Inspiring Innovation and Leadership

KARATINA UNIVERSITY

UNIVERSITY EXAMINATIONS
2024/2025 ACADEMIC YEAR

FOURTH YEAR FIRST SEMESTER EXAMINATION

FOR THE DEGREE OF

BACHELOR OF SCIENCE IN ACTUARIAL SCIENCE

COURSE CODE: ACS 412

COURSE TITLE: SURVIVAL MODELS

DATE: 9TH DECEMBER 2024 TIME: 12 – 2 P.M.

INSTRUCTION TO CANDIDATES

- ANSWER **ALL QUESTIONS** IN SECTION A
- ANSWER **ANY TWO** QUESTIONS FROM SECTION B

SECTION A

Answer ALL QUESTIONS in this section

QUESTION ONE (30 marks)

(a) Define the following terms as used in the analysis of survival data:

(i) Censoring [2 marks]

(ii) Hazard function [2 marks]

(iii) Median Survival [2 marks]

(b) Differentiate between the following terms:

(i) Informative and non - informative censoring [2 marks]

(ii) Right and left censoring [2 marks]

(c) The Weibull survival distribution can be characterized by hazard function $\lambda(t) = \lambda \gamma t^{\gamma-1}$.

Determine the:

(i) Survival function for the Weibull distribution. [3 marks]

(ii) Probability density function for the Weibull distribution. [2 marks]

(d) A clinical trial to evaluate the efficacy of maintenance chemotherapy for Acute Mylogenous Leukemia (AML) was conducted. The following data shows times (weeks) of remission (i.e. freedom from symptoms in a precisely defined sense) of AML patients received chemotherapy:

9, 13, 13+, 18, 23, 28+, 31, 34, 45+, 48, 161+.

Observations with plus (+) sign are right censored.

(i) Calculate the Kaplan – Meier estimator of the survival function for this treatment group. [6 marks]

(ii) Calculate the 95% confidence interval for the survival function for $\hat{S}(23)$ using the Greenwood's formula. [7 marks]

(iii) Compute the median survival time. [2 marks]

SECTION B

ANSWER ANY TWO QUESTIONS FROM THIS SECTION.

QUESTION TWO (20 marks)

The following life table shows data on time to HIV development for a sample of 100 individuals with STD but free of HIV at time 0.

Year Interval	Number of HIV positive	Number Lost to follow-up
0 – 2	1	1
2 – 4	2	1
4 – 6	8	4
6 – 8	5	8
8 – 10	5	18
10 – 12	3	20
12 – 14	8	16

Compute the life table for this data. [20 marks]

QUESTION THREE (20 marks)

Consider the following data set having two treatment groups (0 and 1).

Death Time	Treatment Group
1	0
2	0
3	0
4	0
5+	0
6	0
1+	1
2	1
3+	1
14	1
15	1
16	1

(a) Compute the logrank CMH – type variance and the logrank CMH – Statistic for the data. [10 marks]

(b) Compute the linear rank – type variance and the linear rank type – Statistic for the data. [10 marks]

QUESTION FOUR (20 marks)

(a) A study was done to assess the effects of treatment on the duration of steroid – induced remission for 42 acute leukaemia patients.

The variable definitions are given below:

Variable name	Description
sex	Gender (1 = male, 2 = female)
logwbc	log (base 10) white blood cell count
trt	Treatment (1 = placebo, 0 = treatment)
survtime	Time to relapse (weeks)
status	Status (0 = censored, 1 = relapse)

The following SAS code was used to fit several models to the data.

Model 1:
```
proc lifetest;
strata sex;
time survtime*status (0);
run;
```

Model 2:
```
proc phreg;
model survtime*status (0) = sex  trt  logwbc ;
run;
```

Model 3:
```
proc phreg;
strata sex;
model survtime*status (0) = trt  logwbc ;
run;
```

Model 4:
```
proc phreg;
model survtime*status (0) = sex  trt  logwbc  sextime;
sextime = log (survtime) * sex;
run;
```

(i) Write out the form of the Cox proportional hazards model for Model 2. [4 marks]

(ii) Using this model, demonstrate that the hazard ratio for males versus females satisfies the assumption of "proportionality". [3 marks]

(iii) Which of the models 1 – 4 would ***not*** provide an estimate of the log hazard ratio for the gender effect? [2 marks]

(iv) Which of the models 1 – 4 would ***not*** provide a test statistic (Score, Wald or LR) for the gender effect? [1 mark]

(v) Which of the models 1 – 4 could be used as a basis for *graphically* assessing the proportional hazards assumption for sex, while controlling for the other covariates? [1 mark]

(vi) Describe how this assessment would be done. [3 marks]

(b) Let the event of interest be contracting lung cancer. The lung cancer hazard rate for an x – old male smoker is given by:

$$\lambda(t) = 0.027 + 0.00025(t-40)^2, \quad t \geq 40$$

Assuming that a 40 – year – old male smoker is still alive at his age 50, determine the probability that he survives to age 50 without contracting lung cancer. [6 marks]

QUESTION FIVE (20 marks)

(a) (i) Describe the Cox regression model. [2 marks]

(ii) Why is this a semi parametric model? [2 marks]

(iii) Why is the Cox model often called a proportional hazards model? Explain it carefully. [2 marks]

(b) Consider the lengths of remission (in weeks) of one group of patients who are administered with 6-MP drug in the treatment of patients with melanoma cancer. [Survival times are in weeks]

Group	**Lengths of Remission (in weeks)**
6-MP Drug	6, 6, 6, 6+, 7, 9+, 10, 10+, 11+, 13, 16, 17+, 19+, 20+, 22, 23, 25+, 32+, 32+, 34+, 35+

The plus observations are censored times.

(i) Ignore the censoring indicators for treatment group (6-MP) for the moment (that is, treat all times as failure times). Compute the Kaplan-Meir estimator of the survival function for this treatment group. [5 marks]

(ii) Plot the Kaplan-Meir estimator of the survival functions for this treatment group. [2 marks]

(c) Given the survivorship function $S(t) = \exp(-t^{\gamma})$; derive the:

(i) Probability density function, [2 marks]

(ii) Hazard function, [2 marks]

(iii) Cumulative hazard function. [3 marks]