Chapter 6. Selected Topics on Multivariate Survival Time Analysis

1. Parametric estimation

2. Nonparametric estimation

3. Regression analysis

- Types of multivariate failure times
  - parallel multivariate failure times
    - twin studies
    - family studies
    - group randomized clinical trials
  - serial multivariate failure times
    - recurrent event times of same/different kind/kinds
    - gap times between repeated opportunistic infections
    - serial events in disease progression

1. Parametric estimation

- Bivariate survival function

$$S(t_1, t_2) = \text{Pr}\{T_1 > t_1, T_2 > t_2\}$$

  - if T_1 and T_2 are independent, $S(t_1, t_2) = S_{T_1}(t_1)S_{T_2}(t_2)$, and the usual univariate analysis can be extended
  - complication arises as T_1 and T_2 are correlated

- Shared frailty model
  - assume there is a frailty underlying the bivariate failure times, w, such that conditional on w,

$$\lambda_1(t \mid W = w) = w\lambda_1(t) \text{ and } \lambda_2(t \mid W = w) = w\lambda_2(t)$$

    by way of the proportional hazards model
  - positive dependence; can be extended to negative dependence
  - conditional bivariate survival function

$$S(t_1, t_2 \mid w) = \exp[-w\{\Lambda_1(t_1) + \Lambda_2(t_2)\}]$$

- marginal distribution

$$\begin{aligned} S(t_1, t_2) &= \int_w S(t_1, t_2 \mid w) dG(w) \\ &= E \exp[-w\{\Lambda_1(t_1) + \Lambda_2(t_2)\}] \\ &= L\{\Lambda_1(t_1) + \Lambda_2(t_2)\} \end{aligned}$$

- $L(u) = Ee^{-uW} = \int_w e^{-uw} dG(w)$ is Laplace transform of W

- $\Lambda(t_1, t_2) = \Lambda_1(t_1) + \Lambda_2(t_2)$

- density function

$$f(t_1, t_2) = \frac{\partial^2 S(t_1, t_2)}{\partial t_1 \partial t_2} = \frac{\partial^2 L\{\Lambda(t_1, t_2)\}}{\partial \Lambda_1 \partial \Lambda_2} \left\{ \prod_{j=1}^{2} \lambda_j(t_j) \right\}$$

- Data collected on a pair of censored bivariate failure times are (x_1, δ_1) and (x_2, δ_2), which contributes to the likelihood by

$$\frac{\partial^{\delta_1+\delta_2} L\{\Lambda(t_1, t_2)\}}{\partial \Lambda_1^{\delta_1} \partial \Lambda_2^{\delta_2}} \prod_{j=1}^{2} \lambda_j(t_j)^{\delta_j}$$

- choices of hazard functions
  - symmetry: $\lambda_1(t) = \lambda_2(t)$
  - unrelated: $\lambda_1(t) = \lambda_1(t; \theta_1)$, $\lambda_2(t) = \lambda_2(t; \theta_2)$
  - PHM: $\lambda_1(t) = \theta_1 \lambda_0(t)$, $\lambda_2(t) = \theta_2 \lambda_0(t)$

- For $k \geq 2$, let $\Lambda.(t_1, \ldots, t_k) = \sum_{j=1}^{k} \Lambda_j(t_j)$. Then

$$S(t_1, \ldots, t_k) = L\{\Lambda.(t_1, \ldots, t_k)\}$$

- The density function is

$$f(t_1, \ldots, t_k) = (-1)^k L^{(k)}\{\Lambda.(t_1, \ldots, t_k)\} \prod_{j=1}^{k} \lambda_j(t_j)$$

