(Ch.2) Terminology and Notation

Li-Hsin Chien

September 18, 2024

1 Terminologies and Notations

X: non-negative random variable, failure time, survival time

The probability distribution of X can be specified in three ways (useful in survival applications):

- Survival function: $S(x)$
- Probability density/mass function: $f(x)/p(x)$
- Hazard function: $h(x)$

2 X continuous

$S(x)$: survival function, non-increasing right-continuous function of x, $S(0) = 1$

$$S(x) = Pr(X > x), 0 < x < \infty$$

$f(x)$: probability density function (pdf) of X

$$f(x) = -dS(x)/dx$$

$$S(x) = \int_x^\infty f(s)ds$$

$h(x)$: hazard function

$$\begin{aligned} h(x) &= \lim_{h\to 0+} \frac{Pr(x \leq X < x+h | X \geq x)}{h} \\ &= \frac{f(x)}{S(x)} \\ &= -\frac{d \log S(x)}{dx} \end{aligned}$$

[1] $H(x)$: cumulative hazard function

$$H(x) = \int_0^x h(s)ds$$

[1] $\frac{Pr(x\leq X<x+h|X\geq x)}{h} = \frac{Pr(x\leq X<x+h, X\geq x)}{hPr(X\geq x)} = \frac{Pr(x\leq X<x+h)}{h}\frac{1}{Pr(X\geq x)}$

$$S(x) = \exp\left[-H(x)\right] = \exp\left[-\int_0^x h(s)ds\right]$$

$$f(x) = h(x)\exp\left[-H(x)\right]$$

Ex.1: X follows the exponential distribution with parameter λ. $X \sim Exp(\lambda)$

$$f(x) = \lambda\exp(-\lambda x), \quad x \geq 0$$

$$\begin{aligned} S(x) &= \int_x^\infty \lambda\exp(-\lambda s)ds \\ &= -\exp(-\lambda s)|_x^\infty \\ &= \exp(-\lambda x) \end{aligned}$$

$$h(x) = \frac{f(x)}{S(x)} = \lambda$$

3 X discrete

X: discrete random variable taking values $x_1, x_2, \cdots$ with associated probability mass function

$$p(x_i) = Pr(X = x_i), \qquad i = 1, 2, \cdots,$$

$$S(x) = Pr(X > x) = \sum_{x_j > x} p(x_j)$$

The hazard at x_i is

$$\begin{aligned} h(x_i) &= Pr(X = x_i | X \geq x_i) \\ &= \frac{p(x_i)}{S(x_i^-)} = \frac{p(x_i)}{S(x_{i-1})} \\ &= \frac{S(x_{i-1}) - S(x_i)}{S(x_{i-1})} \\ &= 1 - \frac{S(x_i)}{S(x_{i-1})}, \qquad i = 1, 2, \cdots, \end{aligned}$$

$S(x_0) = 1$,

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

Ex.2: $S(x_3) = \frac{S(x_3)}{S(x_2)}\frac{S(x_2)}{S(x_1)}\frac{S(x_1)}{S(x_0)} = [1 - h(x_3)][1 - h(x_2)][1 - h(x_1)]$

Ex.3: The lifetime X has the pmf $p(x_j) = Pr(X = j) = 1/3, \quad j = 1, 2, 3.$

$$S(x) = Pr(X > x) = \sum_{x_j > x} p(x_j) = \begin{cases} 1, & \text{if } 0 \leq x < 1, \\ 2/3, & \text{if } 1 \leq x < 2, \\ 1/3, & \text{if } 2 \leq x < 3, \\ 0, & \text{if } x \geq 3. \end{cases}$$

$$h(x) = \begin{cases} \frac{1/3}{1} = 1/3, & \text{if } x = 1, \\ \frac{1/3}{2/3} = 1/2, & \text{if } x = 2, \\ \frac{1/3}{1/3} = 1, & \text{if } x = 3, \\ 0, & \text{otherwise.} \end{cases}$$

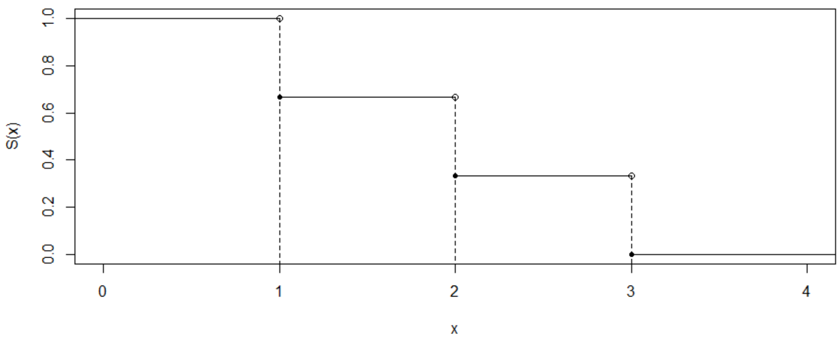

Figure 1: Survival curve(right continuous)

4 The Mean Residual Life Function and Median Life

Mean residual life

The mean residual life (expected residual life) at time x: $\text{mrl}(x) = E(X - x|X > x)$

For a continuous random variable:

$$\text{mrl}(x) = \frac{\int_x^\infty (t-x)f(t)dt}{S(x)} = \frac{\int_x^\infty S(t)dt}{S(x)}.$$

[2]

When $x = 0$, $\text{mrl}(0) = E(X)$

$$\mu = E(X) = \int_0^\infty tf(t)dt = \int_0^\infty S(t)dt.$$

$$Var(X) = 2\int_0^\infty tS(t)dt - \left[\int_0^\infty S(t)dt\right]^2$$

[3]

For a discrete random variable:

$$\text{mrl}(x) = \frac{(x_{i+1} - x)S(x_i) + \sum_{j \geq i+1}(x_{j+1} - x_j)S(x_j)}{S(x)}, \text{ for } x_i \leq x < x_{i+1}$$

[2]Use the integration by part: $\int_x^\infty (t-x)f(t)dt = \int_x^\infty S(t)dt - (t-x)S(t)|_x^\infty$. Note that $lim_{t\to\infty} tS(t) = 0$ since $E(X) < \infty$.

[3]Use the integration by part: $\int_x^\infty 2tS(t)dt = \int_x^\infty t^2 f(t)dt - t^2 S(t)|_x^\infty$

Median life

The 0.5th quantile (median life) is found by solving the equation $S(x_{0.5}) = 0.5$.

$$S(x_{0.5}) \leq 0.5, \quad \text{i.e.} \quad x_{0.5} = \inf\{t : S(t) \leq 0.5\}$$

Homework 1.(Ch.2)

1. (# 2.13) Suppose that X has a geometric distribution with probability mass function

$$p(x) = p(1-p)^{x-1}, \quad x = 1, 2, \cdots$$

   (a) Find the survival function of X. (Hint: Recall that for $0 < \theta < 1, \sum_{j=k}^{\infty} \theta^j = \theta^k/(1-\theta)$)
   (b) Find the hazard rate of X. Compare this rate to the hazard rate of an exponential distribution.

2. (# 2.12) Let X have a uniform distribution on the interval 0 to θ with density function

$$f(x) = \begin{cases} \frac{1}{\theta}, & \text{for } 0 \leq x \leq \theta \\ 0, & \text{otherwise} \end{cases}$$

   (a) Find the survival function of X.
   (b) Find the hazard rate of X.
   (c) Find the mean residual-life function.