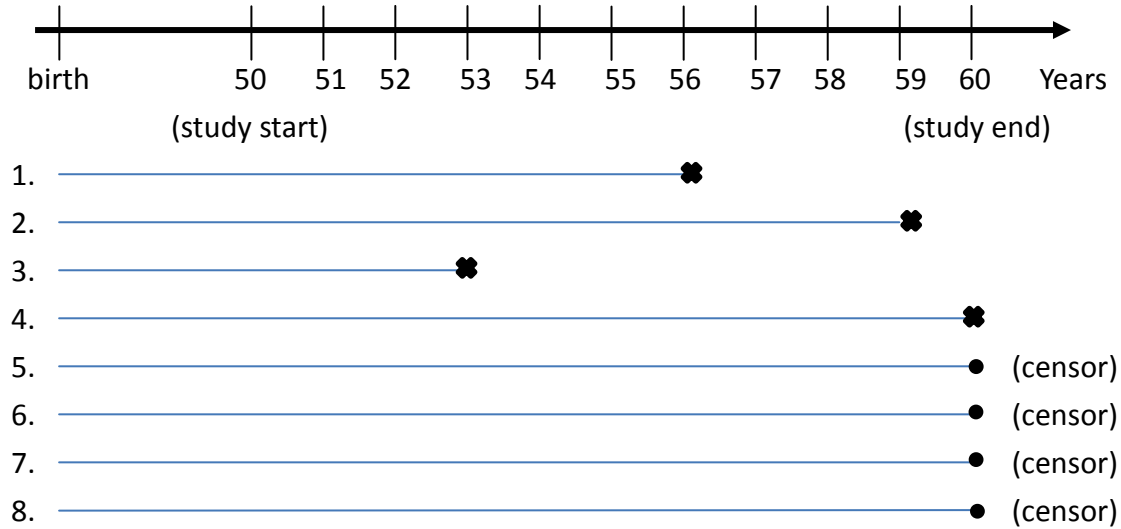


**Chapter 3, Exercise 3.7 (p.89); Appendix A Example A.1 (p.444)**

3.7 (a)

To estimate the distribution of the ages at which postmenopausal woman develop breast cancer. The samples of eight 50-year-old women were given yearly mammograms for a period of 10 years.



✱ : the presence of a tumor   ● : absence of a tumor  
 (55, 56], (58, 59], (52, 53], (59, 60],  $\geq 60, \geq 60, \geq 60, \geq 60$ .

Because study start is fixed

If some woman death before 50 years old, then she can't enter the study

(i) Left truncated occurs

Because onset time of breast cancer is that it occurs between examinations.

(ii) Interval censor occurs

If some woman presence of a tumor after 60 years old, then we only know  $X > 60$

(iii) Right censor occurs

3.7 (b)  $X$  : the age at which breast cancer develops  $X \sim Weib(\alpha, \lambda)$

$$L(\alpha, \lambda) = \Pr(55 < X \leq 56 | X \geq 50) \Pr(58 < X \leq 59 | X \geq 50) \Pr(52 < X \leq 53 | X \geq 50) \times \\ \Pr(59 < X \leq 60 | X \geq 50) \Pr(X > 60 | X \geq 50)^4 \\ = \left\{ \frac{S(55) - S(56)}{S(50)} \right\} \left\{ \frac{S(58) - S(59)}{S(50)} \right\} \left\{ \frac{S(52) - S(53)}{S(50)} \right\} \left\{ \frac{S(59) - S(60)}{S(50)} \right\} \left\{ \frac{S(60)}{S(50)} \right\}^4$$

,where  $S(t) = \exp(-\lambda t^\alpha)$ , survival function of  $X$

$$= \{ \exp(-\lambda 55^\alpha) - \exp(-\lambda 56^\alpha) \} \{ \exp(-\lambda 58^\alpha) - \exp(-\lambda 59^\alpha) \} \times \\ \{ \exp(-\lambda 52^\alpha) - \exp(-\lambda 53^\alpha) \} \{ \exp(-\lambda 59^\alpha) - \exp(-\lambda 60^\alpha) \} \{ \exp(-\lambda 60^\alpha) \}^4 / \\ \{ \exp(-\lambda 50^\alpha) \}^8$$

3.8  $X$  : time to death  $X \sim \text{exp}(\lambda)$   $C$  : right-censoring  $C \sim \text{exp}(\theta)$  ,  $X \perp C$

