GSERM - 2018
Survival Model Extensions

June 8, 2018 (late morning session)

Topics

- Stratification
- Duration Dependence
- Separation
- Proportional Hazards
- Competing Risks
- Cure Models
- Repeated Events
- Frailty Models

Stratification

- Allow different groups to have <u>different baseline hazards</u>
- Akin to different intercepts, but more flexible.
- *Assumes covariate effects are otherwise identical*
- Uses:
    - Unit/group heterogeneity
    - Nonproportional hazards
    - Simple models for duration dependence

Duration Dependence

1. *State Dependence*
   - E.g., Institutionalization / Degradation

Positive State Dependence $\longrightarrow$ Negative Duration Dependence

Negative State Dependence $\longrightarrow$ Positive Duration Dependence

2. *Unobserved / Unmodeled Heterogeneity*
   - $h(t|\mathbf{X}_i) \neq h(t|\mathbf{X}_j)$ for $\mathbf{X}_i = \mathbf{X}_j$
   - Adverse selection in the sample / data
   - Result: "Spurious" duration dependence

Knowing / detecting the difference is important...

Separation

"Separation" $=$ "perfect prediction"

Intuition:

```
      Dems
Yeas    0   1
   0 178  34
   1   0 219
```

$\Pr(Y = 1|X = 0) =?$
Fixes: Exact logistic regression; Firth's bias correction...

Survival $\rightarrow$ Firth-type corrections (Heinze et al.)

Proportional Hazards

For two individuals A and B:
$$h_A(t) = Ch_B(t)$$
where C is the hazard ratio between A and B.

Proportionality:

- "Flat" hazards $\rightarrow$ parallel
- Rising hazards $\rightarrow$ diverging
- Falling hazards $\rightarrow$ converging

Common models assume hazards are proportional...

- Implies covariate effects are constant over time...
- Must be checked / diagnosed
- If violated: Log-time interactions...

Competing Risks

R multiple kinds of events:

$$T_i \in T_{i1}, ..., T_{iR}$$

Observed duration:

$$T_i = \min(T_{i1}, ... T_{iR})$$

Event indicator:

$$D_i = r \text{ iff } T_i = T_{ri}$$

R censoring indicators:

$$C_{ir} = \begin{cases} 1 \text{ if observation } i \text{ experienced event } r \\ 0 \text{ otherwise} \end{cases}$$

The point:

- If risks are **independent**: Separate analyses for each event type.
- Dependent risks are harder...

Cure Models

Standard models assume:

$$\int_0^\infty f(t)\,dt = 1 \;\forall\; i.$$

I.e., *all observations will (eventually) experience the event of interest.*

Cure models relax this assumption...

- **Mixture** and **non-mixture** types of cure models
- Allow modeling of both $h_i(t)$ and $\Pr(C_i = 0)$...
- Can be finicky...

Multiple / Repeated Events

Usual assumption: Events are not “absorbing” $\rightarrow$ capable of repetition

Raises (at least) two issues:

- Dependence across events

- Parameter variability

Relaxing that assumption...

- First issue: *variance correction* approaches...
- Second: *covariate-by-stratum interactions*.

See e.g. Box-Steffenemeier and Zorn (2002)...

“Frailty” Models

$$h_i(t) = \lambda_i(t)\nu_i$$

- $\nu_i = 1 \approx$ “baseline,”
- $\nu_i > 1 \rightarrow i$ has a greater-than-average hazard,
- $\nu_i < 1 \rightarrow$ the opposite.

Essentially “random effects” for survival data...

- Available for both parametric and Cox models
- Usual random effects assumption ($Cov(\mathbf{X}_{it}, \nu_i) = 0$) assumption is required
- Can be very data-intensive...