Chapter 2

Cox Model Introduction

Fall, 2017 at WHU

Basic Specifications

- $\widetilde{T}$: the potential failure time (≥ 0);
- C: the potential censoring time;
- $T = \min(\widetilde{T}, C)$: the observed time;
- $\Delta = I(T \leq C)$: the censoring indicator;

$$\Delta = \begin{cases} 1 & \text{failed} \\ 0 & \text{censored} = \begin{cases} \text{study ends} \\ \text{lost} \\ \text{withdraw} \end{cases} \end{cases}$$

- Z: the p-dimensional covariate

Basic Specifications (Cont'd)

- The instantaneous rate at which failures occur for items that are surviving at time t, given Z:

$$\lambda(t|Z) = \lim_{h \to 0^+} \frac{P(t \leq \widetilde{T} < t + h | \widetilde{T} \geq t,\ Z)}{h}.$$

- $\lambda(t|Z)$ is called the hazard function of $\widetilde{T}$ given Z.

Basic Specifications (Cont'd)

- The cumulative hazard function:

$$\Lambda(t|Z) = \int_0^t \lambda(s|Z)ds$$

- $\lambda(t|Z) = \frac{f(t|Z)}{S(t|Z)} = -\frac{\log S(t|Z)}{dt}$,
- $S(t|Z) = e^{-\int_0^t \lambda(s|Z)ds} = e^{-\Lambda(t|Z)}$,
- $f(t|Z) = \lambda(t|Z)S(t|Z) = \lambda(t|Z)e^{-\Lambda(t|Z)}$

Cox Model

The Cox model for censored survival data specifies the hazard rate with covariate takes the form as:

$$\lambda(t|Z) = \lambda_0(t)\exp(Z'\beta)$$

- β: the regression parameters of interest;
- $\lambda_0(t)$: the unspecified baseline hazard function.

Cox Model – Survival Function

The survival function of the Cox model:

$$\begin{aligned} S(t|Z) &= \exp\left[-\int_0^t \lambda(s|Z)ds\right] \\ &= \exp\left[-\int_0^t \lambda_0(s)ds \cdot e^{Z'\beta}\right] \\ &= \exp\left\{-\Lambda_0(t)e^{Z'\beta}\right\} \\ &= \left[\exp\left\{-\Lambda_0(t)\right\}\right]^{\exp(Z'\beta)} \\ &= \left\{S_0(t)\right\}^{\exp(Z'\beta)} \end{aligned}$$

- $S_0(t)$ is the baseline survival function.

Cox Model – Density Function

The density function of the Cox model:

$$f(t|Z) = \{\lambda_0(t)\exp(Z'\beta)\}\exp\left\{-\exp(Z'\beta)\int_0^t \lambda_0(s)ds\right\}$$

- the Cox model is called the semiparametric regression model.

Estimation Procedures – Partial Likelihood

The partial likelihood for the inference of β is given by Cox (1972):

$$L_P(\beta) = \prod_{i=1}^{n} \left[\frac{e^{Z_i'\beta}}{\sum_{l \in R(T_i)} e^{Z_l'\beta}} \right]^{\Delta_i}$$

where the at-risk set at time t

$$R(t) = \{j : \ T_j \geq t\}.$$

Estimation Procedures – Partial Likelihood

The corresponding log-likelihood function is

$$l_P(\beta) = \sum_{i=1}^{n} \Delta_i \left[Z_i'\beta - \log\left\{ \sum_{l \in R(T_i)} e^{Z_l'\beta} \right\} \right].$$

The estimator of β is defined as,

$$\widehat{\beta} = \arg\max \; l_P(\beta).$$

Estimation Procedures – Score Equation

The score equation is

$$U(\beta) = \sum_{i=1}^{n} \Delta_i \left[Z_i - \frac{\sum_{l \in R(T_i)} Z_l e^{Z_l' \beta}}{\sum_{l \in R(T_i)} e^{Z_l \beta}} \right] = 0$$

The estimator $\widehat{\beta}$ can be obtained by solving the above score equation.