Let  $T = \min(X, C)$  and  $\delta = \begin{cases} 1 & \text{if } X \leq C \\ 0 & \text{if } X > C \end{cases}$

$$\begin{aligned} \text{(a) } \Pr(\delta = 1) &= \Pr(X \leq C) = \int_0^{\infty} \int_x^{\infty} \lambda \exp(-\lambda x) \theta \exp(-\theta c) dc dx \\ &= \int_0^{\infty} \int_x^{\infty} \lambda \exp(-\lambda x) \theta \exp(-\theta c) dc dx = \int_0^{\infty} \lambda \exp(-\lambda x) \{ \exp(-\theta x) \} dx \\ &= \lambda \int_0^{\infty} \exp \{ -(\lambda + \theta)x \} dx \\ &= \frac{\lambda}{\lambda + \theta} \end{aligned}$$

$$\Rightarrow \delta \sim \text{bernulli} \left( \frac{\lambda}{\lambda + \theta} \right)$$

$$\begin{aligned} \text{(b) } \Pr(T \leq t) &= \Pr(\min(X, C) \leq t) = 1 - \Pr(\min(X, C) > t) \\ &= \Pr(\min(X, C) \leq t) = 1 - \Pr(X > t, C > t) \quad \because X \perp C \\ &= 1 - \Pr(X > t) \Pr(C > t) \\ &= 1 - \exp(-\lambda t) \exp(-\theta t) \\ &= 1 - \exp \{ -(\lambda + \theta)t \} \end{aligned}$$

$$T \sim \text{exp}(\lambda + \theta)$$

(c)

$$F_{T,\delta}(t,1) = F_T(t) = F_T(t)G(1),$$

$$\text{where } G(1) = \Pr(\delta \leq 1) \quad \delta \sim \text{bernulli} \left( \frac{\lambda}{\lambda + \theta} \right)$$

$$F_{T,\delta}(t,0) = \int_0^t S(x)g(x)dx,$$

where  $S(t)$  is survival function of  $X$  ,  $g(t)$  is density of  $C$

$$\begin{aligned} &= \int_0^t \exp(-\lambda x) \theta \exp(-\theta x) dx = \theta \int_0^t \exp \{ -(\lambda + \theta)x \} dx \\ &= \frac{\theta}{\lambda + \theta} [1 - \exp \{ -(\lambda + \theta)t \}] \\ &= G(0)F_T(t), \end{aligned}$$

where  $G(0) = \Pr(\delta \leq 0)$

$$\because F_{T,\delta}(t,0) = F_T(t)G(0) \text{ and } F_{T,\delta}(t,1) = F_T(t)G(1)$$

$\therefore T \perp \delta$

$$(d) T \sim \exp(\lambda + \theta) \quad , \quad \delta \sim \text{bernulli}\left(\frac{\lambda}{\lambda + \theta}\right)$$

$$f(t)g(\delta) = \left(\frac{\lambda}{\lambda + \theta}\right)^\delta \left(\frac{\lambda}{\lambda + \theta}\right)^{1-\delta} (\lambda + \theta) \exp\{-(\lambda + \theta)t\}$$

$$\begin{aligned} \Rightarrow l(\alpha) &= \log L(\alpha) = \sum_i \delta_i \log\left(\frac{\lambda}{\lambda + \theta}\right) + \sum_i (1 - \delta_i) \log\left(\frac{\theta}{\lambda + \theta}\right) + n \log(\lambda + \theta) - \sum_i (\lambda + \theta)t_i \\ &= \sum_i \delta_i \log(\lambda) + n \log\left(\frac{\theta}{\lambda + \theta}\right) - \sum_i \delta_i \log(\theta) + n \log(\lambda + \theta) - \sum_i (\lambda + \theta)t_i \\ &= \sum_i \delta_i \log(\lambda) + n \log(\theta) - \sum_i \delta_i \log(\theta) - \sum_i (\lambda + \theta)t_i \end{aligned}$$

$$l'(\alpha) = \frac{\sum_i \delta_i}{\lambda} - \sum_i t_i = 0$$

$$\Rightarrow \hat{\lambda} = \frac{\sum_i \delta_i}{\sum_i t_i}$$

$$l''(\alpha) = -\frac{\sum_i \delta_i}{\lambda^2} \leq 0$$

$$\begin{aligned} E(\hat{\lambda}) &= E\left(\frac{\sum_i \delta_i}{\sum_i T_i}\right) = E\left(\sum_i \delta_i\right) E\left(\frac{1}{\sum_i T_i}\right) \quad \because \delta \perp T \\ &= n \frac{\lambda}{\lambda + \theta} E\left(\frac{1}{S}\right) \quad , \quad S = \sum_i T_i \end{aligned}$$

$$\because T_1, T_2, \dots, T_n \sim \exp(\lambda + \theta) \Rightarrow \sum_i T_i \sim \text{Gamma}(n, \lambda + \theta)$$

$$\begin{aligned} E\left(\frac{1}{S}\right) &= \int_0^\infty s^{-1} \frac{(\lambda + \theta)^n}{\Gamma(n)} s^{n-1} \exp\{s(\lambda + \theta)\} ds = \frac{(\lambda + \theta)^n}{\Gamma(n)} \frac{\Gamma(n-1)}{(\lambda + \theta)^{n-1}} \int_0^\infty \frac{(\lambda + \theta)^{n-1}}{\Gamma(n-1)} s^{(n-1)-1} \exp\{s(\lambda + \theta)\} ds \\ &= \frac{(\lambda + \theta)}{n-1} \end{aligned}$$

$$\Rightarrow E(\hat{\lambda}) = n \frac{\lambda}{\lambda + \theta} \frac{(\lambda + \theta)}{n-1} = \frac{n\lambda}{n-1}$$

