Midterm exam, Survival Analysis I, 2016 Spring [+30 points]
Name: 張雅玟 Chang, Ya-Wen 103225008.

~~22~~
27

- **Not only answer but also derivations**

+15

Q1[+15]. Consider the AFT regression model: $Y = \log(T) \sim S_0\left(\dfrac{y-\boldsymbol{\beta}'\mathbf{x}}{b}\right)$, where $S_0(w)$ is a survival function with $f_0(w) = -dS_0(w)/dw$. We observe censored data $y_i = \min(\log T_i, \log C_i)$, $\delta_i = I(T_i \le C_i)$, and $\mathbf{x}_i' = (x_{i1}, \ldots, x_{ip})$, for $i = 1, \ldots, n$.

+2 1. [+2] Write down the log-likelihood using $z_i = \dfrac{y_i - \boldsymbol{\beta}'\mathbf{x}_i}{b}$ and $r = \sum_{i=1}^{n} \delta_i$.

$$L(\beta, b) = \prod_{i=1}^{n}\left[\frac{1}{b} f_0(z_i)\right]^{\delta_i} \cdot \left[S_0(z_i)\right]^{1-\delta_i}$$

$$\ell(\beta, b) = -r\log b + \sum_{i=1}^{n}\left[\delta_i \log f_0(z_i) + (1-\delta_i)\log S_0(z_i)\right] \ \#$$

+3 2. [+3] Write down the score function using $\dfrac{\partial}{\partial z}\log f_0(z)$ and $\dfrac{\partial}{\partial z}\log S_0(z)$.

$$\frac{\partial z_i}{\partial \beta} = -\frac{x_{ij}}{b}, \quad j=1,\ldots,p$$

$$\frac{\partial z_i}{\partial b} = \frac{-z_i}{b}$$

$$\frac{\partial \ell(\beta,b)}{\partial \beta} = -\frac{1}{b}\sum_{i=1}^{n}\left[\delta_i \frac{\partial}{\partial z_i}\log f_0(z_i) + (1-\delta_i)\frac{\partial}{\partial z_i}\log S_0(z_i)\right]\cdot x_{ij}, \quad j=1,\ldots,p \ \#$$

$$\frac{\partial \ell(\beta,b)}{\partial b} = \frac{-r}{b} - \frac{1}{b}\sum_{i=1}^{n}\left[\delta_i \frac{\partial}{\partial z_i}\log f_0(z_i) + (1-\delta_i)\frac{\partial}{\partial z_i}\log S_0(z_i)\right]\cdot z_i \ \#$$

+5 3. [+5] Write down the Hessian matrix using $\dfrac{\partial^2}{\partial z^2}\log f_0(z)$ and $\dfrac{\partial^2}{\partial z^2}\log S_0(z)$.

$$\frac{\partial \ell(\beta,b)^2}{\partial \beta'\partial \beta} = \frac{1}{b^2}\sum_{i=1}^{n}\left[\delta_i \frac{\partial^2}{\partial z_i^2}\log f_0(z_i) + (1-\delta_i)\cdot\frac{\partial^2}{\partial z_i^2}\log S_0(z_i)\right]\cdot x_{ij}\cdot x_{ik}, \quad j=1,\ldots,p,\ k=1,\ldots,p \ \#$$

$$\frac{\partial \ell(\beta,b)^2}{\partial b\,\partial \beta} = \frac{\partial \ell(\beta,b)^2}{\partial \beta\,\partial b} = \frac{1}{b^2}\sum_{i=1}^{n}\left[\delta_i \frac{\partial}{\partial z_i}\log f_0(z_i) + (1-\delta_i)\frac{\partial}{\partial z_i}\log S_0(z_i)\right]\cdot x_{ij}$$
$$+ \frac{1}{b^2}\sum_{i=1}^{n}\left[\delta_i \frac{\partial^2}{\partial z_i^2}\log f_0(z_i) + (1-\delta_i)\frac{\partial^2}{\partial z_i^2}\log S_0(z_i)\right]\cdot x_{ij}\cdot z_i \ \#$$