Estimation Procedures – Survival Function

The estimator of the baseline survival function is:

$$\widehat{S}_0(t) = \prod_{t_i < t} \left[1 - \frac{\exp(Z_i' \widehat{\beta})}{\sum_{l \in R(T_i)} \exp(Z_l' \hat{\beta})} \right]^{\exp(Z_i' \widehat{\beta})}.$$

The above Kalbfleisch-Prentice method is an extension of Kaplan-Meier estimator.

Estimation Procedures – Survival Function

The estimator of the baseline survival function is:

$$\widehat{S}_0(t) = \prod_{t_i < t} \left[- \frac{\Delta_i}{\sum_{l \in R(T_i)} \exp(Z_l' \widehat{\beta})} \right].$$

The estimator of the baseline cumulative hazard function is:

$$\widehat{\Lambda}_0(t) = \sum_{t_i < t} \left[\frac{\Delta_i}{\sum_{l \in R(T_i)} \exp(Z_l' \widehat{\beta})} \right].$$

The Breslow method on the survival function is based on the Nelson-Aalen estimator.

Cox Model – Time-Dependent Covariates

- For some other variables, their values may change along the course of a particular life event;
- posing potential threats to the validity of the time-dependent assumption on covaritates.

Cox Model – Time-Dependent Covariates

The hazard function:

$$\lambda(t, Z(t)) = \lambda_0(t) \exp\left\{Z(t)'\beta\right\}.$$

The partial likelihood function:

$$L_P(\beta) = \prod_{i=1}^{n} \left[\frac{e^{Z_i(T_i)'\beta}}{\sum_{l \in R(T_i)} e^{Z_l(T_i)'\beta}} \right]^{\Delta_i}.$$

Cox Model – Time-Dependent Covariates

The partial likelihood fucntion:

$$l_P(\beta) = \sum_{i=1}^{n} \Delta_i \left[Z_i(T_i)'\beta - \log \left\{ \sum_{l \in R(T_i)} e^{Z_l(T_i)'\beta} \right\} \right],$$

The score equation:

$$U(\beta) = \sum_{i=1}^{n} \Delta_i \left[Z_i(T_i) - \frac{\sum_{l \in R(T_i)} Z_l(T_i) e^{Z_l(T_i)'\beta}}{\sum_{l \in R(T_i)} e^{Z_l(T_i)'\beta}} \right] = 0.$$

Asymptotic Properties

Theorem 1 (Consistency)

Under general regularity conditions, there exists, with probability going to one as $n \to \infty$, a sequence $\{\widehat{\beta}_n\}$ of solutions to the score equation such that

$$\widehat{\beta}_n \xrightarrow{P} \beta_0.$$

Asymptotic Properties

Theorem 2 (Asymptotic Normality)

The following asymptotic distributional properties hold:

$$\sqrt{n}(\widehat{\beta} - \beta_0) \xrightarrow{d} N(0,\ \Sigma(\beta_0)^{-1})$$

Simulation

Step 1: generate data

$$\lambda(t|X_1,\ X_2) = \lambda_0(t)\exp(\beta_1 X_1 + \beta_2 X_2)$$

- Set $n = 100,\ 200$;
- Set $\beta_1 = -0.5,\ \beta_2 = 0.693$;
- $X_1 \sim Bernoulli(0.5),\ X_2 \sim N(0,1)$;
- Set $\lambda_0(t) = 1$, then generate $\widetilde{T} \sim E(e^{\beta_1 X_1 + \beta_2 X_2})$;
- $C \sim U(0,c)$;
- $T = min(\widetilde{T}, C)$;
- $\Delta = I(\widetilde{T} \le C)$.

