MACHAKOS UNIVERSITY

UNIVERSITY EXAMINATIONS 2018/2019

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS, STATISTICS AND ACTUARIAL SCIENCE

THIRD YEAR SECOND SEMESTER EXAMINATION FOR

BACHELOR OF SCIENCE (COMMUNITY AND RESOURCE MANAGEMENT)

BACHELOR OF SCIENCE (FASHION DESIGN AND MARKETING)

HCU 301: INTRODUCTORY STATISTICS

DATE: 18/4/2019   **TIME: 8:30 – 10:30 AM**

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

QUESTION ONE (COMPULSORY) (30 MARKS)

a)   Explain the meaning of the following terms as applied in Statistics

   i)   Population

   ii)   Null hypothesis   (2 marks)

b)   Differentiate between EACH of the following terms:

   i)   Descriptive and inferential statistics   (2 marks)

   ii)   Type I error and Type II error   (2 marks)

c)   The data below relates to the number of successful sales made by a salesmen employed by a large marketing firm in a particular quarter

No of sales	0 to 4	5 to 9	10 to 14	15 to 19	20 to 24	25 to 29
No of salesmen	1	14	23	21	15	6

Calculate the mean and standard deviation   (6 marks)

d)   In a random sample of 64 patients in community-based dispensary, the mean waiting time for being served is 3 minutes is with a standard deviation of 1.5 minutes. Construct a 99% confidence interval for the average waiting time in the dispensary.   (5 marks)

e)   Given that $H_0: \mu = 100, H_a: \mu < 100, n = 36, \bar{x}=94$, s=30, test the null and alternate hypothesis at a significance of α=0.05.   (3 marks)

f)   Determine the values of a, b, c, d, e from the following ANOVA Table

	Sum of Squares	Degrees of Freedom	Mean Squares	F-Ration
Factor	354.67	b	c	e
Error	a	9	d	
Total	676.67	11		

(5 marks)

g)   A set of cards, identical in size and shape, is numbered 1 to 15 inclusive. A card is drawn at random from the pack and its number n is noted

    A is the event: n is prime

    B is the event: n is ≤ 10

    Calculate (i) P(A),(ii) P(B),(ii) P(A∧B)   (5 marks)

QUESTION TWO (20 MARKS)

A TV network is concerned about the high cost of producing many of its programs. A study is conducted to relate the production costs for 30 minutes of programming (in hundreds of thousands of dollars) to the ratings that the program gets in the national ratings survey. The results are shown below:

Production cost	1.2	1.6	1.8	2.5	2.7	3.0	3.5	4.4
Ratings	3.3	3.9	5.7	4.2	4.5	8.2	6.1	4.6

i)   Calculate the correlation coefficient between the ratings and the production cost.

(10 marks)

ii)   Interpret the results in (i) above   (2 marks)

iii)   Determine the least squares regression equation relating the ratings to the production costs   (6 marks)

iv)   Using the regression line obtained in (iii) above estimate the ratings when the costs is 6.5 dollars.   (2 marks)

QUESTION THREE (20 MARKS)

a) Explain the meaning of the following sampling techniques:-

      i)   Random sampling

      ii)   Stratified sampling

      iii)   Quota sampling

      iv) Cluster sampling   (8 marks)

b) The following are the heights in meters achieved by athletes competing in the pole vault event in a track and field competition:5.2,5.6,4.9,5.3,5.8,4.8,5.0,5.2,5.4,4.8,4.4,5.1,5.5,4.9,5.2,5.7,5.0,5.3,4.9,4.8

      i)   Construct a frequency distribution table with class interval by 4.4-4.6,…etc

                                                  (6 marks)

      ii)   Use the table in b(i) above to calculate the: i) mode   (3 marks)

                                            ii) Median.   (3 marks)

QUESTION FOUR (20 MARKS)

a) The following table shows the number of household members in certain town in 2010

No of House hold Members	Percentage
1	18
2	32
3	20
4	19
5	7
6or more	4

      i)   Calculate the mean and standard deviation of the number of households.

                                                  (5 marks)

      ii)   Assuming the data is based on a single random sample of 445; calculate a 95% confidence interval for the mean household size.   (5 marks)

b) Past records suggest that the heights of graduates of a certain college (at the time of their graduation) fit a normal distribution with mean 165 cm and standard deviation 6 cm. Use this information to determine:-

 i) The percentage of graduates whose heights is less than 170 cm.   (4 marks)

 ii) The percentage of graduates whose heights is between 170 cm and 175 cm.

                                                  (6 marks)

QUESTION FIVE (20 MARKS)

a)   Explain the meaning of each of the following terms as used in probability theory.

   i)   Random experiment
   ii)   An event
   iii)   Mutually exclusive events
   iv)   Independent events.   (8 marks)

b)   A consumer research organization conducts a survey of drivers to determine if there is any difference in their choice of brand of Japanese-made car based on their gender. These are results:

	Toyota	Mtsubishi	Subaru
Women	70	80	150
Men	40	60	100

   i)   Construct the corresponding cross-tabular contingency table for the expected frequencies.   (6 marks)
   ii)   Determine the value of the chi-square statistic   (6 marks)