What to do today (2022/02/17)?

Part I. Preliminaries

Part II. Parametric Inference

Part III. Nonparametric/Semi-parametric Inference

Part III.1. Introduction and Overview: Motivation

Part III.3. Kaplan-Meier Estimator

Part III.3. Nonparametric Tests

Part III.4. Cox Proportional Hazards Model

Part III.4.1 Modeling

Part III.4.2 Inference

Part III.4.3 Extensions

Brief Summary of Part III

A. Practical Example

B. Discussion about Homeworks 1, 2, and 3

C. Talk on Survival Analysis by A.A. Tsiatis

Part IV. Advanced Topics

Part III.4.2C Cox PH Model: Estimation of $h_0(\cdot)$

$$\hat{h}_0(t) = \begin{cases} \frac{d_l}{\sum_{j \in \mathcal{R}_{(l)}} e^{\beta z_j}} & t = V_l \\ 0 & otherwise \end{cases}$$

$$\hat{H}_0(t) = \int_0^t \hat{h}_0(u; \beta) du = \sum_{V_l \leq t} \hat{h}_0(V_l; \beta)$$

Finally, $\hat{H}_0(t) = H_0(t; \hat{\beta})$ (**Breslow estimator**)

Part III.4.2C Cox PH Model: Estimation of $h_0(\cdot)$

Remarks

- **Breslow estimator**: NPMLE
  - uniform consistency, weak convergence
- If $\beta = 0$, $H_0(t) = H(t)$
  - $\hat{H}(t) = \sum_{V_l \le t} \frac{d_l}{N_l}$: **Nelson-Aalen estimator**
  - $\hat{S}(t) = \exp(-\hat{H}(t))$: **Fleming-Harrington estimator**, an alternative to Kaplan-Meier estimator $\hat{S}_{KM}(t)$

Part III.4.2C Cox PH Model: Estimation of $h_0(\cdot)$

Example III.4. (cont'd) $n = 5$ indpt subjects and

$$Z = \begin{cases} 1 & \textit{treatment} \\ 0 & \textit{placebo} \end{cases}$$

(u_i, δ_i, z_i) : $(16,1,1)$, $(13,0,0)$, $(21,1,1)$, $(11,1,0)$, $(12,1,1)$
V_l : 11, 12, 13, 16, 21; $Z_{(l)}$: 0, 1, 0, 1, 1;
$\mathcal{R}_l$: $\{1,2,3,4,5\}, \{1,2,3,5\}, \{1,2,3\}, \{1,3\}, \{3\}$,
$\hat{\beta} = \frac{1}{2}\log 2 - \log 3$.

$$\hat{h}_0(t;\beta) = \begin{cases} \frac{d_l}{\sum_{j \in \mathcal{R}_{(l)}} e^{\beta z_j}} & t = V_l \\ 0 & \textit{otherwise} \end{cases}$$

Thus, for $t = 11, 12, 13, 16, 21$, $\hat{h}_0(t;\beta)$ is

$$\frac{1}{2+3e^{\beta}}, \ \frac{1}{1+3e^{\beta}}, \ \frac{0}{1+2e^{\beta}}, \ \frac{1}{2e^{\beta}}, \ \frac{1}{e^{\beta}}.$$

$$\hat{H}_0(t) = \int_0^t \hat{h}_0(u;\hat{\beta})du = \sum_{V_l \le t} \hat{h}_0(V_l;\hat{\beta}).$$

Part III.4.3 Cox PH Model: Extensions

III.4.3A. To include strata

e.g. $Z = \begin{cases} 1 & treatment \\ 0 & placebo \end{cases}$, $W = \begin{cases} 1 & male \\ 0 & female \end{cases}$, $\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$

To consider the Cox PH model with covariates Z, W:

$$h(t|Z, W) = h_0(t)e^{\beta_1 Z + \beta_2 W},$$

four hazards proportional to each other

Alternatively, stratify the subjects according to the categories of W, and consider the Cox PH model:

$$h(t|Z, W = w) = h_{w0}(t)e^{\beta Z}, \quad w = 1, 2, ..., K$$

K pairs of proportional hazards: the same effect across different strata

Inference: in stratum k $\Longrightarrow L_P^{(k)}(\beta)$; over all $\prod_{k=1}^{K} L_P^{(k)}(\beta)$

What if different treatment effects in difference strata?

