



# MACHAKOS UNIVERSITY

University Examinations 2019/2020 Academic Year

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS AND STATISTICS

FOURTH YEAR SPECIAL /SUPPLEMENTARY EXAMINATION FOR  
BACHELOR OF EDUCATION

SST 302: PROGRAMMING LANGUAGE FOR STATISTICS II

DATE: 21/1/2021

TIME: 8.30-10.30 AM

**INSTRUCTION:** Answer question one and any other two Questions

## QUESTION ONE (30 MARKS)

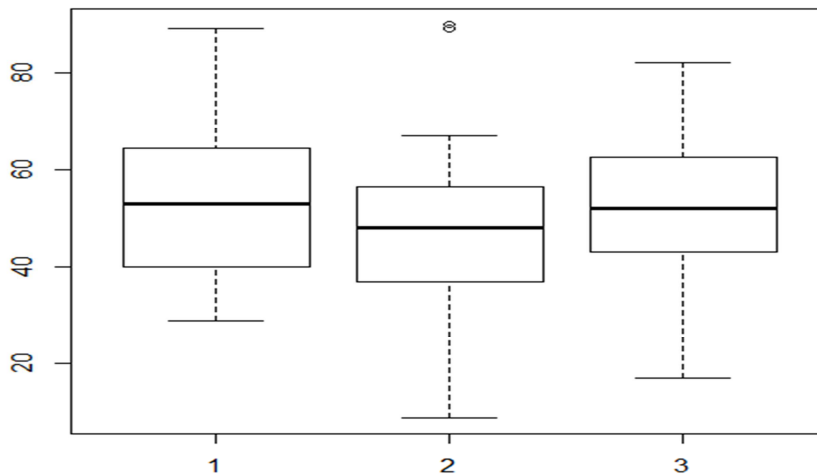
- a) By citing examples differentiate between the discrete random variables and continuous random variables (4 marks)
- b) Highlight three characteristics of the normal probability distribution and write R script for plotting normal curve ranging between -3 to +3 (6 marks)
- c) Suppose that men's neck sizes are approximately normally distributed with a mean of 16.2 inches and variance of 0.81 square inch. Using R script determine the probability that the neck size of a randomly chosen man lies between 13.5 and 18.9 inches (6 marks)
- d) Write R script for randomly choosing a sample of size five for numbers between 50 and 100 inclusive (3 marks)
- e) Suppose  $A = \{2, 3, 5, 7, 11, 13, 17\}$  and  $B = \{2, 4, 6, 8, 9, 14, 18\}$  write the R codes for
  - i.  $A \cup B$
  - ii.  $A \cap B$
  - iii.  $A \setminus B$
  - iv.  $A' \cup (A \cap B)$  (8 marks)
- f) Write the R code for choosing 25 random numbers from 300 to 375. (3 marks)

## QUESTION TWO (20 MARKS)

- a) Suppose in a certain shipment of 250 Pentium processors there are 17 defective processors. A quality control consultant randomly collects 5 processors for inspection to determine whether or not they are defective. By letting  $X$  denote the number of defectives in the sample, write the R code for determining  $P(x = 3)$  (4 marks)
- b) If  $X$  is the number of heads that show up after tossing a fair coin thrice, write the R codes for determine
- The sample space
  - the,  $E(x)$ ,  $Var(x)$  and standard deviation ( $x$ ) (8 marks)
- c) Three manufacturing machines A, B and C produces 70%, 20% and 10% of the output in a day. It has being observed that machines A, B and C produce 0.003, 0.004 and 0.005 of defective items respectively. Write the R codes for determining the probability of picking a defective item produced by machine B at random. (8 marks)

## QUESTION THREE (20 MARKS)

- a) Write R script for simulating data with 70% chance of choosing “male” for a committee of 20 members. (5 marks)
- b) The figure below gives summary of class 1, 2 and 3 end of term 1, 2019 performance for Tanya Academy.
- Write the R script for plotting the figure
  - Use the figure to estimate the; Quartiles, median, the minimum and maximum score for each class.
  - Which class had the highest dispersion? (8 marks)



c) Given the data set below

$Y = \{11, 13, 17, 34, 22, 19, 27, 23, 39, 44, 51, 43, 42, 37, 26\}$

Write it in vector form and give the R codes for determining its;

- i. Arithmetic mean
- ii. Sample Median
- iii. Generic summary for the data and model fits
- iv. Sample quintiles
- v. Sample variance and standard deviation

(7 marks)

#### QUESTION FOUR (20 MARKS)

a) A continuous random variable  $x$  has a probability function given by:

$$f(x) = \begin{cases} \frac{3}{32} (6x - x^2 - 5) & \text{for } 1 \leq x \leq 5. \\ 0 & \text{elsewhere} \end{cases}$$

Using the R commands

- i. Show that the function  $f(x)$  is a probability density function (p.d.f.)
  - ii. Determine the mean and the variance of the random variable  $x$  in the probability distribution. (12 marks)
- b) Suppose scores in an exam follow normal distribution with mean 80 and standard deviation 5. Using R script determine;

- i. The minimum score that you should get to be in the top 10%?
- ii. The maximum score that you should get to be in the last 25% (8 marks)

#### QUESTION FOUR (20 MARKS)

a) Write the R code for

- i. Computing;  $2.3^8 + \ln(7.5) - \cos(\pi / \sqrt{2})$
- ii. Generating the even numbers between 30 and 50. (6 marks)

b) Use the R codes

- i. To simulate 20 observations from the binomial distribution with  $n = 15$  and  $p = 0.2$ .
- ii. To estimate the 20th percentile of the gamma distribution with  $\alpha = 2$  and  $\theta = 10$ . (6 marks)

c) A random variable  $X$  with pdf given as 
$$P(x) = \begin{cases} \frac{1}{\Gamma(p/2)2^{p/2}} X^{p/2} e^{-x/2}, & X > 0 \\ 0 & \text{Otherwise} \end{cases}$$

is said to have a chi-square distribution with  $p$  degrees of freedom. Write the R code used to produce the Chi square distribution figure below (8 marks)

