Chapters 1 & 3

PubH 7450

©Wei Pan

Email: weip@biostat.umn.edu

Http: www.biostat.umn.edu/~weip

Outline

- Introduction: survival analysis
  censoring and truncation
- Examples
- Read Chapter 1, §3.1-3.4 (Type II censoring is not required).

Introduction

- Survival analysis: stat analysis of survival/failure/event times
- Outcome of interest: survival/failure/event *time*;
  can be any event of interest, e.g., death, development or relapse of a disease, quitting smoking,...
- Feature: incomplete obs's
- 1) Censoring
  X: event time of interest, e.g. death
  C: censoring time; e.g., end of study, loss to follow-up
- i) right censoring: only know $X_i > C_i$ with observed C_i.
  Data: (T_i, δ_i), $i = 1, ..., n$
  $T_i = min(X_i, C_i)$, $\delta_i = I(X_i \leq C_i)$
- ii) left censoring: only know $X_i < C_i$ with observed C_i.
  Data:

- iii) right- & left-censoring = "double" censoring
  (in AIDS studies, there is another use of "double censoring".)

- iv) interval censoring: only know $X_i \in [L_i, R_i]$ with L_i and R_i observed.
  $L_i < \infty$, $R_i = \infty \Longrightarrow$ right-censoring.
  $L_i = 0$, $R_i < \infty \Longrightarrow$ left-censoring.

- 2) Truncation

- i) left truncation: observe Y_{L_i}; only observe (exact or censored) X_i iff $X_i > Y_{L_i}$
  Left-truncated data: (Y_{L_i}, X_i), $i = 1, ..., n$
  L-T + R-C data: (Y_{L_i}, T_i, δ_i), $i = 1, ..., n$

- ii) right truncation: observe Y_{R_i}; only observe (exact or censored) X_i iff $X_i < Y_{R_i}$

Examples

- §1.2. Remission duration from a clinical trial for acute leukemia
  Table 1.1:
  trt (6-MP): 6, 6, 6, 6+, 7, 9+, 10, 10+,... (in months)
  placebo: 1, 1, 2, 2, 3,...
  Event: relapse; censoring: end of study
  Interpretatioon:
  6: replased at month 6;
  6+: leukemia not returned yet at month 6 (at the end of the study).

- §1.17. Time to the first use of marijuana
  191 high school boys were asked: "When did you first use marijuana?"; an answer could be one of
  1) exact age at the time, X_i;
  2) never used;

C_i: current age; only know $X_i > C_i$; right-censoring;
3) used but did not remember when;
C_i: current age; only know $X_i < C_i$; left-censoring.

- §1.16. Death times of elderly residents of a retirement community
Channing House Data: 1964-1975; followed 462 seniors who lived in a retirement center.
Event of interest: death time
Censoring: end of study, loss to follow-up (moving out the center)
Left-truncation: Y_{L_i} is subject i's age at the beginning of the study; subject i was in the study iff $X_i > Y_{L_i}$.
- §1.18. Time to cosmetic deterioration of breast cancer patients
Goal: to compare cosmetic effects of two groups: radiotherapy alone vs radiotherapy plus chemotherapy
Eevent of interest: first appearance of moderate or severe

breast retraction

Study design: observed inititially every 4-6 months; the interval length b/w visits gradually increased $\Longrightarrow$ ...

Three possibilities:

1) already happened at the first visit–left-censoring;

2) between two visits–interval-censoring;

3) not yet at the last visit–right-censoring.

Data:

(radiotherapy only): (0,7],...,(4, 11], (6, 10],..., 32+, 46+ (in months)

- §1.19. Time to AIDS (right truncation)

  Sampling: restrospective – all subjects had developed AIDS by June 30, 1986.

```
----+--------------+------------| \\
    HIV+           AIDS         June 30, 1986
```

HIV: Human Immunedeficiency Virus

AIDS: Acquired Immune Deficiency Syndrom

In 80's, a pressing public health question was important to predict the prevalence of the epidemic $\Longrightarrow$ predict *incubation* period. (e.g. by a random sampling $\Longrightarrow$ Pr(HIV+).)

Event time: X=(AIDS time) - (HIV+ time), incubation time

For the data, the cohort were infected by a contaminated blood transfusion, hence (HIV+ time) was known.

Y_R=(June 30, 86) - (HIV+ time)

Observed X_i iff $X_i \leq Y_{R_i}$ (why?)