



# MERU UNIVERSITY OF SCIENCE AND TECHNOLOGY

P.O. Box 972-60200 – Meru-Kenya.

Tel: +254(0) 799 529 958, +254(0) 799 529 959, +254 (0)712 524 293

Website: [www.must.ac.ke](http://www.must.ac.ke) Email: [info@mucst.ac.ke](mailto:info@mucst.ac.ke)

---

## UNIVERSITY EXAMINATIONS 2023/2024

FOURTH YEAR SECOND SEMESTER EXAMINATION FOR DEGREE OF BACHELOR  
OF SCIENCE IN ACTUARIAL SCIENCE

### SMS 3466: SURVIVAL ANALYSIS

**DATE: APRIL 2023**

**TIME: 2 HOURS**

---

**INSTRUCTIONS: Answer Question ONE and any other TWO questions.**

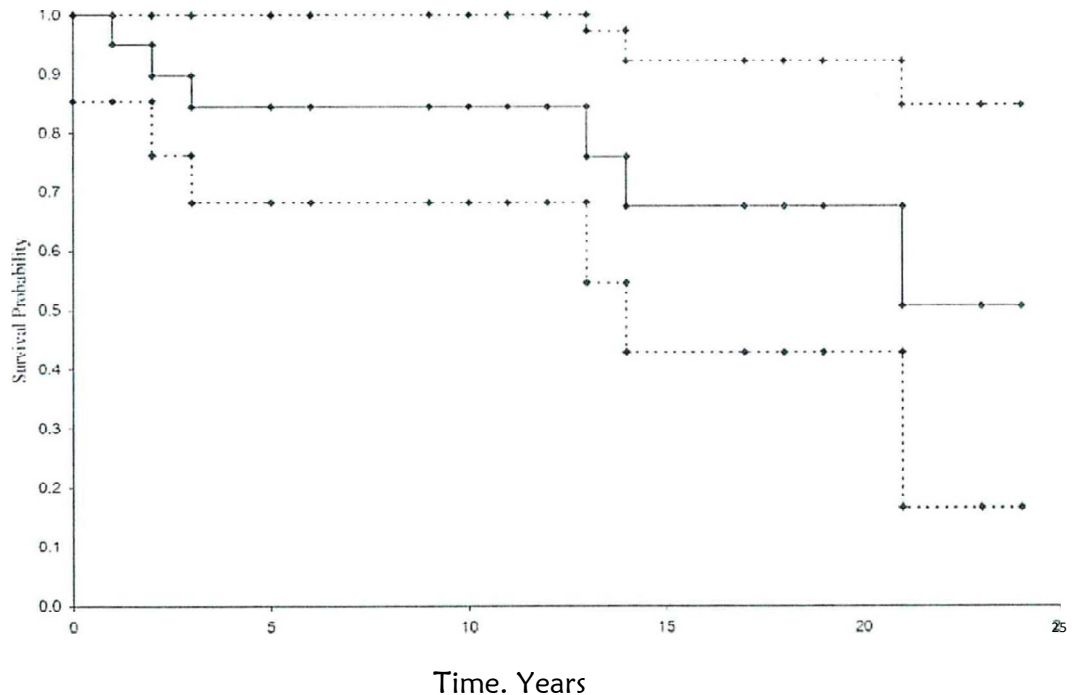
---

#### QUESTION ONE (30 MARKS)

- a) Define the term 'survival analysis' [1 Mark]
- b) With the help of an illustrative example, discuss the following terms s used in survival analysis:
- i Right censoring [2 Marks]
  - ii Interval censoring [2 Marks]
- c) The figure below summarizes the estimates and confidence intervals for a certain survival distribution. The Kaplan-Meier survival curve is shown as a solid line, and the 95% confidence limits are shown as dotted lines.



## Kaplan-Meier Survival Curve With Confidence Intervals



- i. Estimate the median and upper quartile of survival time [2 Marks]
  - ii. State two key features of survival data that are evident in the chart above. [3 Marks]
- d) State 3 assumptions for appropriate use of the Cox proportional hazards regression model [3 Marks]
- e) Write down the equation of the cox proportional hazards model in which the hazard function depends on duration  $t$  and a vector of covariates  $\mathcal{z}$ . Define all other terms you use. [2 marks]
- f) The hazard rate for a certain population is given by the Gompertz failure law as  $h(t) = e^{\beta t}, t > 0$   
Find expressions for the;
- i. Cumulative hazard function (I-I(t)) [2 marks]
  - ii. Survival function (s(t)) [2 Marks]
  - iii. Probability density function f(t) of T [1 Mark]
- g) Consider a discrete random variable T for the future lifetime in days of a group of insects, with the survival function given as;

