Introduction (regression models)

Accelerated Failure Time Model (AFT model)

X : survival time

$$Y = \ln X$$

$$Y = \mu + \gamma^{T} Z + \sigma W \quad , W \sim \text{random error}$$

Proportional Hazards model (PH model)

$$h(x \mid z) = h_0(x) \exp(\beta^{T} z)$$

The Cox model is a special case of PH model when $h_0(x)$ is unspecified (nonparametric)

various forms of AFT and Cox models

(1) AFT : $S(x|z) = S_0(x \exp -\gamma^T z)$

Cox : $S(x|z) = [S_0(x)]^{\exp \beta^T z}$

(2) ☆ AFT : $Y = \mu + \gamma^T Z + \sigma W$

Cox : $\ell n[-\ell n\, S(x|z)] = \beta^T z + \ell n[-\ell n\, S_0(x)]$

(3) AFT : $h(x|z) = h_0(x e^{-\gamma^T z}) \exp(-\gamma^T z)$

☆ Cox : $h(x|z) = h_0(x) \exp(\beta^T z)$

Cox model : easy to do inference but hard to interpret

△The Cox model

Cox(1972)

$$h(t|z) = h_0(t)\exp \beta^T z$$

$h_0(t)$: Not specified → Nonparametric

$$\frac{h(t|z=\zeta_1)}{h(t|z=\zeta_0)} = \frac{h_0(t)\exp \beta^T \zeta_1}{h_0(t)\exp \beta^T \zeta_0} = \exp \beta^T(\zeta_1 - \zeta_0) \Rightarrow$$

This ratio is constant

So Cox model also called PH model.

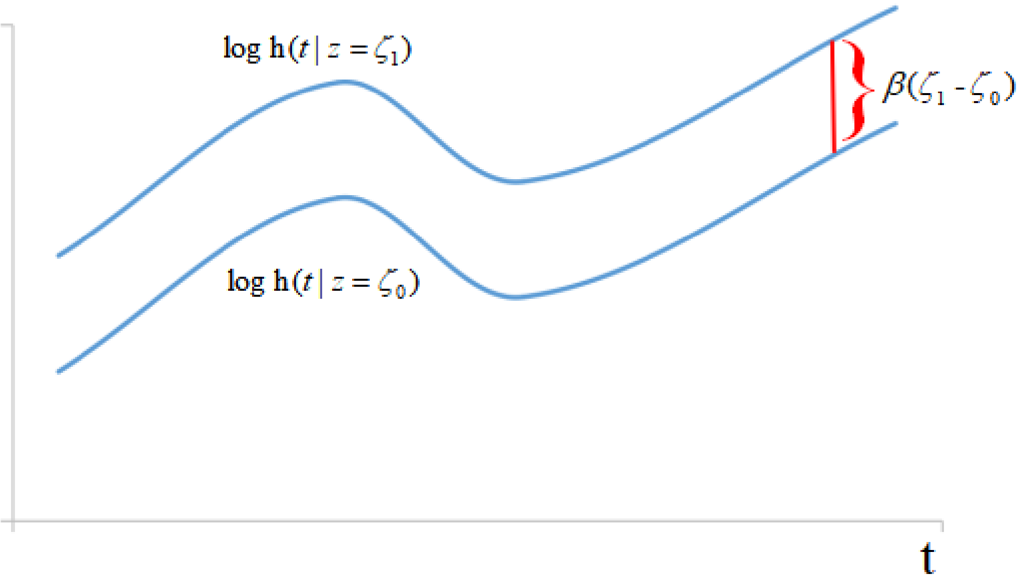
$\zeta_1 = 1$ Male

$\zeta_0 = 0$ Female

$$\zeta_1 - \zeta_0 = 1 \Rightarrow \exp\beta(\zeta_1 - \zeta_0) = \exp \beta$$

M relative risk to F

$$\begin{cases} \log h(t|z=\zeta_1) = \log h_0(t) + \beta\zeta_1 \\ \log h(t|z=\zeta_0) = \log h_0(t) + \beta\zeta_0 \end{cases}$$



△Partial likelihoods for distinct event time data

Suppose that there are no ties between event times

Let $t_1 < t_2 < \cdots\cdots < t_D$ denote the ordered event

and Z is a covariate.

