****

MAASAI MARA UNIVERSITY

**REGULAR UNIVERSITY EXAMINATIONS**

**2016/2017 ACADEMIC YEAR**

***FOURD* YEAR, *FIRST* SEMESTER**

**SCHOOL OF SCIENCE**

**BACHELOR OF SCIENCE IN APPLIED STATISTICS WITH COMPUTING**

**COURSE CODE: STA 421**

**COURSE TITLE: STATISTICAL COMPUTING**

**DATE: MONDAY 27TH JAN 2017 TIME: 11:00 – 1:00**

**INSTRUCTIONS TO CANDIDATES**

1. Answer Question **ONE** and any other **TWO** questions
2. No writing on the Question paper
3. Use of mobile phone in the exam room is prohibited

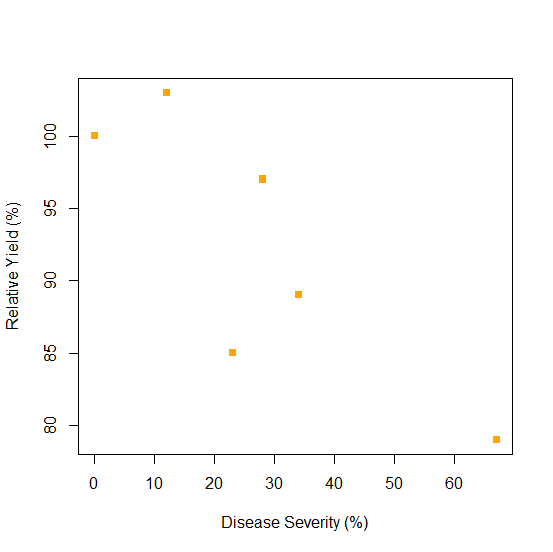
*This paper consists of three printed pages. Please turn over.*

**STA 421:STATISTICAL COMPUTING**

**QUESTION ONE *[COMPULSORY, 30 MARKS]***

1. Explain the use of the following notation in STATA output.
2. // **[2 marks]**
3. /\*\*/ **[2 marks]**
4. State and explain two modes of running STATA files. **[4 Marks]**
5. Write a detail R code to plot the graph below using the following data **[5 Marks]**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Disease Severity (%) | 0 | 12 | 67 | 34 | 23 | 28 |
| Relative Yield (%) | 100 | 103 | 79 | 89 | 85 | 97 |



1. Write a command of saving and importing data from excel to R **[4 marks]**
2. Write a command used in STATA to yield the following results and thereafter interpret the results **[5 marks]**





1. Cross tabulation between the status of a disease and a putative exposure have the following results.

Diseased Non-Diseased

Exposed 15 20

Non-Exposed 20 22

Create the table in R and perform chi-square and Fishers exact tests. **[5 marks]**

1. Explain the meaning of every statement in R programming language given by ## **[4 marks]**

zx=read.csv("D:\\rkoech.csv",header=T)####

zx=as.matrix(zx)####

zx

X11=zx[ ,-c(5,6)]####

X11

Sc=M11-(M12%\*%solve(M22)%\*%M21) #####

Sc

**QUESTION TWO *[20 MARKS]***

1. The data below shows visits to a website in two separate months. Create a Box

Plot for Month 1 as well as a side-by-side box plot for both months. **[7 marks]**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| Month I | 4 | 40 | 15 | 20 | 32 | 30 | 51 | 29 | 25 | 83 | 24 | 25 | 18 |
| Month II | 7 | 4 | 7 | 55 | 6 | 8 | 9 | 12 | 32 | 5 | 7 | 14 | 9 |

1. Write R code on how to modelled balanced incomplete block designs **[10 marks]**
2. Let ‘y’ be a series of integers running from 1 to 1000. Compute the sum of elements of ‘y’ which are multiples of 7. **[3 marks]**

**QUESTION THREE *[20 MARKS]***

1. Write a code in R to display area between -4 and 4 on the x-axis and label the y axis as and main title: The Standard Normal Density Function  **[8 marks]**
2. To the plot obtained in above, insert the grid line at x=0 and y=0 in grey color, define the area equal to or greater than 1.96 as well as shading it in grey. **[8 marks]**
3. Point and label the area shaded above as  with the text being above the arrow **[4 marks]**

**QUESTION FOUR *[20 MARKS]***

1. Use the data below to answer the following questions;

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 121 | 82 | 100 | 151 | 68 | 58 | 84 | 57 | 139 |
| 95 | 145 | 64 | 201 | 101 | 163 | 119 | 104 | 110 |
| 60 | 78 | 94 | 62 | 83 | 67 | 93 | 92 | 110 |
| 113 | 118 | 203 | 25 | 123 | 70 | 48 | 95 | 42 |

Write R scripts for plotting the following;

1. Histogram **[3 marks]**
2. Quantile plot **[2 marks]**
3. Normality plot **[2 marks]**
4. Stem and leaf plot **[3 marks]**
5. Use R codes to plot exponential graph below by considering all information during coding **[10 marks]**

