



**JARAMOGI OGINGA ODINGA UNIVERSITY OF SCIENCE &
TECHNOLOGY UNIVERSITY EXAMINATIONS 2012/2013**

**2ND YEAR 1ST SEMESTER EXAMINATION FOR THE DEGREE OF
BACHELOR OF SCIENCE ACTUARIAL**

(MAIN)

COURSE CODE: SAS 205

COURSE TITLE: STATISTICAL COMPUTING I

DATE: 12 /8/13

TIME: 2 .00 - 4.00 PM

DURATION: 2 HOURS

INSTRUCTIONS

- 1. This paper contains five (5) questions.**
- 2. Answer question 1 (compulsory) and ANY other TWO questions.**

3. Write all answer in the booklet provided.

Question one 30marks

- a) Define a compiler in statistical computing. 2marks
- b) Explain in details any four areas in which statistical computing can be used. 8marks
- c) Describe what is meant by simulation modeling. 1mark
- d) Describe the advantages and disadvantages of using simulation modeling in complex analysis. 4marks
- e) Illustrate how variables are assigned using R software in statistical Computing. 2marks
- f) Explain Data archiving and the reason for archiving data. 4marks
- g) In R programming variable X can be assigned the a value a shown ,

`X<-1:5`

Find the function of the above expression: 4marks

- i) Sum
 - ii) Minimum
 - iii) Mean
 - iv) Standard deviation
- h) An experiment was performed with seven hop plants. One half of each plant was pollinated and the other half was not pollinated. The yield of the seed of each hop plant is tabulated as follows.

Pollinated	.78	.76	.43	.92	.86	.59	.68
UNpollinated	.21	.12	.32	.29	.30	.20	.14

Carry out the appropriate statistical hypothesis test to determine if the pollinated half of the plant gives a higher yield in seed than the un-pollinated half. You may assume normality. 5marks

Question two 20marks

a) Explain the following phases of a compiler

6 marks

- i. Lexical analysis
- ii. Syntax analysis
- iii. Semantic analysis

b) Explain the characteristics of the operating systems for mainframe computers and personal computers.

4 marks

a) Safaricom Company must buy rechargeable batteries for its servers. Four brands are considered: A, B, C and D. Six batteries of each brand are tested to determine the number of hours before recharging is necessary. The results are given below.

Brand A	100	116	107	114	128	122
Brand B	134	163	144	147	137	132
Brand C	135	102	113	113	104	121
Brand D	140	127	124	142	138	141

i) Produce boxplots of the data, and draw appropriate conclusions.

5 marks

ii) Produce a suitable graph to investigate whether the variability in the results differs from brand to brand, and report your conclusions.

5 marks

Question three 20mks

a) Using R software syntax generate a linear graph showing the relationship between x and y. 5 marks

X	1	2	3	4	5	6
----------	----------	----------	----------	----------	----------	----------

Y	2	4	6	5	4	7
----------	----------	----------	----------	----------	----------	----------

- b) Using a diagram explain the context of a compiler. 5 marks
- c) Given a sequence X_1, X_2, \dots, X_n of $U[0; 1]$ -distributed pseudo random numbers, we can use a scatter plot of $(X_i; X_{i+1})$ for $i = 1; \dots; n-1$ in order to try to assess whether the X_i are independent. Using R software Create such a plot using the built-in random-number generator of R: 5 marks
- d) As a statistician explain the guideline formulated for data entry into software. 5 marks

Question Four 20mks

- a) Explain the purpose and importance of Chi Square as a nonparametric statistic. 6 marks
- b) What are the **three** main purposes/function of an operating system? 3 marks
- c) Explain null and alternative hypothesis. 3 marks
- d) A genetics engineer was attempting to cross a tiger and a cheetah. She predicted a phenotypic outcome of the traits she was observing to be in the following ratio 4 stripes only: 3 spots only: 9 both stripes and spots. When the cross was performed and she counted the individuals she found 50 with stripes only, 41 with spots only and 85 with both. According to the Chi-square test, did she get the predicted outcome? 8 marks

Question Five 20mks

- a) A researcher selected a sample of 150 seniors from each of three area high schools and asked each senior, “Do you drive to school in a car owned by either you or your parents?” The data are shown in the table. At $\alpha = 0.05$, test the claim that the proportion of students who drives their own or their parents’ cars is the same at all three schools.

	School 1	School 2	School 3	Total
Yes	18	22	16	56
No	32	28	34	94
	50	50	50	150

Compute the test values by developing appropriate algorithm of the above question and comment on its result.

(20 marks)