defn

- $\tilde{T}_i$: survival time (event of interest) for i^{th} individual ☆
- T_i: observed ~~surv~~ ~~ev~~ time
- C_i: censoring time.

survival function

$$\cancel{S(t) = P(\tilde{T}_i > t)}$$
$$\cancel{F(t) = P(\tilde{T}_i < t)} \Rightarrow \cancel{S(t)}$$

$$S(t) = P(\tilde{T} > t)$$
$$F(t) = P(\tilde{T} < t) \text{ [by defn]} \Rightarrow S(t) = 1 - F(t)$$

hazard function

- continuous case

$$\lambda(t) = \lim_{\delta \to 0} \frac{P(t \le \tilde{T} < t + \delta \mid \tilde{T} > t)}{\delta} \quad \approx P(\tilde{T} = t)$$

$$= \lim_{\delta \to 0} \frac{P(t \le \tilde{T} < t + \delta)}{\delta \cdot P(\tilde{T} > t)}$$

$$= \lim_{\delta \to 0} \frac{F(t+\delta) - F(t)}{\delta} \cdot \frac{1}{S(t)}$$

$$= \frac{f(t)}{S(t)} = \frac{f(t)}{1 - F(t)} \quad [-\ln(1-F(t))] \text{ 求导}$$

$\therefore$ $\cancel{\lambda(t)} = \cancel{-\ln[1-F(t)]} = \cancel{-\ln S(t)}$

$\therefore$ ~~exp{-~~

integrate both sides:

$$\underbrace{\int_0^t \lambda(u)\,du}_{①} = \int_0^t \frac{f(u)}{1-F(u)}\,du = -\ln(1-F(u))\Big|_0^t = -\ln S(u)\Big|_0^t$$

$$= -\ln S(t) + \ln S(0) \overset{S(0)=1}{=} \underbrace{-\ln S(t)}_{②} \Rightarrow ① = ②$$

• Discrete case [$h(x_j)$ is a conditional prob.]

$$h(x_j) = P(X = x_j \mid X > x_{j-1}) \quad j-1: \text{alive},\ j: \text{event}$$

$$= \frac{P(X = x_j)}{P(X > x_{j-1})} = \frac{F(x_j) - F(x_{j-1})}{S(x_{j-1})} = \frac{[1 - S(x_j)] - [1 - S(x_{j-1})]}{S(x_{j-1})}$$

$$= \frac{S(x_{j-1}) - S(x_j)}{S(x_{j-1})} = 1 - \frac{S(x_j)}{S(x_{j-1})}$$

$$1 - h(x_j) = \frac{S(x_j)}{S(x_{j-1})} \quad ①$$

As $S(x) = \prod_{x_j \le x} \frac{S(x_j)}{S(x_{j-1})} = \frac{S(x_j)}{S(x_{j-1})} \cdot \frac{S(x_{j-1})}{S(x_{j-2})} \cdots \frac{S(x_1)}{\cancel{S(x_0)}}$ $\quad S(x_0) = 1$

$\therefore \cancel{S(x)} = \cancel{\prod_{x_j \le x}}$

multiply both side every time point till x.

$$\prod_{x_j \le x} (1 - h(x_j)) = \prod_{x_j \le x} \frac{S(x_j)}{S(x_{j-1})} = S(x)$$

~~intuitively, survival till time x means~~

The likelihood

- what we observed $T_i = \min\{\tilde{T}_i, C_i\}$ i^{th}
- what we're interested in. $\tilde{T}_i$, $\lambda(t)$.

↳ if E_i event occurs, denoted by $E_i = 1$.

$$p(t_i \le T_i < t_i + dt, E_i = 1 \mid X_i, \theta) \Rightarrow \text{what's been observed}$$

⇒ i.e., $\tilde{T}_i \in (t_i, t_i + dt)$ $\quad C_i > t_i$

$$= \underbrace{p(C_i > t_i)}_{\text{under independent censoring}} \; p(t_i \le \tilde{T}_i < t_i + dt \mid X_i, \theta)$$

censoring dist $= p(C_i > t_i) \cdot \lambda_i(t_i;\theta)\, dt \cdot S_i(t_i;\theta)$

$= F(t_i + dt) - F(t_i)$ $\quad = f(t)\,dt$

$= \dfrac{[F(t_i+dt) - F(t_i)] \cdot dt}{dt}$ $\quad = \lambda(t) S(t) \cdot dt$ ☆

(Note: we assume censoring time does not involve θ. i.e., informative censoring mechanism)

↳ if censored, $E_i = 0$

$$p(t_i \le T_i < t_i + dt, E_i = 0 \mid X_i, \theta) \Rightarrow \text{observed censoring}$$

⇒ censored @ $(t_i, t_i + dt)$, & $\tilde{T}_i > t_i$

$$p(t_i \le C_i < t_i + dt)\, p(\tilde{T}_i > t_i \mid X_i; \theta)$$

$$= p(t_i \le C_i < t_i + dt)\, S_i(t_i; \theta)$$

[Transfer observed time → event of interest, which is $\tilde{T}$, parametrized by $\lambda(t_i;\theta)$ $S_i(t_i;\theta)$]

Total likelihood

$$L(\theta) = \prod_{i=1}^{n} p(t_i \le T_i \le t_i + dt, E_i = e_i \mid X_i; \theta)$$

$$\propto \prod_{i=1}^{n} \left[\lambda_i(t_i;\theta)^{e_i} \cdot S_i(t_i;\theta)\right] \quad ☆$$

$$\ell(\theta) = \sum_{\text{uncensored/event}} \log \lambda_i(t_i;\theta) + \sum_{\text{all}} \log S_i(t_i;\theta)$$

likelihood for discrete ~~variable~~ time

$L(\theta) = \lambda_i(t_i;\theta)^{e_i} \cdot S_i(t_i;\theta)$

Let look @ a toy example

individual	t_1	t_2	t_3	t_4
1	$(1-h_1)$	$(1-h_2)$	h_3	
2	h_1			
3	$(1-h_1)$	h_2		
4	$(1-h_1)$	$(1-h_2)$	$(1-h_3)$	$(1-h_4)$

individual 1

$\therefore L(\theta) = \cancel{(1-h_1)}\, h_3 \cdot \underbrace{(1-h_1)\cdot(1-h_2)}$ $[S(t) = \prod_{x_j \le x} (1-h(x_j))]$

2 → h_1

3 → h_2 $\cancel{(1-h_1)}\,(1-h_1)$

4 → $(1-h_1)(1-h_2)(1-h_3)(1-h_4)$

rearrange by time ⟹ column

$h_1 \cdot h_2 \cdot h_3\, (1-h_1)(1-h_1)(1-h_1)\cdot(1-h_2)\cdot(1-h_2)$
$\cdot(1-h_3)\cdot(1-h_4)$

∴ if we multiply by each individual [row-based] is actually equal to that if we look by columns coz the likelihood is the multiplication of all cell values. It's easier to write out the likelihood by column

$$\prod_{x_j \le x} (h_j)^{d_j} (1-h_j)^{n_j - d_j}$$