$R(t_i)$: risk set at time t_i

(The set of all individuals who are still under study at time just prior to t_i)

$$\text{Hazard function} = \begin{cases} \dfrac{f(t)}{S(t)} & \text{for continuous time} \\ P(X = t \mid X \geq t) = \dfrac{P(X = t)}{P(X \geq t)} & \text{for discrete time} \end{cases}$$

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq X \leq t + \Delta t \mid X > t)}{\Delta t}$$

$$\text{Likelihood function} : \begin{cases} \prod_{i=1}^{n} f(t_i) \\ \prod_{i=1}^{n} P(X = t_i) \end{cases}$$

Nonparametric methods treat time as discrete ☆☆☆

$$\Rightarrow L(\beta) = \prod_{i=1}^{n} P(X_i = t_i) = \prod_{i=1}^{n} p(\text{individual } i \text{ die at } t_i) = \boxed{\text{We can not handle this}}$$

However we can handle for individual i

$$\Rightarrow p(\text{individual i died at } t_i \mid \text{one died at } t_i)$$

$$= \frac{p(\text{individual i died at } t_i \mid \text{Survival to } t_i)}{p(\text{one death at } t_i \mid \text{Survival to } t_i)}$$

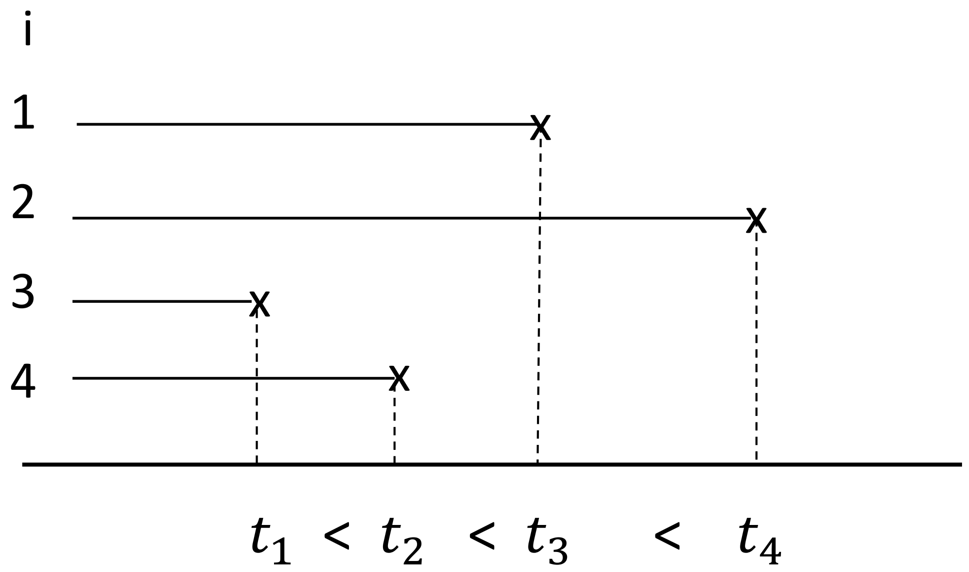
$$= \frac{h(t_i \mid z_i)}{\sum_{j \in R(t_i)} h(t_i \mid z_j)}$$

$$= \frac{h_0(t_i) \exp \beta z_i}{\sum_{j \in R(t_i)} h_0(t_i) \exp \beta z_j} = \frac{\exp \beta z_i}{\sum_{j \in R(t_i)} \exp \beta z_j}$$

$$L(\beta) = \prod_{i=1}^{D} \frac{\exp \beta z_i}{\sum_{j \in R(t_i)} \exp \beta z_j}$$

