Competing Risks[1]

STA312 Spring 2019

[1]See last slide for copyright information.

Background Reading

- Chapter 9 (Especially 9.2.4) in *Applied Survival Analysis Using R* by Dirk Moore

Competing Risks: Common in practice

- Death from cancer, death from other cause, censored.
- For kidney patients: death, transplant, censored.
- For Ph.D. students: Graduate, withdraw, disappear, censored.
- Marriage: Divorce, separation, widowed, censored.

Three Approaches

There are probably more

- Analyze the data for each outcome in turn, treating the others as censored.
- Variation on the first method, with stratification and frailty.
- "Sub-distribution" approach.

One-at-a-time

Analyze the data for each outcome in turn, treating the others as censored.

- Straightforward and easy, but
- Is the censoring mechanism independent of the failure process?
- If x_k affects outcome 2 differently from outcome 3, it's hard to test.

Variation on one-at-a-time

Therneau and Grambsch, pp. 175-177

- Create multiple lines of data for each participant, one for each outcome (except censored).
- For all the outcomes that did *not* happen, the outcome is recorded as censored.
- So for example,
  - If there are 3 outcomes in addition to censoring, each case contributes 3 lines.
  - At most one line has $\delta = 1$. The others have $\delta = 0$; they are censored.
  - "Endpoint" is a variable with different values for the 3 lines.
  - Stratify on endpoint.
  - Also, tie the lines together with a random effect for `id`.
  - This is meant to take care of lack of independence.
  - Different regression coefficients for the strata (outcomes) are possible.

Example

Of the last method described

```
> age1 = age * (endpoint=='death')
> age2 = age * (endpoint='transp')
> coxme(Surv(time, status) ~ hemoglobin + age1 + age2
                            + strata(endpoint) + (1|id) )
```

Objection?

- The random effect is affecting all the outcomes in the same way.
- Maybe there should be a different “frailty” for each outcome.
- Does it matter?

Sub-distribution method

Fine and Grey (1999)

- The “sub-distribution” function $F_k(t)$ is like a cdf, but only applies to outcome (cause of death) k.
- Instead of approaching one as $t \to \infty$, it approaches a limiting probability that the person will die of cause k.
- Corresponding to the sub-distribution function is the sub-distribution hazard

$$h_k(t) = \lim_{\Delta \to 0} \frac{P(t < T_k < t + \Delta \,|\, E)}{\Delta},$$

where the event $E = \{T_k > t\} \cup \{T_{k'} \leq t, k' \neq k\}$.

The meaning of this notation is “. . . the risk set includes not only those currently alive and at risk for the kth event type, but also those who died earlier of other causes.” (Our text, p. 129) Fine and Grey call this “unconventional.”

Sub-distributions and sub-hazard functions

- The sub-hazard is related to the sub-distribution function by

$$h_k(t) = -\frac{d}{dt} \log(1 - F_k(t))$$

- $h_k(t) = h_{0,k}(t)\, e^{\mathbf{x}^\top \boldsymbol{\beta}_k}$: Baseline hazard times regression function.
- In theory, there is a separate set of regression coefficients for each outcome.
- In software, each set is estimated in a different run.
- The method extends to time-varying covariates, in both theory and software.

Copyright Information