$$\frac{\partial \ell(\beta,b)^2}{\partial b^2} = \frac{r}{b^2} + \frac{1}{b^2}\sum_{i=1}^{n}\left[\delta_i \frac{\partial}{\partial z_i}\log f_0(z_i) + (1-\delta_i)\frac{\partial}{\partial z_i}\log S_0(z_i)\right]\cdot z_i$$
$$+ \frac{1}{b^2}\sum_{i=1}^{n}\left[\delta_i \frac{\partial}{\partial z_i}\log f_0(z_i) + (1-\delta_i)\frac{\partial}{\partial z_i}\log S_0(z_i)\right]\cdot z_i$$
$$+ \frac{1}{b^2}\sum_{i=1}^{n}\left[\delta_i \frac{\partial^2}{\partial z_i^2}\log f_0(z_i) + (1-\delta_i)\frac{\partial^2}{\partial z_i^2}\log S_0(z_i)\right]\cdot z_i^2$$
$$= \frac{r}{b^2} + \frac{2}{b^2}\sum_{i=1}^{n}\left[\delta_i \frac{\partial}{\partial z_i}\log f_0(z_i) + (1-\delta_i)\frac{\partial}{\partial z_i}\log S_0(z_i)\right]\cdot z_i$$
$$+ \frac{1}{b^2}\sum_{i=1}^{n}\left[\delta_i \frac{\partial^2}{\partial z_i^2}\log f_0(z_i) + (1-\delta_i)\frac{\partial^2}{\partial z_i^2}\log S_0(z_i)\right]\cdot z_i^2 \ \#$$

$$\Rightarrow \text{Hessian matrix} = \begin{bmatrix} \dfrac{\partial \ell(\beta,b)^2}{\partial \beta'\partial\beta} & \dfrac{\partial \ell(\beta,b)^2}{\partial\beta\,\partial b} \\[2ex] \dfrac{\partial \ell(\beta,b)^2}{\partial b\,\partial\beta} & \dfrac{\partial \ell(\beta,b)^2}{\partial b^2} \end{bmatrix}.$$

+9

~~+4~~

Q3.[+10] Let $S_0(w) = \exp(-e^w)$ in the same AFT regression model as Q1.

+2 1. [+2] Prove that the AFT regression model satisfies the proportional ~~hazard~~ AFT model.

$T \sim Weib(\alpha, \beta) \Rightarrow S(t|x) = \exp\left\{-\left(\frac{t}{\alpha(x)}\right)^{\delta}\right\} = S_0^*\left(\left(\frac{t}{\alpha(x)}\right)^{\delta}\right)$, where $S_0^*(w) = \exp(-w)$

$h(t|x) = \delta \cdot t^{\delta-1} \cdot \alpha(x)^{-\delta}$, where $\delta = \beta$, $\alpha(x) = \exp\{\beta' x\}$.

$= \delta \cdot t^{\delta-1} \cdot \exp\{-\delta\beta' x\}$

$= h_0(t) \cdot \exp\{\beta_{PH}' x\}$, where $\begin{cases} \beta_{PH} = -\delta\beta \\ h_0(t) = \delta \cdot t^{\delta-1} \end{cases}$ $\Rightarrow$ It satisfies the PH model. #

2. [+1] Write down the score function.

$\frac{\partial}{\partial w_i} \log f_0(w_i) = 1 - e^{w_i}$

$\frac{\partial^2}{\partial w_i^2} \log f_0(w_i) = -e^{w_i}$

$\frac{\partial}{\partial w_i} \log S_0(w_i) = -e^{w_i}$

$\frac{\partial^2}{\partial w_i^2} \log S_0(w_i) = -e^{w_i}$

+1

$f_0(w) = \exp(w - e^w)$. Let $w = \frac{y - u(x)}{b}$, $u(x) = \beta' x$, $y = \log t$, $r = \sum_{i=1}^{n} \delta_i$

$l(\beta, b) = -r \log b + \sum_{i=1}^{n} [\delta_i \log f_0(w_i) + (1-\delta_i) \log S_0(w_i)] = -r \log b + \sum_{i=1}^{n} [\delta_i (w_i - e^{w_i}) - (1-\delta_i) e^{w_i}]$.

