MERU UNIVERSITY OF SCIENCE AND TECHNOLOGY

P.O. Box 972-60200 – Meru-Kenya.

Tel: 020-2069349, 061-2309217. 064-30320 Cell phone: +254 712524293, +254 789151411

Fax: 064-30321

Website: www.must.ac.ke Email: info@mucst.ac.ke

University Examinations 2013/2014

FIRST YEAR, FIRST SEMESTER EXAMINATION FOR THE DEGREE OF MASTER OF SCIENCE IN APPLIED STATISTICS

STA 3114: SURVIVAL AND CLINICAL DATA ANALYSIS

DATE: DECEMBER 2013 **TIME: 3HOURS**

INSTRUCTIONS: *Answer question **one** and any other **two** questions*

QUESTION ONE - (30 MARKS)

a) i) Define censoring in survival data. (2 Marks)

ii) Suggest 3 possible reasons for censoring. (3 Marks)

b) Suppose the time T is a random variable defined by the survival function

$S(t) = e^{-\lambda t}; t > 0, \lambda > 0$

i. Obtain the probability density function for T. (3 Marks)

ii. Find the cumulative distribution function for T. (2 Marks)

iii. Find the Harzard function of T. (2 Marks)

iv. Find the cumulative Harzard function of T. (2 Marks)

c) Consider the data below showing survival times for 21 individuals. The plus sign indicates censored time

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

Obtain the Kaplan Meier estimate of the survival function and plot it. (6 Marks)

d) Consider the output data below obtained from a Cox proportional Harzards model.

	Coefficients	Se(coefficients)
Treatment 1	-0.3189	0.0669
Treatment 2	0.0172	0.0667

i. Express the data in a Cox proportional Harzards model and interpret the parameters. (6 Marks)

ii. Obtain the 95% confidence interval for the Harzard Ratio for both Treatment 1 and Treatment 2. (4 Marks)

QUESTION TWO (20 MARKS)

a) Suppose T is a discrete lifetime that comes from a Geometric distribution.
$f(t) = \theta(1-\theta)^t; t = 0,1,2, \ldots.$

Obtain

i. The survival function of T. (3 Marks)

ii. The Hazard function of T. (2 Marks)

iii. The cumulative Hazard function. (2 Marks)

iv. For $\theta = 2$, plot both h(t) and H(t) versus time. Comment. (3 Marks)

b) An electronic firm experiments with two types of light bulbs that they manufacture. The following data shows time in months obtained from the lifespan of the light bulbs.

Group A	1	2	3^+	4	4^+	7^+	8	9	11	15^+
Group B	1^+	2	2	3	3	4	4	5^+	6	7^+

i. Test if the hazard rate in the two groups differs at $\propto = 0.05$. use the Log rank test on
$H_0: h_1(t) = h_2(t)$ versus
$H_1: h_1(t) \neq h_2(t)$ (5 Marks)

ii. Conduct a test of hypothesis on the hazard rates for the two groups based on the Wilcoxon test. Choose $\propto = 0.05$. Compare results here with those of the Log rank test above. (5 Marks)

QUESTION THREE (20 MARKS)

a) Let T be survival time defined by a survival function
$S(t) = \frac{t}{1+\lambda^{\propto}};\ \lambda_1 \propto = constants$

i. Obtain f(t), h(t) and H(t), i.e probability density, hazard and cumulative hazard functions respectively. (4 Marks)

ii. Show that the p^th^ percentile of the survival time is given by
$t_p = (1-p)(1+\lambda^{\propto})$. (2 Marks)

iii. Obtain the median survival time if $\lambda = \propto = 0.5$. (2 Marks)

b) The remission time in weeks for patients on drug treatment and on control treatment is recorded in weeks. The plus sign indicates censored survival time

Treatment	15	19^+	21	22^+	30	35
Control	7	8	10^+	11	17^+	19

i. Obtain the Kaplan Meier estimate of the survival function for the two groups and plot them on the same axis. Compare the two groups. (6 Marks)

ii. Conduct a Logrank test of hypothesis ($\propto = 0.05$) to compare the two groups. Is the drug effective? (6 Marks)

QUESTION FOUR (20 MARKS)

a) Using the Delta method, prove that;
"If a random variable Y$\sim N(\mu, \sigma^2)$, then a function of Y given by g(Y) is approximately normal mean $g(\mu)$ and variance $[g'(\mu)]^2\sigma^2$".

b) The Kaplan Meier estimator of a function is given by

$$\hat{S}(t) = \prod_{t_{j<t}} \frac{r_j - d_j}{r_j}; r_j = risk\ set, d_j = number\ of\ deaths\ at\ t_j$$

If $\hat{P}_j \sim N\left(p_j,\ p_j \frac{(1-p_j)}{r_j}\right), E\left(\hat{S}(t) = S(t)\right)$, show that

$$\text{var}\left(\hat{S}(t_j)\right) = [S(t)]^2 \sum_{t_j \le t} \frac{d_j}{r_j\left(r_{j-d_j}\right)}$$

QUESTION FIVE (20 MARKS)

Conduct a multivariate Log rank test of hypothesis to compare the survival in the three groups below. choose $\propto = 0.05$.

Group 1	20	21	23	24	24	26	26	27	28	30
Group 2	26	28	29	29	29	30	31	31	32	-
Group 3	31	32	34	35	36	38	38	39	-	-

(20 Marks)