Simulation (Cont'd)

Step 2: parameter estimation

The score equation:

$$U(\beta) = \sum_{i=1}^{n} \Delta_i \left[Z_i - \frac{\sum_{l \in R(T_i)} Z_l e^{Z_l' \beta}}{\sum_{l \in R(T_i)} e^{Z_l' \beta}} \right] = 0,$$

Hessian matrix:

$$H(\beta) = \sum_{i=1}^{n} \Delta_i \left[\frac{\sum_{l \in R(T_i)} Z_l^{\otimes 2} e^{Z_l' \beta}}{\sum_{l \in R(T_i)} e^{Z_l' \beta}} - \left\{ \frac{\sum_{l \in R(T_i)} Z_l e^{Z_l' \beta}}{\sum_{l \in R(T_i)} e^{Z_l' \beta}} \right\}^{\otimes 2} \right].$$

Newton-Raphson Algorithm:

$$\beta^{(m+1)} = \beta^{(m)} + H^{-1}(\beta^{(m)}) U(\beta^{(m)}).$$

Simulation (Cont'd)

Step 3: standard error estimation

$$\sqrt{n}(\widehat{\beta} - \beta_0) \to N(0, \Sigma^{-1}(\beta_0))$$

1. $\sqrt{n}(\widehat{\beta} - \beta_0) = \left\{\frac{1}{n}H(\beta_0)\right\}\left\{\frac{1}{\sqrt{n}}U(\beta_0)\right\} + o_P(1)$;
2. $\frac{1}{n}H(\beta_0) \to \Sigma(\beta_0)$;
3. $\frac{1}{\sqrt{n}}U(\beta_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i \to N(0, \Sigma(\beta_0))$,

   where $\eta_i = \Delta_i\left[Z_i - \frac{\sum_{l\in R(T_i)} Z_l e^{Z_l'\beta}}{\sum_{l\in R(T_i)} e^{Z_l'\beta}}\right]$;
4. $\widehat{\Sigma}(\widehat{\beta}) = \left\{\frac{1}{n}H(\widehat{\beta})\right\}^{-1}\widehat{E}(\eta^2)\left\{\frac{1}{n}H(\widehat{\beta})\right\}^{-1}$,

   where $\widehat{E}(\eta^2) = \frac{1}{n}\sum_{i=1}^{n}\eta_i^2$.
5. $\widehat{se} = \sqrt{\widehat{\Sigma}(\widehat{\beta})}$.

Simulation (Cont'd)

Step 4: interval estimator

$$\sqrt{n}(\widehat{\beta} - \beta_0) \sim N(0, \Sigma^{-1}(\beta_0))$$

$$P\left(\widehat{\beta} \in \left[\beta_0 \pm z_{\alpha/2}\sqrt{\frac{\widehat{\Sigma}(\widehat{\beta})}{n}}\right]\right) = 1 - \alpha$$

If

$$\widehat{\beta} \in \left[\beta_0 \pm z_{\alpha/2}\sqrt{\frac{\widehat{\Sigma}(\widehat{\beta})}{n}}\right]$$

cp=1, otherwise cp=0.

Simulation (Cont'd)

Step 5: simulation

1. give β_0 and n;
2. generate data by Step 1;
3. calculate $\widehat{\beta}^{(b)}, \widehat{se}^{(b)}, cp^{(b)},\ b = 1, \cdots, 1000$, go back to S2.
4. Mean$= \frac{1}{n}\sum_{b=1}^{1000}\widehat{\beta}^{(b)}$,

   SD$= \sqrt{\frac{1}{n-1}\sum_{b=1}^{1000}\left[\widehat{\beta}^{(b)} - \frac{1}{n}\sum_{b=1}^{1000}\widehat{\beta}^{(b)}\right]^2}$,

   SE$= \frac{1}{n}\sum_{b=1}^{1000}\widehat{se}^{(b)}$

   CP$= \frac{1}{n}\sum_{b=1}^{1000}cp^{(b)}$