W1-2-60-1-6

JOMO KENYATTA UNIVERSITY

OF

AGRICULTURE AND TECHNOLOGY

UNIVERSITY EXAMINATIONS 2017/2018

SECOND YEAR SECOND SEMESTER EXAMINATIONS FOR THE DEGREE OF BACHELOR OF SCIENCE IN STATISTICS, BACHELOR OF SCIENCE IN FINANCIAL ENGINEERING, BACHELOR OF SCIENCE IN BIOSTATISTICS, BACHELOR OF SCIENCE IN ACTURIAL SCIENCE & BACHELOR OF SCIENCE IN OPERARTION RESEARCH

STA 2205: STATISTICAL PROGRAMMING

DATE: AUGUST 2018 **TIME: 2 HOURS**

INSTRUCTIONS TO CANDIDATES:

1. Answer questions ONE (section A) and any two questions in section B
2. Be neat and show all your workings
3. All questions except question one carry equal marks

This paper consists of 6 printed pages

STACS Examination board 2017/2018

QUESTION ONE (30 MARKS)

(a) The elasticity of demand E is the percentage rate of decrease of demand per percentage increase price. We obtain it from the equation according to the following formula:

$$E = -\frac{dq}{dp}\cdot\frac{p}{q}$$

Where the demand equation expresses demand, q as a function of unit p. To find the unit price that maximizes revenue, we express E as a function of p, set E=1, and then solve for p. A factory has the demand equation

$$q = -\frac{-5}{3000}p^2 + p + 1$$

i) Write a program in R that computes $\frac{dq}{dp}$ from the above equation. (3 marks)

ii) Given that

$$E = \frac{-\left(\frac{-10}{3000}p + 1\right)p}{\frac{-5}{3000}p^2 + p + 1}$$

Write a program in R that computes the unit price that maximizes revenue.(3 marks)

(b) Let $X \sim \chi^2(10)$ and $Y \sim t(5)$

(a-1) Write a function in R that does the following:

i) Generates 500 random variates of X and Y. (3 marks)

ii) Computes the mean of $\frac{Y}{X}$ as generated in (a-1)-(i) above. (2 marks)

iii) Returns the mean from (ii) above. (2 marks)

(a-2) write a program in R that calls the function created in (a-1) above 1000 times and stores the means in a vector called OurMeans. (3 marks)

(c) Consider the following system of linear equations

$$3x + y - 6z = -10$$
$$2x + y - 5z = -8$$
$$6x - 3y + 3z = 0$$

Write an R program to solve for the values of $x, y,$ and z. (2 marks)

(d) A hospital administrator wished to study the relation between patient satisfaction (Y) and patient's age (X_1, in years) and severity of illness (X_2, an index). The administrator randomly selected 4 patients. Use the R output below to answer the following questions.

```
Call:
lm(formula = Y ~ X1 + X2)

Residuals:
     1      2      3      4
-3.991  3.591  4.589 -4.190

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  166.545     46.091   3.613    0.172
X1            -4.610      3.802  -1.213    0.439
X2            -3.361      1.439  -2.336    0.257

Residual standard error: 8.212 on 1 degrees of freedom
Multiple R-squared:  0.8695,    Adjusted R-squared:  0.6085
F-statistic: 3.331 on 2 and 1 DF,  p-value: 0.3613
```

i) Identify the type of regression fitted on this data. Explain. (1 mark)

ii) Write down the fitted model. (1 mark)

iii) Comment on the signifance of the regression coefficients at 5% level of significance. Give an interpretation of these coefficients. (3 marks)

(e) Determine the output of the following program (2 marks)

```
b<-c(4,8,9)

while(length(b)-1<12

{

position<-length(b)-1;

w<-b[position]+b[position-1]

b<-c(b,w)

}
```

(f) Consider the following equation: $x^2 + 6x + 34 = 0$, write a function in R to evaluate the value of x. (1 mark)

(g) The following data represent the total number of abenant crypt foci observed in seven rats that had been administered a single dose of the carebioxgen azoxyymethane and sacrificed after six weeks (thanks to Ranjana Bird, Faculty of Human Ecology University of Manitoba for the use of these data):

87, 53, 72, 90, 78, 85, 83

Write a function in R that does the following:

i) Draws, with replacement, a sample of three (3) from the above data set. (2 marks)

ii) Generates seven (7) random observations from the poisson distribution with a mean equal to that of the sample drawn in (i) above. (2 marks)

