$days_to_tumor_recurrence)==1)
TTT=1-TTT
ee=sum(is.na(pData(GSE9891_eset)$days_to_tumor_recurrence)==1&
      is.na(pData(GSE9891_eset)$days_to_death)==0 &
      pData(GSE9891_eset)$recurrence_status=="recurrence")
sum(pData(GSE9891_eset)$recurrence_status=="recurrence" & TTT==1,na.rm = T)
sum(pData(GSE9891_eset)$vital_status=="deceased"& TTT==1)
sum(pData(GSE9891_eset)$vital_status=="living"& TTT==1)
```

```
TTT = (is.na(pData(TCGA_eset)$days_to_death)==1 &
is.na(pData(TCGA_eset)$days_to_tumor_recurrence)==1)
TTT=1-TTT
ee=sum(is.na(pData(TCGA_eset)$days_to_tumor_recurrence)== 1&
            is.na(pData(TCGA_eset)$days_to_death)== 0 &
            pData(TCGA_eset)$recurrence_status=="recurrence")

sum(pData(TCGA_eset)$recurrence_status=="recurrence" & TTT==1 ,na.rm = T)-ee
sum(pData(TCGA_eset)$vital_status=="deceased"& TTT==1)
sum(pData(TCGA_eset)$vital_status=="living"& TTT==1)

jointCox.reg(t.event=dataOvarian$t.event, event=dataOvarian$event,
t.death=dataOvarian$t.death, death=dataOvarian$death,
  Z1=dataOvarian$CXCL12,Z2=dataOvarian$CXCL12,
group=dataOvarian$group,kappa_grid = c(seq(10, 1e+17, length = 30)))

####Reproduce    table5
## proposed method
Estimate_dep = jointCox.reg(t.event=dataOvarian$t.event, event=dataOvarian$event,
t.death=dataOvarian$t.death, death=dataOvarian$death,
  Z1=dataOvarian$CXCL12,Z2=dataOvarian$CXCL12,
group=dataOvarian$group,alpha = 0,kappa_grid = c(seq(10, 1e+17, length = 30)))

exp_beta1=round(exp(Estimate_dep$beta1[1]),2)
se_exp = exp(Estimate_dep$beta1[1])*Estimate_dep$beta1[2]
LB=round(exp(Estimate_dep$beta1[1])-qnorm(0.975)*se_exp,3)
UB=round(exp(Estimate_dep$beta1[1])+qnorm(0.975)*se_exp,3)
##exp(beta1)
exp_beta1
##CI for beta1
c(LB,UB)

exp_beta2=round(exp(Estimate_dep$beta2[1]),2)
se_exp = exp(Estimate_dep$beta2[1])*Estimate_dep$beta2[2]
LB=round(exp(Estimate_dep$beta2[1])-qnorm(0.975)*se_exp,3)
UB=round(exp(Estimate_dep$beta2[1])+qnorm(0.975)*se_exp,3)
```

```
##exp(beta2)
exp_beta2
##CI for beta1
c(LB,UB)
##eta theta tau Maximum penalized log-likelihood
round(Estimate_dep$eta[1],3)
round(Estimate_dep$theta[1],2)
round(Estimate_dep$theta[1]+1,2)
round(Estimate_dep$tau[1],2)
round(Estimate_dep$ML[1],3)
#### independent
Estimate_indep = jointCox.indep.reg(t.event=dataOvarian$t.event,
event=dataOvarian$event, t.death=dataOvarian$t.death, death=dataOvarian$death,
  Z1=dataOvarian$CXCL12,Z2=dataOvarian$CXCL12,
group=dataOvarian$group,alpha = 0,kappa_grid = c(seq(10, 1e+17, length = 30)))

exp_beta1=round(exp(Estimate_indep$beta1[1]),2)
se_exp = exp(Estimate_indep$beta1[1])*Estimate_indep$beta1[2]
LB=round(exp(Estimate_indep$beta1[1])-qnorm(0.975)*se_exp,3)
UB=round(exp(Estimate_indep$beta1[1])+qnorm(0.975)*se_exp,3)
##exp(beta1)
exp_beta1
##CI for beta1
c(LB,UB)
exp_beta2=round(exp(Estimate_indep$beta2[1]),2)
se_exp = exp(Estimate_indep$beta2[1])*Estimate_indep$beta2[2]
LB=round(exp(Estimate_indep$beta2[1])-qnorm(0.975)*se_exp,3)
UB=round(exp(Estimate_indep$beta2[1])+qnorm(0.975)*se_exp,3)
##exp(beta2)
exp_beta2
##CI for beta1
c(LB,UB)
##eta Maximum penalized log-likelihood
round(Estimate_indep$eta[1],3)
round(Estimate_indep$ML[1],3)
```