(We call it partial likelihood function not full likelihood function)

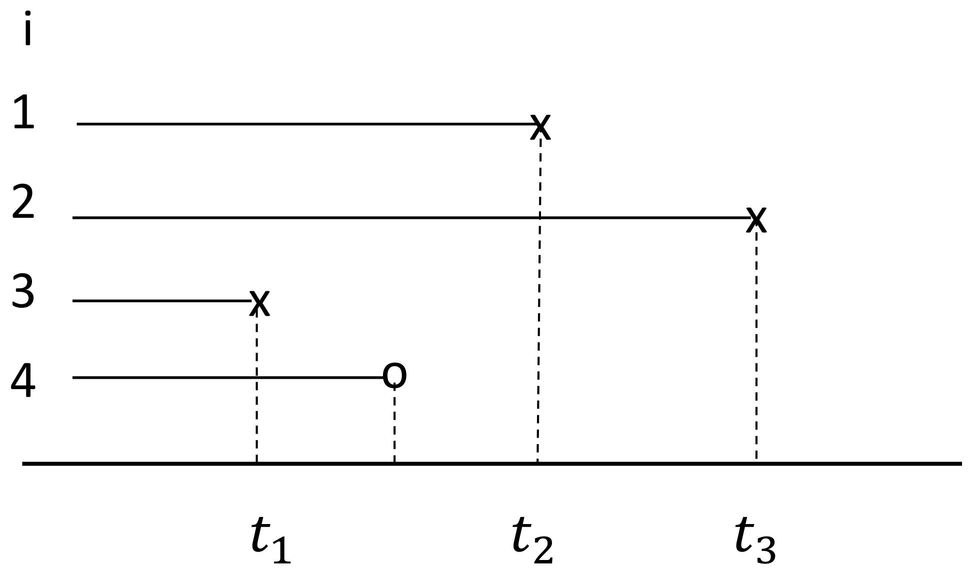
Ex:



R(t_1) = { 1, 2, 3, 4}
R(t_2) = { 1, 2, 4}
R(t_3) = { 1, 2}
R(t_4) = { 2}

$$L(\beta) = \frac{exp\ \beta z_3}{exp\ \beta z_1 + exp\ \beta z_2 + exp\ \beta z_3 + exp\ \beta z_4} \cdot \frac{exp\ \beta z_4}{exp\ \beta z_1 + exp\ \beta z_2 + exp\ \beta z_4} \cdot \frac{exp\ \beta z_1}{exp\ \beta z_1 + exp\ \beta z_2} \cdot \frac{exp\ \beta z_2}{exp\ \beta z_2}$$

Ex:



R(t_1) = { 1, 2, 3, 4}
R(t_2) = { 1, 2}
R(t_3) = { 2}

$$L(\beta) = \frac{exp\ \beta z_3}{exp\ \beta z_1 + exp\ \beta z_2 + exp\ \beta z_3 + exp\ \beta z_4} \cdot \frac{exp\ \beta z_1}{exp\ \beta z_1 + exp\ \beta z_2} \cdot \frac{exp\ \beta z_2}{exp\ \beta z_2}$$

Obtaining β

Let $U(\beta) = \frac{\partial\ log\ L(\beta)}{\partial\beta}$

$I(\beta) = \frac{\partial\ U(\beta)}{\partial\beta}$

$\Rightarrow \hat{\beta}^{(k+1)} = \hat{\beta}^{(k)} - I^{-1}\left(\hat{\beta}^{(k)}\right)U\left(\hat{\beta}^{(k)}\right)$

Starting value k = 0 ⇒ set $\hat{\beta}^{(0)}$ = c

When $\left|\frac{\widehat{\beta}^{(k+1)}-\widehat{\beta}^{(k)}}{\widehat{\beta}^{(k)}}\right| < \varepsilon$, stop

