**JARAMOGI OGINGA ODINGA UNIVERSITY OF**

**SCIENCE AND TECHNOLOGY**

**SCHOOL OF MATHEMATICS AND STATISTICS**

**UNIVERSITY EXAMINATION FOR DEGREE OF BACHELOR OF SCIENCE IN ACTUARIAL SCIENCE WITH IT**

**2ND YEAR 2ND SEMESTER 2019/2020 ACADEMIC YEAR**

**MAIN (RESIT)**

---

**COURSE CODE: SAC 202**

**COURSE TITLE: LIFE TESTING ANALYSIS**

**EXAM VENUE:** **STREAM: (BSc)**

**DATE:** **EXAM SESSION:**

**TIME: 2.00 HOURS**

---

**<u>Instructions:</u>**

1. **Answer question 1 (Compulsory) and ANY other 2 questions**
2. **Candidates are advised not to write on the question paper.**
3. **Candidates must hand in their answer booklets to the invigilator while in the examination room.**

## QUESTION ONE

(a) Define the following words as used in Survival Analysis

i. Time - to- event.

ii. Distribution Function.

iii. Cumulative Hazard.

iv. Censoring.

v. Survival Function. [5 marks]

(b) Consider an exponential density function with parameter $\lambda = 2$. Determine the following;

i. hazard function. [2 marks]

ii. survival probability at time fifteen weeks. [2 marks]

iii. Mortality probability at any time. [2 marks]

(c) The table gives the data for a small sample of employees in a factory. It shows the time in months until the first absence from work. Observations marked +, show the time of leaving for those employees who left employment without being absent from work.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Male employees | $6^+$ | 11 | $13^+$ | 15 | $16^+$ | $19^+$ | 20 | |
| Female employees | $2^+$ | 4 | 7 | $8^+$ | $10^+$ | $12^+$ | 17 | $21^+$ |

calculate the Kaplan Meier estimate of the distribution function. [9 marks]

(d) Most actuaries use the exponential distribution of the form

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0; \\ 0, & \text{otherwise.} \end{cases}$$

as a premise for follow up studies. Derive the survival function and the cumulative hazard function, and comment on the structure of the hazard function. [6 marks]

## QUESTION TWO

(a) The remission data times ( in weeks) for two groups of Leukemia patients was recorded as follows

| Treatment | Placebo |
|---|---|
| 6, 6, 6, 7, 10, 13, 16, 22, 23, 25+, 32+, 32+ | 1, 1, 2, 2, 3, 4, 4, 5, 5, 15, 17 |
| 6+, 9+, 10+, 11+, 17+, 19+, 20+, 34+, 25+ | 8, 8, 8, 8, 11, 11, 12, 12, 22, 23 |

i. Determine the Kaplan Meier estimate of the distribution function of $F(t)$. [6 marks]

ii. Hence estimate the variance at time 23 of the survival function. [5 marks]

iii. Carry out a Mantel-Haenszel test on the remission data at $\alpha = 0.05$. [5 marks]

(b) For what reasons might lives be lost to the investigation if a study is carried out

i. a national investigation into the death from natural causes.

ii. a study of the mortality of life insurance policyholders. [4 marks]

## QUESTION THREE

(a) The population of elderly people in a prison is observed during the period 1 January 1994 to 31 December 1996. The duration of residence (measured to the nearest number of months) is recorded for those who die during the period, for those who are released from the prison during the period and for those who are still in residence on 31 December 1996. The recorded data measured in months are

$6^+$ 6 6 6 7 $9^+$ $10^+$ 10 $11^+$ 13 16 $17^+$ 20 $23^+$

where + indicates those who were released from the prison during the period or who were still in residence on 31 December 1996.

i. State the type(s) of censoring inherent in these data. [2 marks]

ii. Determine the product-limit (Kaplan-Meier) estimate of the survival function, $S(t)$, where $t$ is the duration of residence in the prison. [7 marks]

iii. State the assumptions underlying the estimate in $(ii)$, and explain how each of these assumptions would apply to these data. [4 marks]

(b) An investigation is carried out into the lifestyle of male accountants. A group of 10000 accountants is selected at random on 1 January 2001. Each member of the group supplies detailed personal information as at 1 January 2001 including name, address date of birth and marital status. The same information is collected as at each 1 January in the years 2002, 2003, 2004 and 2005. The investigation closes in 2005. A PhD student wishes to use the data from this investigation for her thesis on the mortality of married men. Describe the ways in which the available data for this investigation are censored. [4 marks]

(c) A trial for the life of a new battery was carried out by placing 1,000 batteries inside 1,000 fluffy electrical toys. The toys were turned on and left to run for 24 hours. A researcher returned every hour to count how many toys had stopped operating. On close analysis it was discovered that some of the toys had ceased to operate due to mechanical problems and that some of the toys were not switched on in the first place. Some of the toys were still running after 24 hours. For each of the following types of censoring, state with reasons whether it is present in the investigation:

i. left-censoring

ii. interval censoring

iii. type I censoring

iv. non-informative censoring [4 marks]

## QUESTION FOUR

(a) A Cox - PH model was estimated to assess the effects on survival of a person's sex and his or her self-esteem (measured on a three point scale as "low" or "medium" or "high"). The baseline category was males with "low" self-esteem. Write down the equation of the model, using algebraic symbols to represent variables and parameters defining all the symbols that you use. [5 marks]

(b) If $\hat{\mu}_{60} = 0.01$, $\hat{\mu}_{61} = 0.02$, $\hat{\mu}_{62} = 0.03$, estimate the values of $p_{60}$, ${}_2p_{60}$ and ${}_3p_{60}$. [3 marks]

(c) In a clinical trial, 50 patients are observed for two years following treatment with a new drug. The following data show the period in complete months from the initial treatment to the end of observation for those patients who died or withdrew from the trial before the end of the two year period.

Deaths 6, 6, 12, 15, 20, 20, 23

Withdrawals 1, 3, 5, 8, 10, 18

i. Calculate the Nelson-Aalen estimate of the integrated hazard function, $\Lambda_t$. [6 marks]

ii. Hence, or otherwise, estimate the probability of a patient surviving for at least 18 months after the initial treatment. [1 mark]

(d) Show that the survival function of a Cox PH model is given by

$$S(t, \underline{x}) = [S_0(t)]^{\exp(\sum_{i=1}^{p} \beta_i x_i)}.$$

[5 marks]

## QUESTION FIVE

(a) 12 brain tumor patients were randomized into radiation or radiation+chemotherapy. One year after the start of the study, survival time in weeks ere recorded as follows;

| | | | | | | |
|---|---|---|---|---|---|---|
| Group 0 RT | 10 | 26 | 28 | 30 | 41 | $12^+$ |
| Group 1 RT+CT | 24 | 30 | 42 | $40^+$ | $30^+$ | $42^+$ |

Apply a logrank test at $\alpha = 0.05$. [8 marks]

(b) Show that if the hazard function has the form $\alpha\beta(\alpha\, t)^{\beta-1}\exp[(\alpha\, t)^{\beta}]$, then the survival function is $\exp\{-\exp(\alpha\, t)^{\beta} - 1\}$. [6 marks]

(c) Give reasons as to why the Cox Proportional Hazards model is commonly used. [3 marks]

(d) Consider a discrete random variable $T$.

i. Show that $F(t) = \prod_{t_j<t}(1-\lambda_j)$. [3 marks]

ii. Derive the Kaplan Meier estimator of the survival function. [3 marks]