$\frac{\partial l(\beta, b)}{\partial \beta} = \frac{-1}{b} \sum_{i=1}^{n} [\delta_i \cdot (1 - e^{w_i}) - (1-\delta_i) \cdot e^{w_i}] x_{ij} = \frac{-1}{b} \sum_{i=1}^{n} [\delta_i - e^{w_i}] \cdot x_{ij}$, $j = 1, \cdots, p$ #

$\frac{\partial l(\beta, b)}{\partial b} = \frac{-r}{b} - \frac{1}{b} \sum_{i=1}^{n} (\delta_i - e^{w_i}) \cdot w_i$ # $-r = \sum_{i=1}^{n} [\delta_i - e^{\hat{w}_i}] \cdot \hat{w}_i$

3. [+5] Write the observed information matrix using $\hat{b}$, $\mathbf{x}_i$ and $\hat{z}_i = \frac{y_i - \hat{\boldsymbol{\beta}}' \mathbf{x}_i}{\hat{b}}$.

+5

$\frac{\partial l(\beta, b)^2}{\partial \beta' \partial \beta} = \frac{-1}{b^2} \sum_{i=1}^{n} e^{w_i} \cdot x_i x_i'$

$\frac{\partial l(\beta, b)^2}{\partial \beta \partial b} = \underbrace{\frac{1}{b^2} \sum [\delta_i - e^{w_i}] x_{ij}}_{0} - \frac{1}{b^2} \sum_{i=1}^{n} e^{w_i} x_{ij} \cdot w_i$

$\frac{\partial l(\beta, b)^2}{\partial b^2} = \frac{r}{b^2} + \frac{2}{b^2} \sum_{i=1}^{n} [\delta_i - e^{w_i}] \cdot w_i - \frac{1}{b^2} \sum_{i=1}^{n} e^{w_i} \cdot w_i^2$ $\overset{let}{\Rightarrow}$ $z_i = w_i$

$$I(\hat{\beta}, \hat{b}) = \frac{1}{\hat{b}^2} \begin{bmatrix} \sum_{i=1}^{n} e^{\hat{z}_i} \cdot x_i x_i' & \sum_{i=1}^{n} e^{\hat{z}_i} \cdot x_{ij} \cdot \hat{z}_i \\ \sum_{i=1}^{n} e^{\hat{z}_i} \cdot x_{ij} \cdot \hat{z}_i & r + \sum_{i=1}^{n} e^{\hat{z}_i} \cdot \hat{z}_i^2 \end{bmatrix}$$ #

ok.

+1 4. [+2] Let $y_p(\mathbf{x})$ be the conditional p-th quantile for Y given $\mathbf{x}$.

Write down the estimate of $\hat{y}_p(\mathbf{x})$.

$p = P(Y \le y|x) = 1 - S_0(y|x) = F_0(y|x)$.

$y_p(x) = u(x) + F_0^{-1}(p) \cdot b$

$= \beta' x + w_p \cdot b$

$\Rightarrow \hat{y}_p(x) = \hat{\beta}' x + w_p \cdot \hat{b}$ #

, where $w_p = F_0^{-1}(p)$ = ?

(-1)

4. [+5] Write the observed information matrix $I(\hat{\boldsymbol{\beta}}, \hat{b})$ (must be simplified well)