3 tests based on likelihood function

$H_0: \beta = \beta_0$

Wald test:

$$X_w^2 = \left(\hat{\beta} - \beta_0\right)^T I\left(\hat{\beta}\right)\left(\hat{\beta} - \beta_0\right)$$

Score test:

$$X_s^2 = U(\beta_0)^T I(\beta_0) U(\beta_0)$$

LR test:

$$X_{LR}^2 = 2\left[logL\left(\hat{\beta}\right) - logL(\beta_0)\right]$$

$$n \to \infty, \begin{cases} X_w^2 \\ X_s^2 \\ X_{LR}^2 \end{cases} \to \chi_p^2$$

△ When ties are present ⇒ Efron(1977) , Brelsow(1974) , Cox(1972)

correction

△ Testing proportional hazards

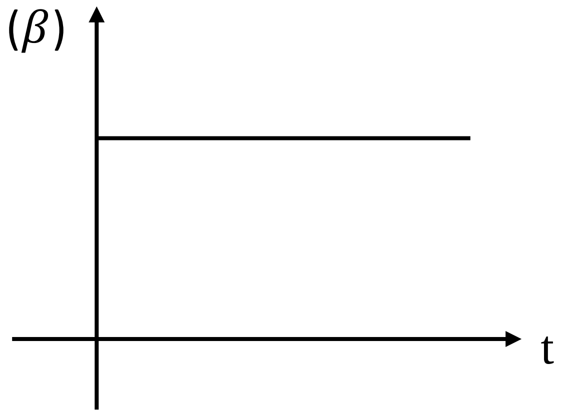
$$S(t|z) = [S_0(t)]^{exp\beta z}$$

Graphic ⇒ $\begin{cases} log[-log\, S(t|z)] = log[-log\, S_0(t)] + \beta z \\ log\, H(t) = log\, H_0(t) + \beta z \end{cases}$

test Cox.zph

$$h(t|z) = h_0(t)\, exp\big(z\beta(t)\big)$$

$\beta(t) = \beta$ ⇒ PH holds



limitation:

① Failure to detect non-proportionality

Ex: $\beta(t)$ qudratic shape

② Sample size can not be too small!

△ Strategies for nonproportional data

suppose that { (i)cox.zph test (Schoenfeld residual plot)
(ii)Diagnostic technique

gives strong evidence of nonproportionality for one or more covariates

What should we do?

① It does not matter (little effect)

② It is not real (outliers)

