The Analysis of Failure Times in the Presence of Competing Risks

Prentice et al. (1978)

Biostat 572 University of Washington

April 2014

Motivations for Competing Risks

- Single cause vs multiple causes for failure process
- examples:
  Flue: H1N1, H3N2, B/Victoria, B/yamagata
  Acute Leukemia Bone marrow transplantation:(i) recurrence, (ii) GVHD, (iii) pneumonia
- Competing Risks analysis on:
  (a) inference on the effects of treatment/exposure on specific types of failure
  (b) interrelations among failure types
  (c) failure rates for some causes given "removal" of some/all other causes

Notations:

$T >= 0$, the time of failure, which may be right censored

$J \in \{1, 2, ...m\}$, the type of failure, which will be unknown if T is censored

$\mathbf{z} = (z_1, ..., z_p)$, regression vector

Cause-Specific Hazard Functions:

Overall failure rate (hazard function Cox 1972):

$$\lambda(t; \mathbf{z}) = \lim_{\Delta t \to 0} \frac{\mathbf{P}\{t \leq T < t + \Delta t \mid T \geq t; \mathbf{z}(\mathbf{t})\}}{\Delta t} \tag{1}$$

Cause-specific hazard function (Prentice and Breslow 1978):

$$\lambda_j(t; \mathbf{z}) = \lim_{\Delta t \to 0} \frac{\mathbf{P}\{t \leq T < t + \Delta t, \mathbf{J} = j \mid T \geq t; \mathbf{z}(\mathbf{t})\}}{\Delta t} \tag{2}$$

$$\lambda(t; \mathbf{z}) = \sum_{1}^{m} \lambda_j(t; \mathbf{z}) \tag{3}$$

Cause-Specific Density Functions:

Overall Survival Density function:

$$f\left(t;\mathbf{z}\right)=\lambda\left(t;\mathbf{z}\right)\mathbf{F}\left(t;\mathbf{z}\right) \tag{4}$$

Cause-specific Survival Density function:

$$f_j\left(t;\mathbf{z}\right)=\lambda_j\left(t;\mathbf{z}\right)\mathbf{F}\left(t;\mathbf{z}\right) \tag{5}$$

Cause-Specific Hazard Functions:

Suppose n subjects : $\left(t_i, \delta_i, j_i; \mathbf{z_i}\right), i = 1, ...n$
Overall Likelihood function:

$$L \propto \prod_{i=1}^{n} \left\{ f\left(t_i; \mathbf{z_i}\right)^{\delta_i} \mathbf{F}\left(t_i; \mathbf{z_i}\right)^{1-\delta_i} \right\} = \prod_{i=1}^{n} \left\{ \lambda\left(t_i; \mathbf{z_i}\right)^{\delta_i} \mathbf{F}\left(t_i; \mathbf{z_i}\right) \right\} \qquad (6)$$

Cause-Specific Likelihood function:

$$L_j \propto \prod_{i=1}^{n} \left\{ \lambda_{j_i}\left(t_i; \mathbf{z_i}\right)^{\delta_i} \mathbf{F}\left(t_i; \mathbf{z_i}\right) \right\} \qquad (7)$$

$$= \left(\prod_{i=1}^{n} \lambda_{j_i}\left(t_i; \mathbf{z_i}\right)^{\delta_i} \prod_{j=1}^{m} exp\left\{ -\int_0^{t_i} \lambda_i\left(\mu; \mathbf{z_i}\right) d\mu \right\} \right) \qquad (8)$$

Cause-Specific Hazard Functions:

Conclusion:

1. Upon rearrangement, the likelihood factors into a separate component for each failure type $j = 1, ..., m$.

2. Moreover, the likelihood factor for $\lambda_j(t; \mathbf{z})$ is precisely the same as would be obtained by regarding all failure types other than j as censored at the individual's failure time. Thus, any of the standard methods for estimating $\lambda(t; \mathbf{z})$ can be applied for inference on $\lambda_j(t; \mathbf{z})$.

Interrelation between Failure Types:

Let $Y_1, ..., Y_m$ to be the latent failure times of each type of failure, then

$$T = min\left(Y_1, ..., Y_m\right) \tag{9}$$

and the correspnding failure type is:

$$\mathbf{J} = \{j \mid Y_j \leq Y_k, k = 1, ..., m\} \tag{10}$$

Define the joint survivor function:

$$Q\left(t_1, ..., t_m; \mathbf{z}\right) = \mathbf{P}\left(Y_1 > t_1, Y_2 > t_2, ..., Y_m > t_m; \mathbf{z}\right) \tag{11}$$

Then, all estimable quantites (such as $\mathbf{F}\left(t; \mathbf{z}\right), \lambda_j\left(t; \mathbf{z}\right)$) can be written in term of Q.

Interrelation between Failure Types:

However,with data of the type $(T, J; \mathbf{z})$, Q is nonidentifiable. Therefore, with z time independent, we can not study the interrelation, or even test for independence among competing failure modes.

Next Step:

- Provide some promising approaches to identify interrelations among failure types.
- Use the acute leukemia example as an illustration
- Failure rate estimation following cause removal

Interrelation between Failure Types:

All estimable quantites (such as $\mathbf{F}(t;\mathbf{z}), \lambda_j(t;\mathbf{z})$) can be written in term of Q:

$$\mathbf{F}(t;\mathbf{z}) = Q(t, t, ..., t; \mathbf{z}), \tag{12}$$

and, provided Q is differentiable along $t_1 = ... = t_m = t$,

$$\lambda_j(t;\mathbf{z}) = \lim_{\Delta t \to 0} \frac{\mathbf{P}(t \leq T_j < t + \Delta t \mid T \geq t; \mathbf{z})}{\Delta t} \tag{13}$$

$$= \frac{-\partial log \mathbf{Q}(\mathbf{t_1}, ..., \mathbf{t_m}; \mathbf{z})}{\partial t_j} \,|_{t_1 = ... = t_m = t} \tag{14}$$

The conclusion from equation (8) shows that the likelihood function can be written entirely in terms of $\lambda_i = (t, \mathbf{z})$'s, it follows that functions of Q other than those given by (14) cannot be estimated.
One promising fix involves the use of time dependent covariate