- The likelihood function contribution should be proportional to

$$(-1)^{\delta.} \frac{\partial^{\delta.} L\{\Lambda(t_1, \ldots, t_k)\}}{\partial \Lambda_1^{\delta_1} \ldots \partial \Lambda_k^{\delta_k}} \prod_{j=1}^{k} \lambda_j(t_j)^{\delta_j},$$

where $\delta. = \sum_{j=1}^{k} \delta_j$

- Gamma frailty model: W follows Gamma distribution $G(\gamma, \theta)$ with density function

$$f_W(w) = \theta^\gamma w^{\gamma-1} \exp(-\theta w)/\Gamma(\gamma)$$

  - $EW = \gamma/\theta$ and $\text{var}(W) = \gamma/\theta^2$
  - γ is shape parameter and θ is scale parameter, usually restricted to $\gamma = \theta$
  - also called Clayton-Oakes model

- Bivariate survival function under Gamma frailty model:

$$S(t_1, t_2) = \theta^\gamma / \{\theta + \Lambda_1(t_1) + \Lambda_2(t_2)\}^\gamma$$

  - alternatively,

$$S(t_1, t_2) = \{S_1(t_1)^{-1/\gamma} + S_2(t_2)^{-1/\gamma} - 1\}^{-\gamma}$$

- Bivariate density function becomes

$$\lambda_1(t_1)\lambda_2(t_2)\{\theta + \Lambda_1(t_1) + \Lambda_2(t_2)\}^{-\gamma-2}\theta^\gamma(\gamma + 1)\gamma$$

- For general $k \geq 2$, the likelihood contribution is

$$\theta^\gamma(\theta + \Lambda.)^{-\gamma-\delta.}\Gamma(\gamma + \delta.)/\Gamma(\gamma) \prod_{j=1}^{k} \lambda_j(t_j)^{\delta_j}$$

- Advantages of Gamma frailty model
  - simplicity in the derivatives of the Laplace transform
  - easy to derive likelihood for any number of correlated events

- Weibull model:

$$\lambda_j(t) = \lambda_j \cdot \kappa t^{\kappa - 1}$$

  - bivariate survival function becomes

$$S(t_1, t_2) = 1/\{1 + (\lambda_1 t_1^{\kappa} + \lambda_2 t_2^{\kappa})/\gamma\}^{\gamma}$$

  - bivariate Burr distribution

- How to measure dependence? general requirements are
  - range: $\rho(T_1, T_2) \in [-1, 1]$
  - identity: $\rho(T, T) = 1$
  - oddity: $\rho(T_1, -T_2) = -\rho(T_1, T_2)$
  - symmetry: $\rho(T_1, T_2) = \rho(T_2, T_1)$
  - independence: $\rho(T_1, T_2) = 0$ for independent T_1 and T_2
  - linear invariance: $\rho(\alpha_1 + \beta_1 T_1, \alpha_2 + \beta_2 T_2) = \rho(T_1, T_2)$
  - continuity: $\rho(T_{1n}, T_{2n}) \to \rho(T_1, T_2)$ given $T_{1n} \to T_1$ and $T_{2n} \to T_2$.

- Some common dependence measures
  - correlation coefficient (Pearson correlation):

$$\rho(T_1, T_2) = \frac{\mathrm{cov}(T_1, T_2)}{\sqrt{\mathrm{var}(T_1)\mathrm{var}(T_2)}}$$

    * $\mathrm{cov}(T_1, T_2) = \int_0^\infty \int_0^\infty \{S(t_1, t_2) - S_1(t_1)S_2(t_2)\}dt_1 dt_2$
    * $S_1(t_1) = S(t_1, 0), S_2(t_2) = S(0, t_2)$
    * $E\{\mathrm{var}(T_2 \mid T_1)\} = (1 - \rho^2)\mathrm{var}(T_2)$:

$$\rho^2 = \frac{\mathrm{var}(T_2) - E\{\mathrm{var}(T_2 \mid T_1)\}}{\mathrm{var}(T_2)}:$$

      variation reduction attributable to T_1
    * not invariant under nonlinear transformation

