JKUAT DEPARTMENT OF STATISTICS AND ACTUARIAL SCIENCES

STA 2433/2493 SURVIVAL ANALYSIS B SC ACTUARIAL SCIENCE CAT 1

MAY 2018

1. (a) The discrete random variable T has times to failure given by $T = 1,2,3,\ldots$ with corresponding survival function S_t and probability mass function f_t for $t = j, j \varepsilon W$. Show that the expected time to failure $E(T)$ can be given by $\sum_{j>0} S_j$.

   (b) Discrete random variable X has PMF f(x) as given below. Determine correct to two decimal places in each case the survival function S(x) for the random variable X, the hazard function h(x) for the random variable X, and E(X).

x	0	1	2	3	4	5
f(x)	0.10	0.15	0.25	0.25	0.15	0.10

2. Continuous random variable X has PDF f(x) given by

$$f(x) = \begin{cases} \cos x & for\ 0 \le x \le \frac{\pi}{2} \\ 0 & elsewhere. \end{cases}$$

   Determine expressions for the survival function, the hazard function and the integrated hazard function.

3. The number of hours T that a certain bacteria survives can be modelled by the survival function S(t) where

$$S(t) = \begin{cases} a - bt^2 & for\ 0 \le t \le \frac{1}{8}, \\ \frac{k}{t^{\frac{5}{3}}} & for\ t > \frac{1}{8}. \end{cases}$$

   (a) It is given that $S\left(\frac{1}{8}\right) = \frac{3}{4}$. Determine the values of the constants k, a and b.

   (b) Hence determine the expectation of life of the bacteria.

4. A trainee in a carpentry workshop is given an assignment of cutting planks of wood and making four joints within a day. The agreement is that as soon as the trainee demonstrates the ability to make four joints within a day there will be an immediate offer of employment. On each day there is a constant probability of 0.8 that the trainee will succeed in this assignment.

   (a) Calculate each of the probabilities $P(X = 1)$, $P(X = 2)$, $P(X = 3)$, and write down the probability $P(X = x)$.

   (b) Determine the survival function and the hazard function for the variable X.

   (c) Determine the expected number of attempts.

   (d) Determine the median number of attempts.

5. A P.E. instructor sends a number of runners out on a cross country race, in which it is expected that every runner will complete the exercises within one hour. The time in hours taken by a runner to complete the race can be modelled by the random variable X which has PDF f(x) given by:

$$f(x) = \begin{cases} \theta x^{\theta - 1} & \text{for } 0 \le x \le 1 \text{ and } \theta \text{ is an unknown parameter} \\ 0 & \text{elsewhere.} \end{cases}$$

(a) Determine expressions for the survival function S(x) and hence the hazard function h(x).

(b) In the sample of n runners' times $(x_1, x_2, x_3, \ldots x_d, x_{d+1}, \ldots\ldots x_n)$ of which the first d times are uncensored, and the remaining times are censored show that the maximum likelihood estimator of θ is a root of the equation

$$\frac{d}{\theta} + \sum_{i=1}^{d} \ln x_i - \sum_{j=d+1}^{n} \frac{x_j^{\theta} \ln x_j}{1 - x_j^{\theta}} = 0 .$$

(c) If there is no censoring show that the maximum likelihood estimator of θ is given by:

$$\hat{\theta} = -\frac{n}{\sum_{j=1}^{n} \ln x_j} .$$

(d) Eight runners took part in the race with times in hours (assume all were uncensored) 0.84, 0.38, 0.79, 0.88, 0.28, 0.95, 0.63,0.61. Determine correct to two decimal places the maximum likelihood estimator of θ .