+5

$$I(\hat{\beta}, \hat{b}) = -\begin{bmatrix} \frac{\partial \ell(\beta,b)^2}{\partial \beta' \partial \beta} & \frac{\partial \ell(\beta,b)^2}{\partial \beta \partial b} \\ \frac{\partial \ell(\beta,b)^2}{\partial b \partial \beta} & \frac{\partial \ell(\beta,b)^2}{\partial b^2} \end{bmatrix}\Bigg|_{\substack{\beta=\hat{\beta} \\ b=\hat{b}}}$$

$$= -\frac{1}{\hat{b}^2}\begin{bmatrix} \sum_{i=1}^{n}\left[\delta_i \frac{\partial^2}{\partial z_i^2}\log f_0(z_i) + (1-\delta_i)\frac{\partial^2}{\partial z_i^2}\log S_0(z_i)\right]\Big|_{(z_i=\hat{z}_i)} \cdot x_{ij}\cdot x_{ik} & \sum_{i=1}^{n}\left[\delta_i \frac{\partial^2}{\partial z_i^2}\log f_0(z_i) + (1-\delta_i)\frac{\partial^2}{\partial z_i^2}\log S_0(z_i)\right]\Big|_{(z_i=\hat{z}_i)} \cdot x_{ij}\cdot \hat{z}_i \\ \sum_{i=1}^{n}\left[\delta_i \frac{\partial^2}{\partial z_i^2}\log f_0(z_i) + (1-\delta_i)\frac{\partial^2}{\partial z_i^2}\log S_0(z_i)\right]\Big|_{(z_i=\hat{z}_i)} x_{ij}\cdot \hat{z}_i & r+\sum_{i=1}^{n}\left[\delta_i \frac{\partial^2}{\partial z_i^2}\log f_0(z_i) + (1-\delta_i)\frac{\partial^2}{\partial z_i^2}\log S_0(z_i)\right]\Big|_{(z_i=\hat{z}_i)} \cdot \hat{z}_i^2 \end{bmatrix}$$

#

+3

Q2 [+5] Consider left-truncated data u_i, t_i, subject to $u_i \le t_i$ for $i = 1, \ldots, n$.

Assume the Weibull lifetime model $\Pr(T \ge t) = \exp\left\{-\left(\frac{t}{\alpha}\right)^{\beta}\right\}$, where β is known.

Derive the MLE (must be simplified).

$$\left[\begin{array}{l} S(u) = \exp\{-(\frac{u}{\alpha})^{\beta}\} \\ f(t) = \frac{-d}{dt}S(t) = \exp\{-(\frac{t}{\alpha})^{\beta}\}\cdot\frac{\beta}{\alpha}(\frac{t}{\alpha})^{\beta-1} \end{array}\right]$$

$$L(\alpha) = \prod_{i=1}^{n}\frac{f(t_i;\alpha)}{S(u_i;\alpha)}$$

$$\ell(\alpha) = \sum_{i=1}^{n}\left[\log f(t_i;\alpha) - \log S(u_i;\alpha)\right].$$

$$= \sum_{i=1}^{n}\left[-\left(\frac{t_i}{\alpha}\right)^{\beta} + \log\beta - \log\alpha + (\beta-1)[\log t_i - \log\alpha] + \left(\frac{u_i}{\alpha}\right)^{\beta}\right]$$

$$\ell'(\alpha) = \sum_{i=1}^{n}\left(\beta\left(\frac{t_i}{\alpha}\right)^{\beta-1}\cdot\frac{t_i}{\alpha^2} + \frac{1}{\alpha} - \frac{\beta-1}{\alpha} - \beta\left(\frac{u_i}{\alpha}\right)^{\beta-1}\cdot\frac{u_i}{\alpha^2}\right)$$

$$= \frac{n}{\alpha} - \frac{n(\beta-1)}{\alpha} + \sum_{i=1}^{n}\left[\frac{\beta}{\alpha^{\beta+1}}t_i^{\beta} - \frac{\beta}{\alpha^{\beta+1}}u_i^{\beta}\right]$$

$$= \frac{n}{\alpha} - \frac{n(\beta-1)}{\alpha} + \frac{\beta}{\alpha^{\beta+1}}\sum_{i=1}^{n}(t_i^{\beta} - u_i^{\beta}) \overset{\text{set}}{=} 0$$

+3

$$\Rightarrow \hat{\alpha} = \left[\frac{\beta\sum_{i=1}^{n}(t_i^{\beta} - u_i^{\beta})}{n(\beta-2)}\right]^{\frac{1}{\beta}}$$ #

close!