Frailty models for recurrent events and clustered data

NORBIS course, Oslo
Thursday, December 7, 2017
Odd O. Aalen
Dept. of Biostatistics,
Institute of Basic Medical Sciences
University of Oslo

From ABG book
- Chapter 6: Frailty
- Chapter 7: Multivariate frailty
- Only parts of the material is covered

Variation in susceptibility

Similarities in the Age-Specific Incidence of Colon and Testicular Cancers

“With this data, the hypothesis that 100% of the population is susceptible to colon carcinoma can be ruled out.”
Clustered survival data

Shared frailty models

- Survival data often form natural clusters
  - Several members from the same family
  - Repeated observations in the same individual (e.g. measurements of the duration of fillings in teeth)
- The simplest model is the shared frailty one. We assume the shared frailty model where we use the same model as in the univariate case: $Z \cdot \alpha(t)$
- But here we assume that all members of the same cluster (family) have the same value of the frailty $Z$.
- So the frailty measures variation between families.

Estimating frailty

Familial Relative Risk (FRR)

- Given two individuals, A and B, in a family
  \[ \text{FRR} = \frac{P(T_A \leq t \mid T_B \leq t)}{P(T_A \leq t)} \]
- Assume the event is rare, and let the variance of the frailty be $\delta$. Then: $\delta = \text{FRR}-1$
- Examples:
  - Testicular cancer: FRR=6, $\delta = 5$ (Brothers)
  - Colon cancer: FRR=3.1, $\delta = 2.1$ (Monozygotic twins)

Estimating familial risk for testicular cancer (from Morten Valberg)

The study

- 1,135,320 Norwegian families included (Linkage between Statistics Norway and the Cancer Registry of Norway).
- 7,524 contained at least one TC case.
- We construct models that take into account the structure, and size, of a family.
- We calculate relative risks of disease, given any combination of affected family members.
Some results for TC

- The FRR of TC with 95% CIs for an individual A given that up to four of his brothers (B, C, D, and E) develop cancer or not.

<table>
<thead>
<tr>
<th>Affected</th>
<th>Unaffected</th>
<th>FRR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>5.88</td>
<td>4.70, 7.36</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td>5.50</td>
<td>4.56, 6.64</td>
</tr>
<tr>
<td>B, C</td>
<td></td>
<td>5.22</td>
<td>4.30, 6.33</td>
</tr>
<tr>
<td>B, C, D</td>
<td></td>
<td>5.07</td>
<td>4.11, 6.27</td>
</tr>
<tr>
<td>B, C, D, E</td>
<td></td>
<td>2.71</td>
<td>1.93, 3.57</td>
</tr>
<tr>
<td>B, C, D, E</td>
<td></td>
<td>15.60</td>
<td>9.56, 26.11</td>
</tr>
</tbody>
</table>

What are multivariate survival data?

1. Repeated events over time.
2. Each individual may have several "units on test".
3. Data from families.

Example: Small bowel motility

- Study cyclic pattern of motility (spontaneous movements) of the small bowel in humans. Focus on MMC complexes which come with irregular intervals (lasting from minutes to several hours).
- Motility is very important from a clinical point of view.
- Data studied by frailty models in (Aalen & Husebye, 1991).

Example: Duration of amalgam fillings

- Study of duration of amalgam fillings in teeth.
- Include several patients who each have many fillings in their teeth.
- A study, including 32 patients, with from 4 to 38 fillings for each patient, was analyzed by means of frailty models by Aalen, Bjertness and Sønju (1995).
Example: Sleep data

- Data from Max Planck institute in Munich concerning sleep patterns.
- Analysing tendency to fall asleep, wake up, have REM periods etc.
- Clearly, many occurrences of events each night.

Data for shared frailty model

- Consider $m$ independent clusters
- For cluster $i$, let $T_{ij}$, $j = 1, \ldots, n_i$, denote the observation times for the $n_i$ units
- Let $D_{ij}$ be binary variables that are equal to 1 if survival time $(i,j)$ is uncensored and equal to 0 if it is censored. Write $D_{i*} = \sum_j D_{ij}$
- Consider only a gamma frailty distribution
- Explicit likelihood function is given in ABG, p. 279.

Empirical Bayes estimate

- The frailty in a cluster can be estimated by an empirical Bayes estimate
- Define: $V_i = \sum_j A(T_{ij})$. The estimate is:

$$\hat{Z}_i = E(Z_i | H_i) = \frac{\delta^{-1} + D_{i*}}{\delta^{-1} + V_i}$$

Amalgam fillings

- Large variation in risk between individuals
Questions

- What is the difference between simple (univariate) frailty models and multivariate frailty models?
- Why is it easier to estimate the latter ones?