JARAMOGI OGINGA ODINGA UNIVERSITY OF SCIENCE AND TECHNOLOGY

SCHOOL OF MATHEMATICS AND ACTURIAL SCIENCE

UNIVERSITY EXAMINATION FOR DEGREE OF BACHELOR OF SCIENCE ACTURIAL

2nd YEAR 1st SEMESTER 2016/2017 ACADEMIC YEAR

MAIN REGULAR

COURSE CODE: SAS 205

COURSE TITLE: STATISTICAL COMPUTING I

EXAM VENUE: STREAM: (Bsc. **Actuarial Science with IT)**

DATE: **EXAM SESSION:**

TIME: 2.00 HOURS

Instructions:

1. **Answer questions one and any other two only.**
2. **Candidates are advised not to write on the question paper.**
3. **Candidates must hand in their answer booklets to the invigilator while in the examination room.**

QUESTION ONE (30 MARKS)

a)What is R? (2 marks)

b)Differentiate between the following giving an example in each case

i. Quantitative and qualitative variables (3 marks)
ii. Observations and variables (3 marks)

c)How would you check whether a vector contains missing values? (2 marks)

d)Match with the most appropriate choice available in (i) and (vi)

_____ counts the number of items for all levels of a factor (2 marks)

_____ adds text to the margin of a plot (2 marks)

_____ an argument of a function par() (2 marks)

_____ adds points to a plot (2 marks)

i. points
ii. boxplot
iii. mfrow
iv. table
v. mtext
vi. text

e)List 4 advantages of using R programming language over other programming languages (4 marks)

f)Differentiate the following giving an example in each case

i. Nominal and ordinal scales (4 marks)
ii. Numeric and categorical data types (4 marks)

QUESTION TWO (20 Marks)

In the following lines "species is a dataframe of species abundances in a series of plots. The columns are the species names and the rows are different plots. The dataframe "envir" contains the environmental data. The code works, but it is not commented. Please comment after the "#"

a)Species.pa = ifelse(species>0,1,0) #_____ (3 marks)

b)Diversity = rowSums(species.pa) #_____ (3 marks)

c)Diversity[1] #_____(2 marks)

d)plot(diversity~envir$temperature, ylab='', xlab='') #_____(3 marks)

e)mtext("temperature", side=1,line=3) #_____(3 marks)

f)If x is a factor with n levels and y is a length n vector, explain what happens if you compute y[x] (3 marks)

Two vectors x and y are defined as follows

x = c(3,2,4)

y = c(1,2)

What will be output of vector z that is defined as z = x*y (3 marks)

QUESTION THREE (20 Marks)

A real estate agent wanted to know to what extent the selling price of a home is related to its size. To acquire this information he took a sample of 12 homes that had recently sold, recording the price in thousands of dollars and the size in hundreds of square feet. The data is provided below.

size	23	18	26	20	22	14	33	28	23	20	27	18
price	315	229	355	261	234	216	308	306	289	204	265	195

i. Use a graphical technique to describe the relationship between size and price (4 marks)
ii. What is the shape of the relationship (4 marks)
iii. Interpret the graph in (i) (2 marks)
iv. Obtain the estimated regression parameters (4 marks)
v. Provide the estimated regression function for the data (2 marks)
vi. Interpret the estimated regression parameters (2 marks)
vii. Write R-codes that would provide the above outputs (2 marks)

QUESTION FOUR (20 Marks)

(a)The following data are the random sample of observations recorded from an industry producing juice per hour in different days

y_i :
50,55,62,67,45,68,70,62,73,64,75,55,50,68,64,60,66,60,56,59,60,60,63,67,66,68,70,65,54,55,70, 66,67

i. Provide an exploratory analysis with a boxplot using the above data (6 marks)
ii. Interpret the boxplot in (i) (2 marks)
iii. Give a confirmatory test for normality (2 marks)
iv. Under the assumption $y_i \sim N(\mu, \sigma^2)$, test the significance of H0: $\mu = 55$ against H_1: $\mu \neq 55$ at $\alpha = 0.05$ (6 marks)
v. Provide R-codes (i) and (ii) (4 marks)

QUESTION FIVE (20 Marks)

(a)A uniform distribution function is given by

$$f(x) = \begin{cases} \dfrac{1}{b-a}; & a < x < b \\ 0; & elsewhere \end{cases}$$

(i)Obtain E(x) and var(x) (8 marks)

(ii)Give an R-code that would be used to generate 1000 random numbers when a=0 and n=1. (2 marks)

(b)Suppose a binomial probability mass function is given by

$$f(x) = \begin{cases} \binom{n}{k} p^x (1-p)^{n-x}; & x = 0,1,2,\cdots,n \\ 0; & elsewhere \end{cases}$$

(i)Show that E(x)=np and var(x)=npq (8 marks)

(ii)Give an R-code that would be used to generate 100 random numbers when p=0.72 and n=400. (2 marks)