- Kendall's coefficient of concordance
  * for two pairs, (T_{11}, T_{12}) and (T_{21}, T_{22})

$$\tau = E\mathsf{sign}\{(T_{11} - T_{21})(T_{12} - T_{22})\}$$

  * let $p = \mathsf{Pr}[\{(T_{11} - T_{21})(T_{12} - T_{22})\} > 0]$, then $\tau = 2p - 1$
  * alternatively,

$$\tau = 4 \int_0^1 \int_0^1 f(t_1, t_2) S(t_1, t_2) dt_1 dt_2 - 1$$

  * two pairs required for valid interpretation
  * invariant under time-transformation
  * able to accommodate censoring

- Spearman's correlation coefficient

$$\rho_s = 12 \int_0^1 \int_0^1 S(S_1^{-1}(u), S_2^{-1}(v)) du dv - 3$$

- median concordance

$$\kappa = E\text{sign}\{(T_1 - \text{median}(T_1))(T_2 - \text{median}(T_2))\}$$

- integrated hazard

$$\rho_h = \int_0^1 \int_0^1 \log u \log v f(u, v) du dv - 1$$

- local dependence

$$\rho(t) = \frac{S(t) S_{12}^{(2)}(t)}{S_1^{(1)}(t) S_2^{(1)}(t)}$$

- Examples:

  - Kendall's τ in shared frailty model:

$$\tau = 4\int_0^\infty sL(s)L^{(2)}(s)ds - 1$$

    or

$$\tau = 4\int_0^\infty q(v)/q'(v)dv + 1,$$

    where $q(v) = L^{-1}(v)$

  - median concordance in Gamma frailty model

$$\kappa = 4(2^{1+1/\gamma} - 1)^{-\gamma} - 1$$

- Shared frailty model:

$$\lambda_{ij}(t) = W_i \lambda_j(t)$$

  - W_i: shared frailty
  - $\lambda_j(t)$: parametric, usually Weibull

- Other shared frailty models
  - positive stable frailty
  - power variance function frailty
  - lognormal frailty

- Inferences on shared frailty model
  - Parametric models: using the likelihood function specified in marginal distributions
    * Examples: Gamma-Weibull and Stable-Weibull
  - Simple estimates:
    * Example: Gamma-frailty model, $\gamma = (1/\tau - 1)/2$
    * $\widehat{\gamma} = (1/\widehat{\tau} - 1)/2$

– EM-algorithm

* Data: (X, Δ) and frailty W. The full likelihood function becomes

$$\mathscr{L}_{(X,\Delta),W} = \mathscr{L}_{(X,\Delta)|W} \times \mathscr{L}_W$$

* Term I:

$$\prod_i \prod_j \left\{ W_i \lambda_0(X_{ij}) e^{\beta Z_{ij}} \right\}^{\Delta_{ij}} \exp\left\{ -\int_0^{X_{ij}} W_i \lambda_0(u) e^{\beta Z_{ij}} du \right\}$$

* Term II: $(\gamma = \theta)$

$$\prod_i \gamma^{\gamma} W_i^{\gamma-1} e^{-\gamma W_i} / \Gamma(\gamma)$$

$*$ Algorithm:

1. Augmentation: estimate β and $\Lambda(\cdot)$ without frailty

2. E-step: insert the means based on current parameter estimates by

$$E(W_i \mid \mathscr{F}_\infty) = \frac{\tilde{\gamma}_i}{\tilde{\theta}_i} = \frac{\gamma + \sum_j \Delta_j}{\gamma + \sum_j \Lambda_j(X_{ij})}$$

   and

$$E(\log W_i \mid \mathscr{F}_\infty) = \frac{\Gamma'(\tilde{\gamma}_i)}{\Gamma(\tilde{\gamma}_i)} - \log \tilde{\theta}_i$$

3. M-step: maximize $\mathscr{L}_1$ for β by way of MPLE and calculate $\Lambda(\cdot)$; maximize $\mathscr{L}_2$ for γ

4. iterate E-M steps till converged

- Other inference approaches
  - full marginal estimates
  - full conditional estimates
  - penalized likelihood
- Asymptotic theory: ongoing research
  - identifiability
  - boundary problem
  - censoring and nonparametric hazards

• Reference:

