What to do this week (2022/02/15, 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**
- *A Brief Summary of Part III*
  - *A. Practical Example*
  - *B. Talk on Survival Analysis by A.A. Tsiatis*

Part IV. Advanced Topics

Part III.4.1 Cox Proportional Hazards Model: Modeling

Cox Proportional Hazards Model: (Cox, JRSSB 1972)
The hazard function of event time $T|Z = z$ is

$$h(t|z) = h_0(t)e^{\beta z}, \quad t > 0$$

The conditional survivor function is

$$S(t|z) = \exp(-\int_0^t h_0(u)e^{\beta z} du) = \exp(-H_0(t)e^{\beta z}), \quad t > 0$$

Part III.4.1 Cox Proportional Hazards Model: Modeling

Remarks

- the hazard ratio $h(t|Z = z_1)/h(t|Z = z_0) = e^{\beta(z_1 - z_0)}$ for all $t > 0$ *proportional*!
- $Z = \begin{cases} 1 & \textit{treatment} \\ 0 & \textit{placebo} \end{cases}$, e^{β}: treatment effect
- $Z_1 = \begin{cases} 1 & \textit{treatment} \\ 0 & \textit{placebo} \end{cases}$, $Z_2 = \begin{cases} 1 & \textit{male} \\ 0 & \textit{female} \end{cases}$, $\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$, $\mathbf{Z} = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$, $h(t|\mathbf{Z}) = h_0(t)e^{\boldsymbol{\beta}'\mathbf{Z}}$: relative impacts of the treatment to female and male are the same.

Part III.4.2A Cox Proportional Hazards Model: Estimation of β

Often is interested to estm β in the Cox PH model, for comparison/evaluate/assess effect ... ...

With right-censored event times along with the covariates

$$\{(U_i, \delta_i, Z_i) : i = 1, \ldots, n\}$$

from n indpt subjects and indpt censoring $T_i \perp\!\!\!\perp C_i$

$$L(\beta, h_0(\cdot) \big| data) = \prod_{i=1}^{n} \Big(h_0(u_i) e^{\beta z_i} \Big)^{\delta_i} \exp(-H_0(u_i) e^{\beta z_i})$$

$$L(\beta, h_0(\cdot) \big| data) = L_1(\beta \big| data) L_2(\beta, h_0(\cdot) \big| data)$$

$\Longrightarrow$ **the Cox partial likelihood function** (Cox, Biometrika 1975)

Part III.4.2A Cox Proportional Hazards Model: Estimation of β

the Cox partial likelihood function (Cox, Biometrika 1975)

$$L_1(\beta \big| data) = \prod_{i=1}^{n} \Big(\frac{e^{\beta z_i}}{\sum_{l \in \mathcal{R}_i} e^{\beta z_l}} \Big)^{\delta_i}$$

the risk set at time u_i: $\mathcal{R}_i = \{j : u_j \geq u_i\}$

$\Longrightarrow$ the MPLE (maximum partial likelihood estimator) of β:

$$\hat{\beta} = argmax_{all\ \beta} L_1(\beta \big| data)$$

With some conditions, as $n \to \infty$

- $\hat{\beta} \to \beta$ a.s.
- $\sqrt{n}(\hat{\beta} - \beta) \to N(0, ?)$ in distn

Part III.4.2A Cox Proportional Hazards Model: Estimation of β

Example III.4. $n = 5$ indpt subjects and $Z = \begin{cases} 1 & \textit{treatment} \\ 0 & \textit{placebo} \end{cases}$

(u_i, δ_i, z_i) : $(16, 1, 1)$, $(13, 0, 0)$, $(21, 1, 1)$, $(11, 1, 0)$, $(12, 1, 1)$
$\Longrightarrow \hat{\beta} = \frac{1}{2} \log 2 - \log 3$

Part III.4.2A Cox Proportional Hazards Model: Estimation of β

Remarks

- *implementation*
  - to use $\log L_1(\beta) = \sum_{i=1}^n \delta_i \{\beta z_i - \log(\sum_{l \in \mathcal{R}_i} e^{\beta z_l})\}$

    or $U(\beta) = \partial \log L_1(\beta) / \partial \beta = \sum_{i=1}^n \delta_i \{z_i - \frac{\sum_{l \in \mathcal{R}_i} z_l e^{\beta z_l}}{\sum_{l \in \mathcal{R}_i} e^{\beta z_l}}\} = 0$
  - e.g. *R: coxph*

Part III.4.2A Cox Proportional Hazards Model: Estimation of β

Remarks (cont'd)

- *interpretation*
  - recall *likelihood, marginal likelihood, conditional likelihood, partial likelihood*
  - **the Cox partial likelihood function** of β
    - conditional arguments
    - the marginal distn of the rank statistic when no tie, no censored observation, cfs: Kalbfleisch and Prentice (1980, 2011)

Part III.4.2B Cox Proportional Hazards Model: Testing on β

Consider $H_0 : \beta = 0$ vs $H_1 : \beta \neq 0$

the partial score test

$$U(\beta) = \partial \log L_1(\beta)/\partial\beta = \sum_{i=1}^{n} \delta_i \Big[z_i - \frac{\sum_{l \in \mathcal{R}_i} z_l e^{\beta z_l}}{\sum_{l \in \mathcal{R}_i} e^{\beta z_l}} \Big]$$

Based on $U(\beta)/\sqrt{n} \sim AN(0, ??)$ as $n \to \infty$ with some conditions,

$\Longrightarrow$ the partial score testing procedure ...

Part III.4.2B Cox Proportional Hazards Model: Testing on β

Remarks.

- e.g. when $Z = \begin{cases} 1 & \textit{treatment} \\ 0 & \textit{placebo} \end{cases}$

  $U(\beta)\big|_{\beta=0} = \sum_{l=1}^{L} \left(O_l - n_{.l} \frac{N_{1l}}{N_{.l}} \right) = O - E$, the numerator of the logrank test statistic

- the Wald-type, using the MPLE of β and its asymptotic normality?

- the PLRT, using the structure of LRT?

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

To learn about the whole hazard function, to predict ... ...

Thinking ..., if β is known,

- at time V_l, subject j in $\mathcal{R}_{(l)}$ with prob of failing $h_0(V_l)e^{\beta z_j}$;
- the 'average' prob of failing at V_l for all N_l subjects in $\mathcal{R}_{(l)}$ is $h_0(V_l)\sum_{j\in\mathcal{R}_{(l)}} e^{\beta z_j}\Big/N_l$;
- on the other hand, the proportion of failing at V_l is d_l/N_l

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

Thus

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

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

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

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

Remarks

- **Breslow estimator**: NPMLE
  - uniform consistency, weak convergence
- If $\beta = 0$, $H_0(t) = H(t)$
  - $\hat{H}(t) = \sum_{V_l \leq t} \frac{d_l}{N_l}$: **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 Proportional Hazards Model: Estimation of $h_0(\cdot)$

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

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

$$(u_i, \delta_i, z_i): \; (16, 1, 1), \; (13, 0, 0), \; (21, 1, 1), \; (11, 1, 0), \; (12, 1, 1)$$

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)

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*
- **A Summary of Part III.**
  - **A. Practical Example**
  - **B. Talk on Survival Analysis by A.A. Tsiatis**

Part IV. Advanced Topics