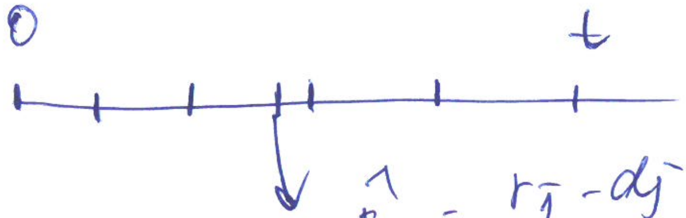
Lecture Week 5

KM estimator:



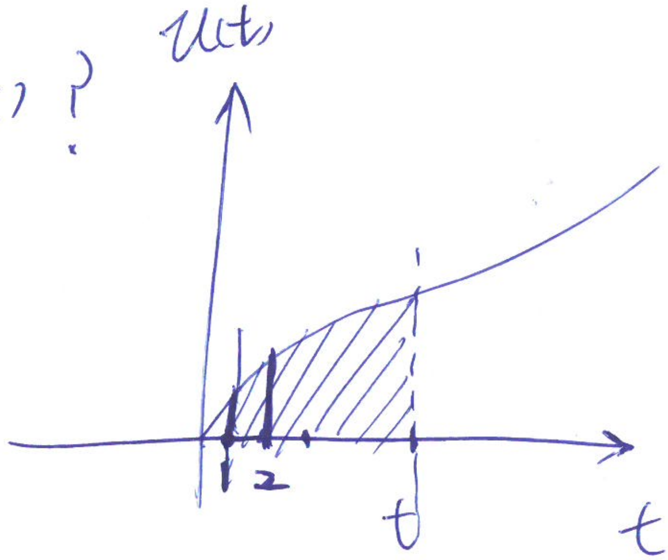
$\hat{p}_j = \frac{r_j - d_j}{r_j}$ $\qquad$ $\hat{S}(t) = \prod_{t_j \le t} \hat{p}_j$

$$S(t) = e^{-\int_0^t u(s)\,ds}$$

→ inspires NA estimator.

NA estimator: $\Lambda(t) = \int_0^t u(s)\,ds$

$S(t) = e^{-\Lambda(t)}$ $\qquad$ $\Lambda(t)$?



sum of harzard →
integral of harzard.

At each time point.

When a death does not occur, → 0.

when a death occurs → $\frac{d_j}{r_j} = \hat{q}_j$

$$\hat{\Lambda}(t) = \sum_{t_j \le t} \hat{q}_j = \sum_{t_j \le t} \frac{d_j}{r_j}$$

$$\hat{S}(t) = \exp\left(-\sum_{t_j \le t} \frac{d_j}{r_j}\right)$$

Conclusion. comparison

KM	t_j	r_j	d_j	$\frac{r_j - d_j}{r_j}$	$\prod_{t_j \le t} \frac{r_j - d_j}{r_j}$
NA	t_j	r_j	d_j	$\frac{d_j}{r_j}$	$\exp\left(-\sum_{t_j \le t} \frac{d_j}{r_j}\right)$

Variance of NA. $Var(\hat{S}(t))$

① find $Var[\hat{\Lambda}(t)]$ first.

$\hat{q}_j = \frac{d_j}{r_j}$. $\quad d_j \sim Bin(r_j, q_j)$

$$Var(\hat{q}_j) = \frac{1}{r_j^2} \cdot r_j \cdot \frac{d_j}{r_j} \cdot \left(1 - \frac{d_j}{r_j}\right)$$

$$= \frac{d_j(r_j - d_j)}{r_j^3}$$

$$\Rightarrow \operatorname{Var}(\hat{\Lambda}(t)) = \operatorname{Var}\left(\sum_{t_j \le t} \hat{q}_j\right)$$

$$= \sum_{t_j \le t} \operatorname{Var}(\hat{q}_j) \quad (\text{iid})$$

$$= \sum_{t_j \le t} \frac{d_j(r_j - d_j)}{r_j^3}$$

② $\hat{S}(t) = e^{-\hat{\Lambda}(t)}$

delta method:

$g(x) = e^{-x}$. $\quad \operatorname{Var}(g(x)) = (g'(u))^2 \cdot \operatorname{var}(x)$

$$E(\hat{\Lambda}(t)) = \sum_{t \le t_j} E\left(\frac{d_j}{r_j}\right) = \sum_{t_j \le t} q_j \approx \sum_{t_j \le t} \hat{q}_j = \Lambda(t)$$

$$\Rightarrow \operatorname{Var}(\hat{S}(t)) = (\hat{S}(t))^2 \sum_{t_j \le t} \frac{d_j(r_j - d_j)}{r_j^3}$$

Example

$8, 12, 12^*, 17, 17, 22, 27^*, 30.$

$\hat{S}(t=24) = ?$

t_j	r_j	d_j	$\frac{d_j}{r_j}$	$e^{-\sum \frac{d_j}{r_j}}$
8	8	1	$\frac{1}{8}$	$e^{-\frac{1}{8}}$
12	7	1	$\frac{1}{7}$	$e^{-\frac{1}{8}-\frac{1}{7}}$
17	5	2	$\frac{2}{5}$	$\vdots$
22	3	1	$\frac{1}{3}$	$\vdots$
30	1	1	1	$e^{-\frac{1}{8}-\frac{1}{7}\cdots-1}$

① $\hat{S}(24) = e^{-\left(\frac{1}{8}+\frac{1}{7}+\frac{2}{5}+\frac{1}{3}\right)}$ N.A.