- Hougaard: *Analysis of Multivariate Survival Data*

2. Nonparametric estimation

- Goal and challenges: nonparametric estimation of bivariate survival function

$$S(t_1, t_2) = \mathrm{Pr}\{T_1 > t_1, T_2 > t_2\}$$

  - NPMLE under independent censoring may have uniqueness problem
  - bivariate cumulative hazard function

$$\Lambda(t_1, t_2) = \int_0^{t_1} \int_0^{t_2} \Lambda(du_1, du_2)$$

    does not determine the survival function over $[0, t_1] \times [0, t_2]$

- Consider $S(dt_1, dt_2)$:
  - when S is absolutely continuous, $S(dt_1, dt_2) = \frac{\partial^2 S(t_1,t_2)}{\partial t_1 \partial t_2} dt_1 dt_2$
  - when S is continuous in t_1 but not t_2, $S(dt_1, dt_2) = \frac{\partial S(t_1,\Delta t_2)}{\partial t_1} dt_1$
  - when S is continuous in t_2 but not t_1, $S(dt_1, dt_2) = \frac{\partial S(\Delta t_1,t_2)}{\partial t_2} dt_2$
  - when S is not continuous in either t_1 or t_2, $S(dt_1, dt_2) = S(\Delta t_1, \Delta t_2)$
- Then

$$S(t_1, t_2) = \int_{t_1}^{\infty} \int_{t_2}^{\infty} S(du_1, du_2) \text{ and } \Lambda(dt_1, dt_2) = \frac{S(dt_1, dt_2)}{S(t_1-, t_2-)}$$

- Therefore

$$\int_0^{t_1}\int_0^{t_2} S(du_1, du_2) = \int_0^{t_1}\int_0^{t_2} S(u_1-, u_2-)\Lambda(du_1, du_2)$$

- This implies a nonhomogeneous Volterra equation

$$S(t_1, t_2) = [S_1(t_1) + S_2(t_2) - 1] + \int_0^{t_1}\int_0^{t_2} S(t_1-, t_2-)\Lambda(du_1, du_2)$$

- A simple recursive estimation of $S(t_1, t_2)$
  - consider
    $$S(\Delta t_1, \Delta t_2) = S(t_1-, t_2-) - S(t_1-, t_2) - S(t_1, t_2-) + S(t_1, t_2)$$
  - since $S(\Delta t_1, \Delta t_2) = S(t_1-, t_2-)\Lambda(\Delta t_1, \Delta t_2)$:
    $$S(t_1, t_2) = S(t_1-, t_2) + S(t_1, t_2-) - S(t_1-, t_2-)[1 - \Lambda(\Delta t_1, \Delta t_2)]$$
  - solve $S(t_1, t_2)$ recursively at all failure time grid points in the risk region
  - $\Lambda(\Delta t_1, \Delta t_2)$ can be estimated by
    $$\widehat{\Lambda}(\Delta t_1, \Delta t_2) = \frac{\sum_i I(T_{i1} = t_1, T_{i2} = t_2, \Delta_{i1} = \Delta_{i2} = 1)}{\sum_i I(T_{i1} \geq t_1, T_{i2} \geq t_2)}$$
  - major limitation: inefficient due to poor correspondence between $\widehat{\Lambda}$ and Kaplan-Meier marginals

- Other estimators
  - Dabrowska (1988, Ann Stat)
  - Prentice & Cai (1992, Bmka)

- More reading:
  - Kalbfleisch & Prentice (§10.3 and §10.6)

3. Regression method

- Consider multivariate failure time $(T_1, T_2, \ldots, T_K)$
  - T_k: failure time of the kth type
  - Marginal model assumes that

$$\lambda_{ki}(t \mid Z_{ki}(t)) = \lambda_{k0}(t) e^{\beta_k^T Z_{ki}(t)}$$

  - β_k: covariate effect for kth type of failure; when all the failure times are of the same type, we can further assume $\beta_1 = \ldots = \beta_k$

