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 2021-23, Semester II, Final Exam
DA312: Time Series & Survival Analysis
Dr. Sudipta Das

Student Name (in block letters):	Date: 14 June 2022
Student Roll No:	Max Marks: 100
Signature:	Time: 3hrs

Answers must be properly justified to deserve full credits.

1. (20 points)

   Given a covariate Z, suppose that the log survival time Y follows a linear model with a logistic error distribution, that is,
   $$Y = \ln(X) = \mu + \beta Z + \sigma W$$
   where the pdf of W is given by
   $$f_W(w) = \frac{e^w}{(1+e^w)^2}, -\infty < w < \infty.$$

   (a) (7 points) For an individual with covariate Z, find the conditional survival function of the survival time X, given Z, namely, $S(x|Z)$.

   (b) (3 points) The odds that an individual will die prior to time x is expressed by $[1 - S(x|Z)]/S(x|Z)$. Compute the odds of death prior to time x for this model.

   (c) (10 points) Consider two individuals with different covariate values. Show that, for any time x, the ratio of their odds of death is independent of x.

2. (20 points)

   To estimate the distribution of the ages at which postmenopausal woman develop breast cancer, a sample of eight 50-year-old women were given yearly mammograms for a period of 10 years. At each exam, the presence or absence of a tumor was recorded. In the study, no tumors were detected by the women by self-examination between the scheduled yearly exams, so all that is known about the onset time of breast cancer is that it occurs between examinations. For four of the eight women, breast cancer was not detected during the 10 year study period. The age at onset of breast cancer for the eight subjects was in the following intervals:
   $$(55, 56], (58, 59], (52, 53], (59, 60], \geq 60, \geq 60, \geq 60, \geq 60.$$

   (a) (4 points) What type of censoring or truncation is represented in this sample?

   (b) (7 points) Assuming that the age at which breast cancer develops follows a Exponential distribution with mean parameter μ, construct the likelihood function.

   (c) (9 points) Assuming that the age at which breast cancer develops follows a Weibull distribution with survival function as $e^{-(\lambda x)^\alpha}$, construct the likelihood function.

3. (20 points)

   In a study to assess the time to first exit-site infection (in months) in patients with renal insufficiency, 40 patients utilized a surgically placed catheter. Times to infection (in months) of the patients appears in the following.

   Death Times:
   $$1.5, 3.5, 4.5, 4.5, 5.5, 8.5, 8.5, 8.5, 10.5, 11.5, 15.5, 15.5, 18.5, 26.5$$

Censored Observations:

$2.5, 2.5, 3.5, 3.5, 3.5, 4.5, 5.5, 6.5, 6.5, 7.5, 7.5, 7.5, 7.5, 8.5, 9.5, 10.5, 11.5, 12.5, 12.5, 13.5, 14.5, 14.5, 22.5, 22.5, 25.5, 27.5$

(a) (8 points) Estimate the survival function at 6 months.

(b) (8 points) Estimate the cumulative hazard rate, $H(t)$, at 6 months.

(c) (4 points) Estimate $S(6)$ by a suitable function of $\tilde{H}(t)$ and compare to your estimate in part a.

4. (20 points)

Death times (in weeks) of patients with cancer of the tongue for two types of tumors are given below.

Aneuploid Tumors:
Death Times: $1, 3, 3, 10, 16, 24, 26, 30, 41, 51, 67, 73, 93, 96, 104, 157$
Censored Observations: $61, 74, 80, 81, 87, 87, 89, 97, 101, 120, 150, 231$

Diploid Tumors:
Death Times: $1, 3, 4, 5, 5, 12, 13, 18, 23, 26, 42, 62, 104, 104, 112$
Censored Observations: $8, 67, 76, 104, 176$

Test the hypothesis that the survival rates of patients with cancer of the tongue are the same for patients with aneuploid and diploid tumors using the log-rank test.

5. (24 points)

Let T be a continuous survival time random variable. The survival and hazard functions of T at time t are given as $S(t)$ and $h(t)$. The mean of T is denoted as μ, also. If $h(t)$ is monotonically increasing with t,

(a) (12 points) then show that $[S(t)]^{\frac{1}{t}}$ is decreasing in t and

(b) (12 points) show that $S(t) \geq e^{-\frac{t}{\mu}}$ for $t < \mu$.

This exam has total 5 questions, for a total of 100 points and 0 bonus points.

Best of luck!!