



MACHAKOS UNIVERSITY

University Examinations 2018/2019

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS, STATISTICS AND ACTUARIAL SCIENCE

THIRD YEAR SPECIAL/ SUPPLEMENTARY EXAMINATION FOR
BACHELOR OF SCIENCE IN STATISTICS AND PROGRAMMING

SST 302: PROGRAMMING LANGUAGE FOR STATISTICS II

DATE: 24/9/2019

TIME: 11:00 – 1:00 PM

INSTRUCTION:

*Answer Question **ONE** which is compulsory and any other **TWO** Questions*

QUESTION ONE (COMPULSORY) (30 MARKS)

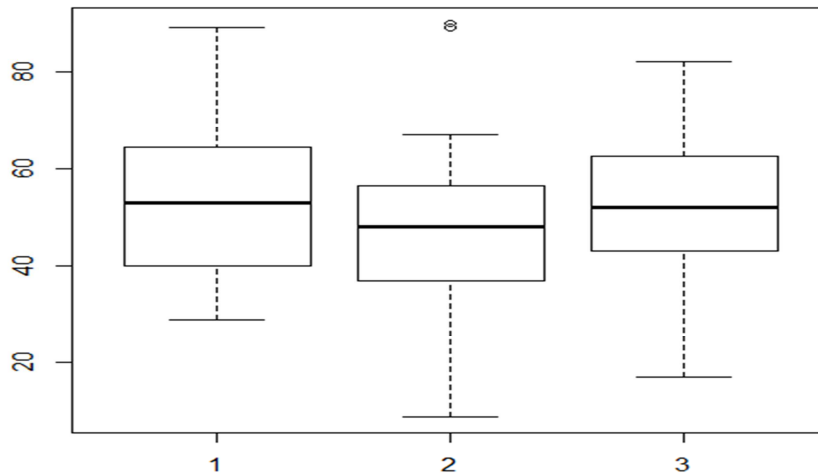
- a) Define a random variable and hence differentiate between the discrete random variables and continuous random variables (6 marks)
- b) Highlight three characteristics of the normal probability distribution and write R script for plotting normal curve ranging between -3 to +3 (6 marks)
- c) Suppose 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 (4 marks)
- e) If X is the number of heads that show up after tossing a fair coin thrice, write the R codes for determine
 - i) The sample space
 - ii) the, $E(x)$, $Var(x)$ and standard deviation (x) (8 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) 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)
- c) 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)

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.
- i. Write the R script for plotting the figure
 - ii. Use the figure to estimate the; Quartiles, median, the minimum and maximum score for each class.
 - iii. Which class had the highest dispersion? (8 marks)



- c) Given the data set below
 $Y \leftarrow c(11, 13, 17, 34, 22, 19, 27, 23, 39, 44, 51, 43, 42, 37, 26)$
 Write the R codes for determining its;
- Arithmetic mean
 - Sample Median
 - Generic summary for the data and model fits
 - Sample quintiles
 - 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

- Show that the function $f(x)$ is a probability density function (p.d.f.)
 - 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;
- The minimum score that you should get to be in the top 10%?
 - The maximum score that you should get to be in the last 25% (8 marks)

QUESTION FOUR (20 MARKS)

- a) Suppose a continuous random variable is defined by the pdf, $\int_0^5 (3x^2 + x - 27.3) \frac{dy}{dx}$, show that it is a pdf and then write the R codes for determining its $E(x)$, $\text{Var}(x)$ and $\text{sd}(x)$ (10 marks)
- b) Use the R codes
- To simulate 20 observations from the binomial distribution with $n = 15$ and $p = 0.2$.
 - To estimate the 20th percentile of the gamma distribution with $\alpha = 2$ and $\theta = 10$. (6 marks)
- c) Write the R code for computing $2.3^8 + \ln(7.5) - \cos(\pi / \sqrt{2})$ (4 marks)