



MACHAKOS UNIVERSITY

SCHOOL OF PURE AND APPLIED SCIENCES DEPARTMENT OF MATHEMATICS AND STATISTICS

University Supplementary Examinations

SIT 110- INTRODUCTION TO STATISTICS

Date:

Time:

Answer question ONE and any other TWO questions

QUESTION ONE (30 MARKS) COMPULSORY

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

- i. Population
- ii. Variable (2 marks)

(b) Differentiate between EACH of the following terms:

- i. Type I error and Type II error (2 marks)
- ii. Point and interval estimation (2 marks)

© .The data given below represents the frequency distribution of marks scored in mathematics by 250 students.

Grade scored	1	2	3	4	5	6	7	8
No. of students	8	20	36	64	48	40	24	10

Determine each of the following measures about the distribution

- i) Median
- ii) Mean
- iii) Standard deviation (7marks)

(d) A study of 49 randomly chosen 8-year-olds shows that they watch television an average of 38 hours per week with a standard deviation of 6.4 hours. Construct a 99% confidence interval for the average time per week that all such children watch television (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 $\alpha = 0.05$. (3marks)

(f) .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: $n \leq 10$

Calculate (i) $P(A)$, (ii) $P(B)$, (iii) $P(A \cap B)$ (4marks)

(g) 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	127.40	b	c	e
Error	a	16	d	
Total	1432.10	18		

(5 marks)

SECTION B

2. (a) Explain the meaning of the following sampling techniques:-

(i) Simple Random sampling

(ii) Stratified sampling

(iii) Judgmental sampling

(iv) Cluster sampling

(8 marks)

(b) A college collects the following set of data on the number of credits that a randomly selected group of students carry and the number of hours they work during the week

Hours worked per week	20	25	30	50	20	23
Number of credits	12	13	12	15	16	16

(i) Determine the linear least squares regression equation for number of credits as a function of the number of hours worked during the week (10marks)

- (ii) Use the equation in (a) above to predict the number of credits for a student who worked for 40 hours per week (2marks)

3(a). At a large departmental store, the average amount of all sales is \$78. On a special holiday promotion day, the total purchase amount of 80 different randomly selected sales has a mean of \$92 with a standard deviation of \$62. Test if the mean sale amount on a promotion day is higher than on a regular day at the $\alpha=0.1$ significance level. (8marks)

(b) The data below shows the number of hours worked in one week by employees in certain IT organization:

46.3 39.2 44.2 41.3 45.1 42.3 43.5 40.0
 45.6 40.6 42.0 42.6 45.6 39.5 43.1 39.7
 46.1 38.9 42.4 42.1 45.0 44.4 42.4 40.8

- (a) Tabulate a frequency distribution table with class intervals by 38.9 – 40.4, ... etc (6 marks)

- (b) Use the table in 2(a) above to calculate the quartile deviation (6 marks)

4 (a) Suppose we have the following data on Age versus the price of Desktop computers

Age	5	7	6	6	5	4	7	6	5	5	2
Price	80	57	58	55	70	88	43	60	69	63	118

Determine the correlation coefficient between age and price of desktop computer (8marks)

(b) 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
6 or more	4

- i. Calculate the mean and standard deviation of the number of households. (7 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)

5.(a) Perform an analysis of variance calculation for the data given below at $\alpha=0.05$ level of significance. (10marks)

Levels of factor

A	B	C
15	18	6
10	20	15
15	22	10

(b) An IT organization conducts a survey of their workers to determine if there is any difference in their choice of brands of Laptops based on their gender. These are results:

	Toshiba	Dell	Hp
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 (4 marks)