



MACHAKOS UNIVERSITY

University Examinations 2022/2023 Academic Year

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS AND STATISTICS

THIRD YEAR SECOND SEMESTER EXAMINATIONS FOR

BACHELOR OF SCIENCE (FASHION DESIGN AND MARKETING)

HCU 301-INTRODUCTORY STATISTICS

DATE: 14/4/2023

TIME: 11:00 – 1:00 PM

INSTRUCTION: Attempt question ONE and any other TWO questions

QUESTION ONE (30 MARKS)

- a) Explain the meaning of the following terms as applied in Statistics
- (i) Population
 - (ii) Null hypothesis (4 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)
 - iii. Point and interval estimation (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 the course of an audit it was found that from a simple random sample of 200 bad debts that the mean debt was £48.50 with a standard deviation of £6.50. Calculate a 95% interval for the mean debt (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 $\alpha = 0.05$. (4 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	205	b	c	e
Error	a	16	d	
Total	287	19		

(5 marks)

QUESTION TWO (20 MARKS)

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

- (i) Random sampling
- (ii) Stratified sampling
- (iii) Purposive sampling
- (iv) Cluster sampling

(8 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. Determine the least squares regression equation relating the ratings to the production costs (10 marks)
- ii. Using the regression line obtained in (iii) above estimate the ratings when the costs is 6.5 dollars. (2 marks)

QUESTION THREE (20 MARKS)

The data below shows the number of hours worked in one week by employees in certain company

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:
- (i) Mode (6 marks)
- (ii) Quartile deviation (8 marks)

QUESTION FOUR (20 MARKS)

- a) A small company is interested in analyzing the effects of advertising on its sales over a five week period as shown below:

Money spend on advertising	2	5	7	10	11
Total sales	10	20	35	50	65

Use the data to determine correlation coefficient between the total sales and the money spend on adverteng. (8 marks)

- 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. (6 marks)

- (ii) Assuming the data is based on a single random sample of 445; calculate a 95% confidence interval for the mean household size. (6 marks)

QUESTION FIVE (20 MARKS)

- a) A company has a computer system that can process 1200 bills per hour. A new system is installed that can process an average of 1260 bills per hour with a standard deviation of 215 bills in a sample of 40 hour. Test if the new system is significantly better than the old one at the 5% level of significance. (6 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	Nissan	Subaru
Women	70	80	150
Men	40	60	100

- (i) State the null hypothesis and alternative hypothesis. (2 marks)
- (ii) Construct the corresponding cross-tabular contingency table for the expected frequencies. (6 marks)
- (iii) Test whether there is any difference in the proportions using the different modes of transport based on gender at $\alpha = 0.01$ level of significance. (6 marks)