# MURANG'A UNIVERSITY OF TECHNOLOGY

## SCHOOL OF PURE, APPLIED AND HEALTH SCIENCES

DEPARTMENT OF MATHEMATICS AND ACTUARIAL SCIENCE

UNIVERSITY POSTGRADUATE EXAMINATION

2024/2025 ACADEMIC YEAR

**FIRST** YEAR **FIRST** SEMESTER EXAMINATION FOR MASTER OF SCIENCE IN STATISTICS

CCR 900 – ADVANCED STATISTICS AND RESEARCH

DURATION: 3 HOURS

**INSTRUCTIONS TO CANDIDATES:**

1. Answer any FOUR questions.
2. Mobile phones are not allowed in the examination room.
3. You are not allowed to write on this examination question paper.

**QUESTION ONE (25 MARKS)**

a. Discuss three methods of data collection. (6 marks)

b. Suppose your company has obtained the measurement of 80 of its employee's height and that they are recorded as follows

| | | | | | | |
|---|---|---|---|---|---|---|
| 160 | 166 | 170 | 173 | 179 | 181 | 184 |
| 166 | 162 | 171 | 179 | 174 | 181 | 185 |
| 163 | 167 | 179 | 174 | 171 | 181 | 186 |
| 168 | 163 | 174 | 172 | 179 | 182 | 186 |
| 164 | 165 | 174 | 173 | 179 | 182 | 187 |
| 164 | 169 | 175 | 173 | 187 | 182 | 180 |
| 164 | 169 | 173 | 175 | 180 | 183 | 188 |
| 165 | 169 | 173 | 175 | 180 | 183 | 188 |
| 165 | 176 | 173 | 170 | 181 | 185 | 183 |
| 165 | 170 | 173 | 176 | 181 | 183 | 189 |

i. Using class interval of 5 and 160 – 164 as the first class, construct a frequency distribution table. (5 marks)

ii. Calculate the:

I. Mean (4 marks)

II. Mode (3 marks)

III. Variance (5 marks)

IV. Standard deviation (2 marks)

**QUESTION TWO (25 MARKS)**

a. State two uses of regression analysis. (2 marks)

b. The following data shows the advertising expenses (in thousand dollars) and the net operating profit (in thousand dollars) in a random sample of six stores.

| | | | | | | |
|---|---|---|---|---|---|---|
| Advertising cost (x) | 1.5 | 1.0 | 2.8 | 0.4 | 1.3 | 2.0 |
| Net operating profit (y) | 3.6 | 2.8 | 5.4 | 1.9 | 2.9 | 4.3 |

i. Prepare a scatter plot for these data and observe if the assumption of linearity was satisfied. (4 marks)

ii. Determine the regression equation for predicting the net operating profit given the advertisement cost. (6 marks)

iii. Use the regression equation to estimate the net operating profit assuming the advertising cost is 3.5. (2 marks)

iv. Construct an ANOVA table. (7 marks)

v. Test whether the regression coefficient is significant at a 5% significance level. (4 marks)

**QUESTION THREE (25 MARKS)**

a. Using $\alpha = 0.05$, perform a run test for randomness on the sample data. (5 marks)

X0XXXX0000X000X000X0

b. The green yield (kg) under five treatments is as tabulated below.

| No. of plots | 1 | 2 | 3 |
|---|---|---|---|
| 1 | 3.17 | 3.44 | 3.15 |
| 2 | 3.40 | 2.88 | 2.69 |
| 3 | 3.5 | 2.97 | 3.10 |
| 4 | 2.87 | 3.24 | |
| 5 | 3.88 | | |

Using the Kruskal Wally's test, test the hypothesis that there is no difference among the three treatments at 10% significance level. (4 marks)

c. State two conditions necessary in the application of the Chi square test. (2 marks)

d. The Coca Cola company sells four brands of sodas in east Africa. To help determine if the same marketing approach used in Kenya can be used in Uganda and Tanzania, one of the marketing analysts want to ascertain if there is an association between the brands of soda preferred and the nationality of the consumer. She first classifies the population according to the brands of soda preferred i.e. Fanta, Sprite, Coke and Krest. Her second classification consists of the three nationalities Kenya, Tanzania and Uganda. The market analyst then interviews around sample of 250 soda drinkers from the criteria and records the observed frequency of drinkers into each of the cells as shown below.

| Nationality | soda performance | | | |
|---|---|---|---|---|
| | Coke | Krest | Sprite | Fanta |
| Kenya | 72 | 8 | 12 | 23 |
| Uganda | 26 | 10 | 16 | 33 |
| Tanzania | 7 | 10 | 14 | 19 |

Based on the sample data, can we conclude at 1% level of significance that there is a relationship between the preference of the soda drinkers and their nationalities. (8 marks)

**QUESTION FOUR (25 MARKS)**

a. What is a sample frame? (1 mark)

b. Describe three sampling procedures. (6 marks)

c. The police service commission in Kenya want to estimate the properties of crime in which firearms are used with a margin error of 0.01 and a 95% confidence interval. Data from the previous year's shows that p = 0.3. determine the sample size required for this estimation. (3 marks)

d. An auditor randomly samples 10 receivables accounts from the 500 accounts of a certain firm. The auditor lists the amount of each account and checks to see whether the underlying documents are in compliance with the stated procedures. The data is given below.
278, 192, 310, 94, 335, 310, 290, 221, 168. Calculate:

i. Mean (2 marks)
ii. Sample variance (2 marks)
iii. Mean of the variance (2 marks)
iv. The 95% confidence interval of the mean. (3 marks)
v. The 95% confidence internal of the total. (3 marks)

**QUESTION FIVE (25 MARKS)**

a. The quality control department of food processing specifies that the mean net weight percentage of a certain food must be 20g experience has shown that a standard deviation of 15g. if a random sample of 15 packages yield a mean weight of 19.5g is this significant evidence at 5% significance that the true weight of the packages has decreased. (7 marks)

b. Two different types of drug A and B were tried on certain patients for increasing weight. Four persons were given drug A and six persons were given drug B. the increase in weight is given as below.

| | | | | | |
|---|---|---|---|---|---|
| Drug A | 8 | 12 | 13 | 9 | |
| Drug B | 10 | 8 | 12 | 15 | 6 |

Do the two drugs differ significantly with regards to their effect on increasing weight at a 1% significance level? (11 marks)

c. A random sample of 15 items is taken and is found to have a mean weight of 16g and a standard deviation of 10g. what is the mean weight of the population with a 95% confidence interval. (3 marks)

d. Data were collected from two cities as regards the starting stipend paid to new management trainees.

| City | monthly stipend | sample | sample size |
|---|---|---|---|
| A | Ksh 1200 | ksh 90 | 100 |
| B | ksh 1500 | ksh 100 | 150 |

Does the data give evidence that the stipend paid in city A is significantly different from city A at 5% significance level? (4 marks)