② $= \exp\left(-\left(\frac{1}{8}+\frac{1}{7}+\frac{1}{5}+\frac{1}{4}+\frac{1}{3}\right)\right)$.

"survfit" function in R uses F.H

F.H.
Fleming-Harrington

Week 3.

Non-censoring :

Para $\begin{cases} \text{MOM} \\ \text{MLE} \end{cases}$

NON-Para — "dist. Free"

$\hat{F}(t)$ cdf.

Week 4~5

Censoring

NON-Para $\begin{cases} \text{KM} \\ - \\ \text{NA} \end{cases}$ $\hat{S}(t) = \begin{cases} \prod \frac{r-d}{r} \\ e^{-\sum \frac{d}{r}} \end{cases}$

C. I

Para — MLE — ?

impact of other covariates ?

COX

$$L(\lambda) = \begin{cases} f(t) & \text{complete} \\ S(c) & \text{censoring} \quad P(T_i \geq c_i) \end{cases}$$

$$L(\lambda) = \prod_{\text{complet}} f(t_i) \prod_{\text{censored}} S(c_j)$$

E.g. $2,\ 2.5^*,\ 3,\ 24,\ 1^*,\ 0.5.$

$f(t) = \lambda \exp(-\lambda t)$ find $\hat{\lambda}$

$$L(\lambda) = \prod_{i=1}^{4} \lambda e^{-\lambda t_i} \prod_{j=1}^{2} e^{-\lambda c_j}$$

$$P(T_i \geq c_j) = e^{-\lambda c_j}$$

$$l(\lambda) = \sum_i \log \lambda e^{-\lambda t_i} + \sum_j \log e^{-\lambda c_j}$$

$$= 4 \log \lambda + \sum_{i=1}^{4} (-\lambda t_i) + \sum_{j=1}^{2} (-\lambda c_j)$$

$$l'(\lambda) = \frac{4}{\lambda} - \sum_{i=1}^{4} t_i - \sum_{j=1}^{2} c_j \Rightarrow$$

$$\hat{\lambda} = \frac{4}{\sum \cdots} = \frac{4}{33}$$

$$\beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} \qquad X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \qquad \beta^T x = \sum_{i=1}^{p} \beta_i \cdot x_i$$

$$S(t) = \exp\Big(-\int_0^t \underbrace{\lambda(s; x)}_{\text{for individual with covariate } x} ds\Big)$$

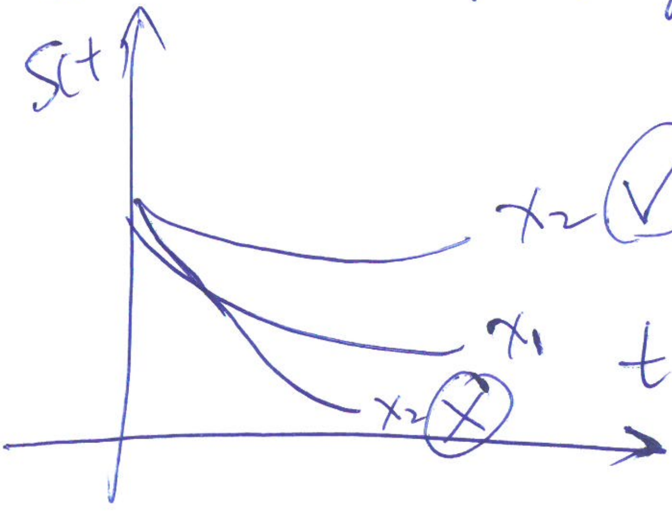
$$= \exp\Big(-\int_0^t \lambda_0(s) \underbrace{\exp(\beta^T x)}_{\text{Cox assumption model}} ds\Big)$$

$$= \exp\Big[\Big(-\int_0^t \lambda_0(s) ds\Big) \cdot \exp(\beta^T x)\Big]$$

$$= \Big[\exp\Big(-\int_0^t \lambda_0(s) ds\Big)\Big]^{\exp(\beta^T x)}$$

$$= S_0(t)^{\exp(\beta^T x)}$$

Can't cross? why



$$S_0(t)^{\exp(\beta^T x_1)} = S_0(t)^{\exp(\beta^T x_2)}$$

$$\Rightarrow x_1 = x_2$$

$$x_1 \neq x_2$$

$$\Rightarrow S_0(t)^{\exp(\beta^T x_1)} \neq S_0(t)^{\exp(\beta^T x_2)}$$

$$\mu_h \approx \frac{{}_h q_x}{h}. \qquad \underline{{}_h q_x \approx \mu_h \cdot h.}$$

$$\frac{P(\text{person with } x_{(s)})}{P(\text{one death} \cdots)} = \frac{\delta \lambda(x_{(s)}, t_s)}{\sum_{j \in R(t_s)} \delta \cdot \lambda(x_{(j)}, t_s)}$$

$$= \frac{\exp(\beta^T x_{(s)})}{\sum_{j \in R(t_s)} \exp(\beta^T x_j)}$$

---

tied deaths. example:

2 covariates. ① age ② blood pressure.

3 deaths at time t

$$x_{(i)} = \begin{bmatrix} \underset{\sim}{age} \\ bp \end{bmatrix} \quad \begin{aligned} &\rightarrow = age1 + age2 + age3. \\ &\rightarrow = bp1 + bp2 + bp3. \end{aligned}$$