- for the kth type of failure, the partial likelihood for β_k is

$$\mathscr{L}_k(\beta) = \prod_{i=1}^{n} \left[\frac{\exp\{\beta^T Z_{ki}(X_{ki})\}}{\sum_{l=1}^{n} Y_{kl}(X_{ki}) \exp\{\beta^T Z_{kl}(X_{ki})\}} \right]^{\Delta_{ki}}$$

- then we can solve $\partial \log \mathscr{L}_k(\beta)/\partial\beta = 0$ for $\widehat{\beta}_k$

- in general, $\widehat{\beta}_1, \ldots, \widehat{\beta}_K$ are correlated. nevertheless

$$n^{1/2}\{(\widehat{\beta}_1^T, \ldots, \widehat{\beta}_K^T)^T - (\beta_1^T, \ldots, \beta_K^T)\}^T \rightarrow_D \mathscr{N}(0, \Sigma)$$

  where Σ can be estimated by sandwich estimator $\widehat{\Sigma}$

- how to estimate $\widehat{\Sigma}$?
  - partial score equation for kth type of failure

$$S_k(\beta) = \int_0^\infty \{Z_{ki}(u) - \bar{Z}_k(u)\}\, dN_{ki}(u),$$

where

$$\bar{Z}_k(u) = \frac{\sum_{j=1}^n Y_{kj}(u) Z_{kj}(u) e^{\beta^T Z_{kj}(u)}}{\sum_{j=1}^n Y_{kj}(u) e^{\beta^T Z_{kj}(u)}}$$

  - by Taylor expansion

$$n^{-1/2} S_k(\beta_k) \doteq \widehat{A}_k(\beta^*) n^{1/2} (\widehat{\beta}_k - \beta_k)$$

where $\widehat{A}_k(\beta)$ is the estimated slope

- consider

$$S_k(\beta, t) = \sum_{i=1}^{n} \int_0^t \{Z_{ki}(u) - \bar{Z}_k(u)\}\, dM_{ki}(u)$$

is local square integrable martingale in t for the kth type

- the usual MCLT does not apply to $n^{-1/2}S_1(\beta_1, t), \ldots, n^{-1/2}S_k(\beta_k, t)$

- however

$$\begin{aligned} S_k(\beta) &= \sum_{i=1}^{n} \int_0^\infty \{Z_{ki}(u) - \bar{Z}_k(u)\}\, dM_{ki}(u) \\ &\doteq \sum_{i=1}^{n} \int_0^t \{Z_{ki}(u) - \bar{z}_k(u)\}\, dM_{ki}(u), \end{aligned}$$

where

$$\bar{z}_k(u) = \frac{E\{Y_k(u) Z_k(u) e^{\beta^T Z_k(u)}\}}{E\{Y_k(u) e^{\beta^T Z_k(u)}\}}$$

- then we can use the usual multivariate central limit theorem to obtain the normal approximation. The asymptotic covariance between $n^{1/2}(\widehat{\beta}_k - \beta_k)$ and $n^{1/2}(\widehat{\beta}_l - \beta_l)$ is

$$\Sigma_{kl} = A_k^{-1}(\beta_k) E\{s_k(\beta_k) s_l(\beta_l)^T\} A_l^{-1}(\beta_l)$$

  where

$$s_k(\beta) = \int_0^\infty \{Z_k(u) - \bar{z}_k(u)\} dM_k(u)$$

- estimate Σ_{kl} by

$$\widehat{\Sigma}_{kl} = \widehat{A}_k^{-1}(\widehat{\beta}_k) \widehat{S}_k(\widehat{\beta}_k) \widehat{S}_l(\widehat{\beta}_l)^T \widehat{A}_l^{-1}(\widehat{\beta}_l)$$

- when $\beta_1 = \ldots = \beta_k$, use the quasi-partial score equation to solve for β:

$$\sum_{k=1}^{K} S_k(\beta) = 0$$

- References:
  - Wei, Lin & Weissfeld (1989, JASA)
  - Kalbfleisch & Prentice (p. 305)