THE CATHOLIC UNIVERSITY OF EASTERN AFRICA

A. M. E. C. E. A

P.O. Box 62157
00200 Nairobi - KENYA
Telephone: 891601-6
Fax: 254-20-891084
E-mail:academics@cuea.edu

MAIN EXAMINATION

SEPTEMBER- DECEMBER 2020 TRIMESTER

FACULTY OF SCIENCE

DEPARTMENT OF MATHEMATICS AND ACTUARIAL SCIENCE

REGULAR PROGRAMME

ACS 401: SURVIVAL MODELS

Date: DECEMBER 2020	**Duration: 2 Hours**
INSTRUCTIONS: Answer Question ONE and any other TWO Questions	

Q1. a) Define the following terms as used in survival analysis;

(i) Survival function **(1 mark)**

(ii) Hazard function **(1 mark)**

b) Describe the following types of censoring and give a suitable example for each.

(i) Interval censoring **(2 marks)**

(ii) Right censoring **(2 marks)**

(iii) Left censoring **(2 marks)**

c) The hazard rate for a certain population is given by the Gompertz failure law as

$h(t) = e^{ßt}, t > 0)$

Find expressions for the;

(i) Cumulative hazard function (H(t) **(2 marks)**

(ii) Survival function (s(t) **(2 marks)**

(iii) Probability density function f(t) of T **(2 marks)**

d) Write down the equation of the cox proportional hazards model in which the hazard function depends on duration t and a vector of covariates z. Define all other terms you use. **(2 marks)**

(ii) Why is the cox P.H model sometimes defined as semi-parametric? **(1 mark)**

e) 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 **(1 mark)**

(iv) Sd (T), the standard deviation of T **(1 mark)**

f) The mortality rates of a given population between 30 and 34 years are estimated by fitting a straight line ɑ + *Bx* to the crude values of Logo (qx/px). The observed deaths were as Follows:

Age (x)	30	31	32	33	34
No. of deaths	335	391	428	436	458

Test whether this model is a good fit for the data given x= - 10.94 and ß = 0.1104.

Assume that the initial exposed to risk were 700,000 in each age. **(7 marks)**

Q2. (i) Define the term "graduation" as used in survival analysis. **(1 mark)**

(ii) Give two reasons for graduating mortality rates. **(2 marks)**

(iii) State two desirable properties of a graduation technique **(2 marks)**

a) A music class provides students with musical instruments tuition. The teacher is concerned about the number of students giving up playing the instruments and wishes to test a new tuition method. Data was collected and a Cox PH model was fitted to model the hazard of giving up playing the instruments. The symmetric 95% confidence intervals based on standard errors for regression parameters are shown below;

Covariates	95% C.I
a) Instruments	
Violin	(-0.05, 0.19)
Piano	0
Trumpet	(0.07, 0.21)
b) Tuition Method	
Traditional method	0
New method	(-0.15,0.05)
c) Gender	
Male	(-0.08, 0.12)
Female	0

The parameter estimates are $ß_1 = 0.07$, $ß_2 = 0.14$, $B_3 = 0.06$ and $B_4 = 0.02$

(i) Write a general expression for this model, defining all the terms used **(5 marks)**

(ii) State the regression parameters fitted in the model **(1 mark)**

(iii) Describe the class of students to which the baseline hazard applies **(2 marks)**

(iv) Does the new method improve chances of students continuing to play their instruments? Explain your reasoning. **(3 marks)**

(v) Based on the model results, find the probability that a boy will still be playing the piano after 4 years if provided with the new tuition method, given the probability that a girl will still be playing the trumpet after 4 years following the traditional method is equal to 0.7. **(4 marks)**

Q3. a) A study is made on the impact of regular exercise and gender on the risk of developing heart diseases amongst 50 - 70-year olds. 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;

$$zi\ (Gender) = \begin{cases} 0, if\ female \\ 1, \quad if\ male \end{cases}$$

$$z2\ (Exercise\ Rate) = \begin{cases} 0, & if\ rare \\ 1, & if\ regular \end{cases}$$

The model results were as follows;

	Model fitted	Maximum log likelihood
1	null model	-1269
2.	Gender only	-1256
3	Gender and Exercise	-1250
4	Gender, Exercise and Interaction	-1246

NB/ the interaction = Gender * Exercise

Covariate	Parameter fitted
Gender	$B_1 = 0.20$
Exercise	$B_2 = -0.30$
Gender * Exercise	$B_o = -0.35$

(i) Give two reasons why the cox PH model is suitable for use in survival data analysis **(2 marks)**

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

(iii) Identify the baseline hazard for this model and state what it represents **(2 marks)**

(iv) 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)**

(v) Interpret your results in (iv) with reference to the baseline hazard and the hazard function for males who rarely exercise. **(2 marks)**

Q4. a) The following are remission times of a disease from a group of patients who received a certain drug and another group who received no treatment.

Group 1 (Treatment given)

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

Group II (no treatment)

1,1,2,2,3,4,4,5,5, 8, 8, 8, 8, 11, 11, 12, 12, 15, 17, 22, 23

Note: + means censored value

Assuming the exponential distribution of survival time, test the hypothesis that the drug administered is not effective. Take ɑ 10%. **(8 marks)**

b.) Given the following data on survival times find the Kaplan-Meir estimate of the integrated hazard and give its variance using the Green-Woods formula. Use t=16

4, 5, 6, 7*, 85, 10, 10, 11, 13, 146, 15, 15", 17, 17, 18, 18", 19", 21, 22

(12 marks)

Q5. a) State two properties of the survival data **(2 Marks)**

b) Given that the random variable for future lifetime I follows a Weibull distribution, with

the hazard function given as;

$h(t) = uß\,(ut)^{B-1}$ where $u > 0$, $ß > 0$,$t > 0$

(i) Find the hazard function when the index parameter ß = 1. What does this hazard function mean? **(2 marks)**

(ii) 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 (1) **(6 marks)**

(iii) Show that the random variable TB has an exponential distribution with parameter uß (6 Marks)

c) (i) Differentiate between truncation and censoring **(2 marks)**

(ii) What are the consequences of the two characteristics of survival data?

(2 marks)