QUESTION TWO (20 MARKS)

a) The dry weight of a crop Y is thought to be related to the organic matter X1, and amount of moisture X2 in the soil by the linear equation $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + e_i$ where $e_i \sim N(0, \sigma^2)$

X_1	5	7	8	10
X_2	18	19	9	9
Y	79	74	104	86

i) Write a program in R that fits a linear regression model on the above data and determines the least squares estimators of β_0, β_1 and β_2. (2 marks)

ii) Write a function in R that determines and returns least squares estimators of β_0, β_1 and β_2 using matrices. (3 marks)

iii) Write a program in R that determines and returns SSE, SSY. Where (5 marks)

$$R^2 = \frac{SSY - SSE}{SSY}$$

$$SSE = \sum_{i=1}^{N} (Y_i - \hat{Y})^2$$

$$SSY = \sum_{i=1}^{N} (Y_i - \bar{Y})^2$$

b) A company manufactures two products, X and Y by using three machines A, B and C. Machine A has 4 hours of capacity available during the coming week. Similarly, the available capacity of machines B and C during the coming week. Similarly, the available capacity of machines B and C during the coming week is 24 hours and 35 hours respectively. One unit of product X requires one hour of machine A, 3 hours of machine B and 10 hours of machine. Similarly one unit of product Y requires 1 hour, 8 hours and 7 hours of machine A, B and C respectively. When one unit of X is sold in the market, it yields a profit of Kshs. 50 per product and that of Y is Kshs. 70 per unit.

i) Formulate the linear program. (4 marks)

ii) Write a well commented program in R that will be used to solve the above program. (6 marks)

QUESTION THREE (20 MARKS)

a) One of the applications of the eigen vectors is in diagonalization of a matrix. Diagonalization means transforming a non-diagonal matrix which is easier to deal with. Let A be a matrix with distinct eigen values and P be a matrix whose columns are the eigen vectors of the eigen values of A. Then, the product $P^{-1}AP$ is a diagonal matrix. Let

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 2 & -2 & 3 \end{pmatrix}$$

Write a program in R that does the following:

i) Computes the diagonalising matrix of A as per instruction above. (3 marks)

ii) Computes the eigen values of A. (3 marks)

iii) Verifies that the results in (i) above is a diagonal matrix with the results in (ii) as its diagonal elements. (4 marks)

b) The lifetime X, of a certain type of of a tire is normal in distribution with an average of three years and a standard deviation of 0.3 yeard. Write an R function that does the following:

i) Generates 1000 random values of X. (2 marks)

ii) Computes the 99th percentile of X. (3 marks)

iii) Repeats (i) and (ii) 1000 times. (3 marks)

iv) Returns the average from (iii) above. (2 marks)

QUESTION FOUR (20 MARKS)

a) Consider the following statistical model for a stock price S_t:

$$S_t = S_0 e^{(\mu - 0.5\sigma^2)t + \sigma W_t}$$

$$W_t = \sum_{i=1}^{t} I_i$$

Where $I_i \sim N(0,1)$

Assuming that $S_0 = 110, \mu = 0.12$ and $\sigma = 0.16$

i) Write an R function that returns S_t for $t = 1, \dots, 100$ (3 marks)

ii) Write an R program that calls the above function and extracts the stock price on the 100th day. (2 marks)

iii) Repeat (ii) 1000 times. (2 marks)

iv) Use (iii) to compute the probability that the stock price on the 100th day will be below120. (2 marks)

v) Use (iii) to compute the 95% confidence interval for the stock price on the 100th day. (1 mark)

b) The current stock price, X_t, of BAS Associates is estimated by the model;

$$\hat{X}_t = 3 + 0.4X_{t-1} - 0.35X_{t-2} + 0.4X_{t-3} + e_t$$

Where X_{t-j} is the stock price j days in the past and $e_t \sim N(0,1)$. Consider the following past stocks prices for BAS Associates

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X_t	21	20	18	19	17	19	20	19	20	15	19	18	14	14	15

a) Write a well commented function in R that accepts a vector of past prices and returns $\hat{X}_t$ using the above equation and the while loop. (3 marks)

b) Write a well commented program in R that:

   i) Uses the above function to compute $\hat{X}_t$, $t = 4, 5, \ldots, 15$ for the data in the above table. (4 marks)

   ii) Plots the time series $\hat{X}_t$ computed in (b(i)) above and superimposes on the same plot the corresponding X_t. (3 marks)