T	0	1	2	3	4	5	6
S(t)	1	0.8	0.6	0.4	0.2	0.1	0

Find;

- i.  $F(t)$ , the cumulative distribution function of T [1 Mark]
  - ii.  $P(T=t)$ , the probability mass function of T [1 Mark]
  - iii.  $E(T)$ , the mean of T [2 Marks]
  - iv.  $Sd(T)$ , the standard deviation of T [2 Marks]
- h) Define the term "graduation" as used in survival analysis. [1 Mark]
- i) State two desirable properties of a graduation technique [2 Marks]

### QUESTION TWO [20 MARKS)

- a) Discuss 3 properties of the survival function  $S(t)$  [3 Marks]
- b) Discuss the properties of the Kaplan Meier estimator of the survival function [2 Marks]
- c) Consider the following data on survival times in weeks of 21 patients who tested positive for the COVID-19 virus.

6\*,6, 6, 6, 7,9\*10, 10\*, 11\*, 13, 16, 17\*, 19\*, 20\*, 22, 23, 25\*, 32\*, 32\*, 34\*,35\*

NB: 6\* means that the observation was censored at week 6

- i. Calculate the Product Limit Estimator for the survival function  $S(t)$  [5 Marks]
- ii. Find the variance of  $S(t)$  [4 Marks]
- iii. Plot the graph of  $S(t)$ ; hence construct a 95% confidence band for  $S(t)$  on the same chart. [6 Marks]

### QUESTION THREE [20 MARKS]

- a) A study is made on the impact of regular exercise and gender on the risk of developing heart diseases amongst 50 — 70-year-old patients. A sample of people was followed from the exact age of 50 years until they either develop

heart diseases or turn 70 years, whichever comes first. The cox PH model was used for this study with the two covariates being defined as;

$$z_1(\text{Gender}) = \begin{cases} 0, & \text{if female} \\ 1, & \text{if male} \end{cases}$$

$$Z_2(\text{exercise rate}) = \begin{cases} 0, & \text{if rare} \\ 1, & \text{if regular} \end{cases}$$

The model results were as follows:

S/N	Model fitted	Maximum 10 likelihood
1	Null model	-1269
2	Gender onl model	-1256
3	Both Gender and Exercise	-1250
4	Gender, Exercise and an interaction term	-1246

Note: the interaction term refers to Gender\* Exercise

Covariate	Parameter estimate
Gender	$B_1 = 0.20$
Exercise	$B_2 - 0.30$
Gender* Exercise	$B_0 - 0.35$

- (i) Give two reasons why the cox PH model IS suitable for use in survival data analysis (2 marks)
- (ii) Identify the baseline hazard for this model (2 marks)
- (iii) Give the hazard functions for a male who does not exercise regularly; a female who exercises regularly; a female who rarely exercises and a male who exercises regularly. (8 marks)
- (iv) Interpret your results in (iii) with reference to the baseline hazard and the hazard function for males who rarely exercise. (2 marks)

- (v) (ii) Perform a statistical test to show that the interaction term is significant in the model. Take  $\alpha = 5\%$  (6 marks)

#### QUESTION FOUR [20 MARKS]

- a) Differentiate between truncation and censoring (2 Marks)
- b) Given that the random variable for future lifetime  $T$  follows a Weibull distribution, with the hazard function given as  $h(t) = u^\beta (ut)^{\beta-1}$  where  $u > 0, \beta > 0$ , and  $t > 0$
- Find and graph the hazard function when the index parameter = 1. What does this hazard function mean? (4 Marks)
  - Find an expression for the cumulative hazard function  $H(t)$ , the survival function  $S(t)$ , and the probability density function  $f(t)$  for the hazard function in part (i) (6 Marks)
  - Show that the random variable  $T^\beta$  has an exponential distribution with parameter  $u^\beta$  (8 Marks)

#### QUESTION FIVE [20 MARKS]

- a) You are given the following data on survival times:
- 16 4, 5, 6, 7, 8, 5, 10, 10, 11, 13, 14, 6, 15, 15", 17, 17, 18, 18", 19", 21, 22
- Find the Nelson - Aalen estimate of the integrated hazard  $H(t)$  (6 Marks)
  - Using the results in part (i) above, find the Nelson Aalen estimate of the survival function  $S(t)$  (4 Marks)
  - Find the variance of the N-A estimate for the cumulative hazard at time  $t=16$  (4 Marks)
  - Plot a graph of  $H(t)$  showing the 95% confidence band as dotted lines (6 Marks)