Part III.4.3 Cox PH Model: Extensions

III.4.3B. Time-varying (time-dependent) covariates:
e.g. air pollution level at time t: $Z(t)$; time till an asthma attack

$$h(t|Z(t)) = h_0(t)e^{\beta Z(t)}$$

Inference: $L_P(\beta) = \prod_{i=1}^{n} \left(\frac{e^{\beta z_i(u_i)}}{\sum_{l \in \mathcal{R}_i} e^{\beta z_l(u_i)}} \right)^{\delta_i}$

Remarks

- interpretation of the effect of $Z(\cdot)$ on T?
- required information of $\{Z_i(t) : t > 0\}$?

Part III.4.3 Cox Proportional Hazards Model: Extensions

III.4.3C. Time-varying (time-dependent) coefficient:
e.g. treatment effect at time t: $\beta(t)$; time till an asthma attack

$$h(t|Z) = h_0(t)e^{\beta(t)Z}$$

Inference:

- assuming a lot of ties of observed event times, or
- specifying $\beta(t)$ into (i) $\beta(t;\alpha)$, (ii) a spline (piecewise-polynomial real function)

Brief Summary of Part III: A.

Analysis of ACTG 116b/117

Study Design

- to compare ddI with low/high dose to AZT
- endpoint – defined AIDS events/AIDS related death
- randomized, double-blinded, multicenter
- 913 HIV infected subjects enrolled and randomly assigned to one of the three arms (AZT, ddI low, ddI high)
- study took about 2 years

This analysis is

- to focus on AZT and ddI low dose groups, and
- to compare the two arms in terms of the time to the 1st AIDS event.

$\Longrightarrow$ censoring = dtdiscont or dtdeath - dtrand;
observed time = min(censoring, dt2AIDS - dtrand)

An Analysis of Data from ACTG116b/117

1. A Glance at Data

- Descriptive Summary:
  - 615 subjects: 311,304 in ddI, AZT
  - observed time range $(0, 729)$ (days)
  - 152 failures observed: 71 vs 81
  - median baseline CD4: 92 (cells/ml); AZT dur: 419 (days)

- Baseline Balance:
  - baseline CD4: medians: 97 vs 84; p-value: 0.0435 (t-test)
  - baseline AZT duration: medians: 427 vs 406.5; p=0.7589 (t-test)

2. Estm on Survior (nonparametric)

- Kaplan-Meier estm:

$$\hat{S}(t) = \prod_{s>t} \Big(1 - \frac{dN_{\cdot}(s)}{\bar{Y}_{\cdot}(s)}\Big) = \prod_{v_j > t} \Big(1 - \frac{d_j}{N_j}\Big)$$

- Nelson-Aalen estm based: $\exp(-\hat{\Lambda}(t))$

$$\hat{\Lambda}(t) = \int_0^t \frac{dN_{\cdot}(s)}{Y_{\cdot}(s)} = \sum_{v_j \leq t} \frac{d_j}{N_j}$$

- the corresponding pointwise CI and CB (c=2.527, 2.636; M=1000)

Determine CB's boundary c:

$$\mathrm{P}\left(\sup_{0<t<729}\left|\frac{\hat{S}(t)-S(t)}{\sqrt{n}S(t)\sigma(t)}\right| \leq c\right) = 95\%$$

Introducing

$$\tilde{Q}_n(t) = \frac{\sqrt{n}}{\hat{\sigma}(t)}\sum_{i=1}^{n}\left(\frac{\delta_i}{Y_{.}(t)}I(t \geq u_i)\right)Z_i$$

Step 1: generate M sets of $Z_1, \ldots, Z_n \sim N(0,1)$, iid and $\perp\!\!\!\perp$ the study data

