Ramakrishna Mission Vivekananda Educational and Research Institute
PO Belur Math, Howrah, West Bengal 711 202
School of Mathematical Sciences
Department of Computer Science

MSc BDA : Batch 2022-24, Semester II, Final Exam
DA312: Time Series & Survival Analysis
Dr. Sudipta Das

*S*tudent Name (in block letters):   Date: 20 June 2023
*S*tudent Roll No: *M*ax Marks: 100
*S*ignature: Time: 3hrs

Answers must be properly justified to deserve full credits.

1. (25 points)

   Suppose that X has a geometric distribution with probability mass function

   $$P(X = x) = p(x) = p(1 - p)^{x-1}, x = 1, 2, \ldots.$$

   (a) (6 points) Find the survival function of X

   (b) (6 points) Find the hazard rate of X.

   (c) (6 points) Find the mean of X

   (d) (7 points) Find the mean residual-life function.

2. (25 points)

   Forty female breast-cancer patients with negative axillary lymph nodes and a minimum 10-year follow-up were selected from The Ohio State University Hospitals Cancer Registry. Of the 40 patients, 10 were immunoperoxidase positive, and the remaining 30 remained negative. Survival times (in months) for both groups of patients are given in the following (+ denotes a censored observation).

   Immunoperoxidase Negative:

   $$19, 25, 30, 30, 46, 51+, 57, 61, 66, 66+, 74, 78, 86, 122+, 123+, 130+, 130+, 134+,$$

   $$143+, 148+, 151+, 152+, 153+, 154+, 156+, 162+, 164+, 165+, 182+, 189+$$

   Immunoperoxidase Positive:

   $$22, 23, 38, 42+, 73, 77, 89, 115+, 144+, 148+$$

   (a) (18 points) Estimate the survival functions with standard errors for both the groups.

   (b) (7 points) Graphically compare the survival functions of the two groups.

3. (25 points)

One of the goals of recent research is to explore the efficacy of triple-drug combinations of antiretroviral therapy for treatment of HIV- infected patients. Because of limitations on potency and the continuing emergence of drug resistance seen with the use of currently available antiretroviral agents in monotherapy and two-drug regimens, triple- combination regimens should represent a more promising approach to maximize antiviral activity, maintain long-term efficacy, and reduce the incidence of drug resistance. Towards this end, investigators performed a randomized study comparing AZT+ zalcitabine (ddC) versus AZT+ zalcitabine (ddC)+ saquinavir. The data, time from administration of treatment (in days) until the CD4 count reached a prespecified level, is given below for the two groups.

AZT+ zalcitabine (ddC)
$85, 32, 38+, 45, 4+, 84, 49, 240+, 87, 75, 102, 39, 12, 11, 80$

AZT+ zalcitabine (ddC)+ saquinavir
$22, 2, 48, 85, 160, 238, 56+, 94+, 51+, 12, 171, 80, 180, 4, 90$

Use the log rank statistic to test if there is a difference in the distribution of the times at which patient's CD4 reaches the prespecified level for the two treatments. [Given, $\chi^2_{0.05,1} = 3.841459$]

4. (25 points)

A study was conducted on the effects of ploidy on the prognosis of patients with cancer of the tongue. Tissue samples were examined to determine if the tumor had a aneuploid or diploid DNA profile. Death times (in weeks) of patients with cancer of the tongue for these two groups of patients are given below.

Aneuploid Tumors:
Death Times:
1, 3, 3, 4, 10, 13, 24, 27, 30, 41, 56, 62, 70, 72, 74, 89, 93, 96, 101, 104
Censored Observations:
67, 74, 79, 81, 88, 89, 97, 101, 104, 108, 170, 400

Diploid Tumors:
Death Times:
3, 4, 13, 27, 30, 41, 56, 62, 129, 181
Censored Observations:
67, 79, 170, 231

To analyze this data create a single indicator variable, $Z = 0$, for Aneuploid tumors and $Z = 1$, for Diploid tumors.

(a) (6 points) Find the contribution of the deaths at time 3 to the likelihood function using the Breslow method of handling ties.

(b) (6 points) Find the Breslow's estimator of the baseline cumulative hazard rate at time 3.

(c) (6 points) Considering both the aneuploid and diploid groups construct the combined likelihood function assuming Cox proportional hazard model with a constant baseline hazard λ. You may denote the coefficient of the regressor as β.

(d) (7 points) Hence, find the maximum likelihood estimate of λ and β.

This exam has total 4 questions, for a total of 100 points and 0 bonus points.

Best of luck!!