### Example A.1

$X \sim \text{Weibull}(\alpha, \lambda = 1)$

To find the maximum likelihood estimator of  $\alpha$

Data: 2.57, 0.58, 0.82, 1.02, 0.78, 0.46, 1.04, 0.43, 0.69, 1.37

$$f(x; \alpha) = \alpha x^{\alpha-1} \exp(-x^\alpha)$$

$$l(\alpha) = \log L(\alpha) = n \log \alpha + (\alpha - 1) \sum_i \log(x_i) - \sum_i x_i^\alpha$$

$$l'(\alpha) = \frac{n}{\alpha} + \sum_i \log(x_i) - \sum_i x_i^\alpha \log(x_i) = 0$$

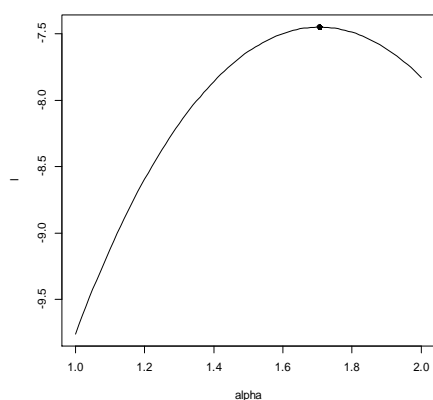
$$l''(\alpha) = -\frac{n}{\alpha^2} - \sum_i x_i^\alpha \{\log(x_i)\}^2 < 0$$

- Bisection method:

Algorithm stop, when  $|l'(\alpha)| < 0.01$

step	$\alpha_L$	$\alpha_U$	$\alpha_N$	$l'(\alpha_L)$	$l'(\alpha_U)$	$l'(\alpha_N)$
1	1.500000	2.000000	1.750000	1.797531	-2.588883	-0.386642
2	1.500000	1.750000	1.625000	1.797531	-0.386642	0.6966227
3	1.625000	1.750000	1.687500	0.6966227	-0.386642	0.1542948
4	1.687500	1.750000	1.718750	0.6966227	-0.1161722	-0.1161722
5	1.687500	1.718750	1.703125	0.01904024	-0.1161722	0.01904024
6	1.703125	1.718750	1.710938	0.01904024	-0.04856852	-0.04856852
7	1.703125	1.710938	1.707031	0.01904024	-0.01476511	-0.01476511
8	1.703125	1.707031	1.705078	0.002137278	-0.01476511	0.002137278

$\Rightarrow \hat{\alpha}_{MLE} = 1.705078$



● secant method:

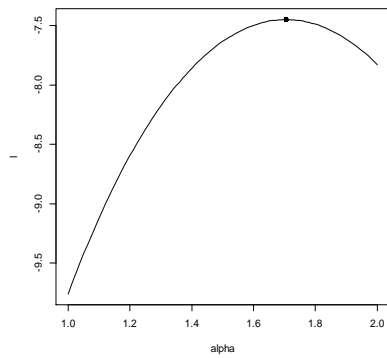
Iteration:

$$\alpha_{i+1} = \alpha_i - l'(\alpha_i) \frac{(\alpha_i - \alpha_{i-1})}{l'(\alpha_i) - l'(\alpha_{i-1})}$$

stop ,when  $|l'(\alpha)| < 0.01$

step	$\alpha_{i-1}$	$\alpha_i$	$l'(\alpha_{i-1})$	$l'(\alpha_i)$	$\alpha_{i+1}$	$l'(\alpha_{i+1})$
1	1.000000	1.500000	7.065037	1.797531	1.670624	0.300472
2	1.500000	1.670624	1.797531	0.300472	1.704870	0.003936636

$\Rightarrow \hat{\alpha}_{MLE} = 1.704870$



● Newton - Raphson:

Iteration:

$$\alpha_{i+1} = \alpha_i - \frac{l'(\alpha_i)}{l''(\alpha_i)}$$

stop ,when  $|l'(\alpha)| < 0.01$

step	$\alpha_{i-1}$	$l'(\alpha_{i-1})$	$l'(\alpha_i)$	$\alpha_i$	$l'(\alpha_{i+1})$
1	1.500000	1.797531	-8.946896	1.700911	0.03820069
2	1.700911	0.03820069	-8.654888	1.705325	1.759597e-06

$\Rightarrow \hat{\alpha}_{MLE} = 1.705325$