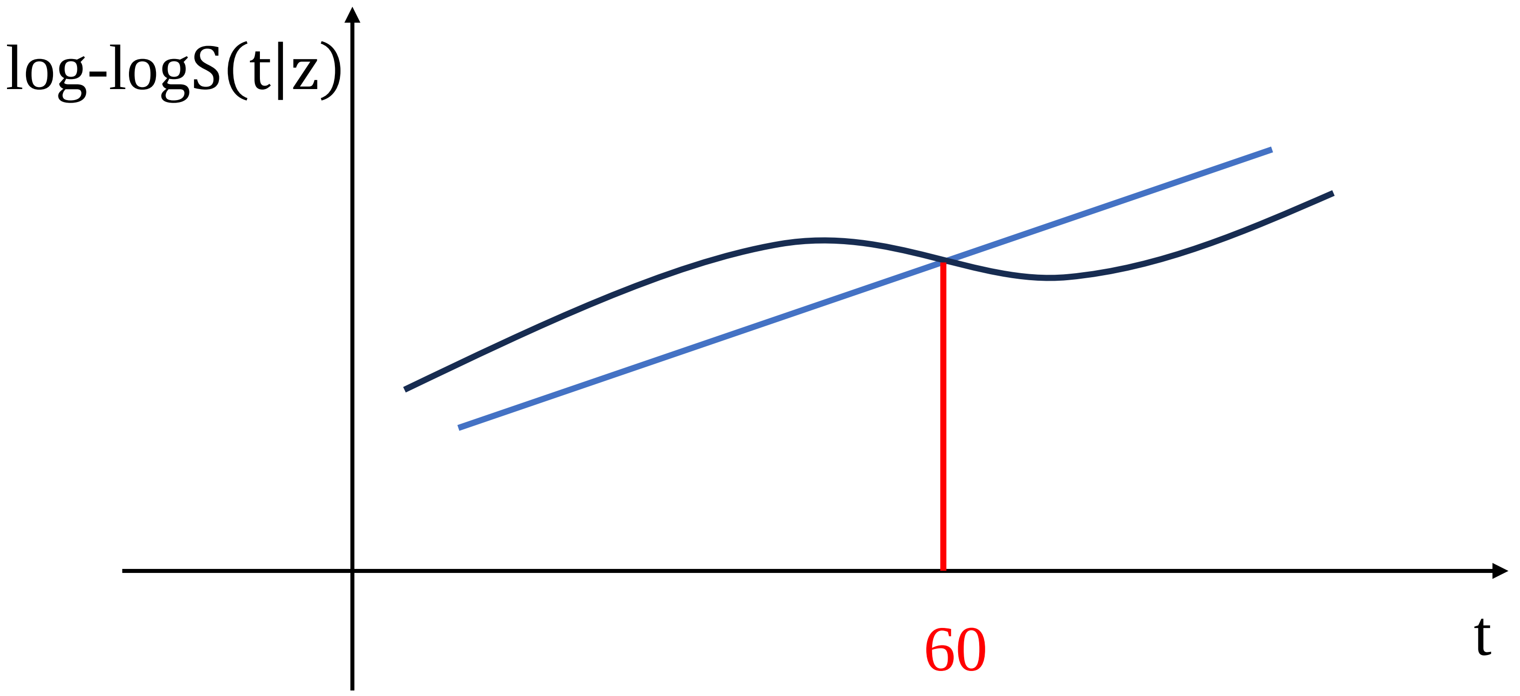
But if the proportionality is large and real , what then?

S1. Stratification:

$h_j(t|z) = h_{0j}(t) exp \beta z$, $j = 1,2, \ldots, s$ # of strata

$n_1 + n_2 + \cdots + n_s = n$

S2. Partition the time axis



coxph(Surv(time,censor) ~$Z_1 + Z_2$, subset (time < 60) , data = data)

S3. Model nonproportionality by time-dependent covariates

$\beta(t)Z = \beta Z^*(t)$ limitation : need to specify functional form of $\beta(t)$ (SAS)

S4. Use a different model

Accelerated Failure Time model(AFT)

Additive hazards model , ….

△ The Cox model with time-dependent covariates

Fixed covariate : $h(t|z) = h_0(t) \exp \beta z$

$$L(\beta) = \prod_{i=1}^{D} \frac{exp\beta Z_i}{\sum_{j \in R(t_i)} exp\beta Z_j}$$

Time-dependent covariates : $h(t|z(t)) = h_0(t) \exp \beta z(t)$

$$L(\beta) = \prod_{i=1}^{D} \frac{exp\beta Z_i(t_i)}{\sum_{j \in R(t_i)} exp\beta Z_j(t_i)}$$

	id	futime	status	death	trt	age	sex	day	start	stop	ascites	hepato	spiders	edema	bili
1	4	1925	2	0	1	54.74059	f	0	0	188	0	1	1	0.5	1.8
2	4	1925	2	0	1	54.74059	f	188	188	372	0	1	1	0.5	1.6
3	4	1925	2	0	1	54.74059	f	372	372	729	0	1	1	0.5	1.7
4	4	1925	2	0	1	54.74059	f	729	729	1254	0	1	1	1	3.2
5	4	1925	2	0	1	54.74059	f	1254	1254	1462	0	0	1	1	3.7
6	4	1925	2	0	1	54.74059	f	1462	1462	1824	0	1	1	1	4
7	4	1925	2	2	1	54.74059	f	1824	1824	1925	1	1	1	1	5.3