Step 2: evaluate $\tilde{Q}_n(t)$ M times

Step 3: obtain M values of $\max_t\{|\tilde{Q}_n(t)|\}$

Step 4: $c = 95\%$ quantile of the M values

3. Tests to Compare the Two Arms (nonparametric)

$H_0 : \Lambda_1 = \Lambda_0$ vs H_1: otherwise

Fleming and Harrington test:

$$\int_0^\infty H_n(t) d\left(\hat{\Lambda}_1(t) - \hat{\Lambda}_0(t)\right),$$

$$H_n(t) = \left[\hat{S}^-(t)\right]^\alpha \frac{Y_{.1}(t) Y_{.0}(t)}{Y_{.1}(t) + Y_{.0}(t)}.$$

- Logrank test ($\alpha = 0$, Mantel-Haenszel test): p=0.0377
- Prentice-Wilcoxon test ($\alpha = 1$, Peto & Peto modification of Gehan-Wilcoxon test): p=0.0422

α	-8	-1	0	1	8
p-values	.369	.037	.038	.042	.115

4. With Cox Proportional Model $\lambda(t|z) = \lambda_b(t)\exp(\beta z)$

- z=indicator of subject in ddI:
  - $\hat{\beta} = -.337$;
  - relative risk $\lambda(t|z=1)/\lambda(t|z=0) = \exp(\hat{\beta}) = .714$
  - partial likelihood ratio test for $H_0 : \beta = 0$: p=.0383
  - baseline estimate (Breslow):

$$\hat{\Lambda}_b(t) = \int_0^t \frac{dN_.(s)}{\sum_{i=1}^n Y_i(s)exp(\hat{\beta}z_i)}$$

  - $\hat{\Lambda}_1(t) = \hat{\Lambda}_b(t) * \exp(-.337)$, $\hat{\Lambda}_0(t) = \hat{\Lambda}_b(t)$

- z: indicator of ddI, baseline CD4, baseline AZT duration
  - $\hat{\beta}_1 = -0.275$, $p = 0.092$
  - $\hat{\beta}_2 = -0.011$, $p < 0.001$
  - $\hat{\beta}_3 = 0.0003$, $p = 0.31$

  - $\Lambda(t)$ of an average subject ($z_2 = 92$, $z_3 = 419$)?

$\hat{\Lambda}_1(t)$
$= \hat{\Lambda}_b(t) * \exp(-.275 - .011 * 92 + .0003 * 419)$,
$\hat{\Lambda}_0(t)$
$= \hat{\Lambda}_b(t) * \exp(-.011 * 92 + .0003 * 419)$

5. With Parametric Models

- exponential? Weibull?
- to consider covariates? $(\Lambda_b(t)\exp(\beta z))$

Weibull: $\lambda(t|z) = \alpha\phi(z;\beta)t^{\alpha-1}$, where $\phi(z;\beta) = e^{\beta_0+\beta_1 z_1+\beta_2 z_2}$

$$\hat{\alpha} = 0.909; \quad \hat{\beta}_0 = -5.006;$$

$$\hat{\beta}_1 = -0.393; \quad \hat{\beta}_2 = -0.013$$

compare with the previous results?

6. Further Investigations

- consider more covariates, such as gender and age;
- dependent censoring? e.g., informative drop-outs/death
- include all experienced AIDS events
- competing risks: death vs AIDS events
- compare the 3 treatments simultaneously
- take interim reviews into account
- ... ...

Part I. Preliminaries

Part II. Parametric Inference

Part III. Nonparametric/Semi-parametric Inference

- *Part III.1. Introduction and Overview*
- *Part III.2. Kaplan-Meier Estimator*
- *Part III.3. Nonparametric Tests*
- *Part III.4. Cox Proportional Hazards Model*
- **Brief Summary of Part III.**
  - *A. Practical Example*
  - **B. Discussion about Homeworks 1, 2, and 3**
  - **C. Talk on Survival Analysis by A.A. Tsiatis**

Part IV. Advanced Topics