



MURANG'A UNIVERSITY OF TECHNOLOGY

SCHOOL OF PURE AND APPLIED SCIENCE

DEPARTMENT OF APPLIED SCIENCE

UNIVERSITY ORDINARY EXAMINATION

2017/2018 ACADEMIC YEAR

**FOURTH YEAR SECOND SEMESTER EXAMINATION FOR THE DEGREE OF
BACHELOR OF SCIENCE IN MATHEMATICS AND COMPUTER SCIENCE**

SMA 2432 – DESIGN AND ANALYSIS OF SAMPLE SURVEYS

DURATION: 2 HOURS

DATE: 27TH APRIL, 2018

TIME: 9.00 – 11.00 A.M.

Instructions to Candidates:

1. Answer **Question 1** and **Any Other Two** questions.
2. Mobile phones are not allowed in the examination room.
3. You are not allowed to write on this examination question paper.

SECTION A – ANSWER ALL QUESTIONS IN THIS SECTION

QUESTION ONE

- a) Explain what is a good sample in measurement terms (3 marks)
- b) Define the following non-probability sampling methods:
- i. Purposive sampling (1 mark)
 - ii. Convenience sampling (1 mark)
 - iii. Quota sampling (1 mark)
- c) State and explain four causes of systematic bias in sampling (4 marks)
- d) In a private library, the books are kept on 130 shelves of similar size. The numbers of books on 15 shelves picked at random were found to be 28, 23, 25, 33, 31, 18, 22, 29, 30, 22, 26, 20, 21, 28 and 25. Estimate the total number Y of books in the library (4 marks)
- e) A sample of 30 students is to be drawn from a population of 300 students belonging to two colleges; A and B. The means and standard deviations of their marks are given below:

	Total number of students	\bar{y}_i	S_i
College A	200	30	10
College B	100	60	40

Use this information to confirm that Neyman's allocation scheme is a more efficient scheme when compared to proportional allocation (5 marks)

- f) Suppose the following summarized information is made available

$$n = 25 \quad N = 275 \quad \bar{x} = 9.2 \quad \bar{y} = 2.6$$
$$\sum x_i^2 = 2200 \quad \sum x_i y_i = 500 \quad \sum y_i^2 = 170$$

Estimate R and $\text{Var}(R)$ (5 marks)

- g) A production line makes 10,000 units a day. How can the quality control department take a systematic sample of 5% of these? (3 marks)
- h) Suppose that we have a population of size $N=4$ whose population units are 1, 2, 3, 4 and that we require a sample of size $n=2$ from the population. Assuming we use simple random sampling without replacement (SRSWOR). Find
- i. The number of possible samples (1 mark)
 - ii. Specify the samples (2 marks)

SECTION B – ANSWER ANY TWO QUESTIONS IN THIS SECTION

QUESTION TWO

- a) Let Y_i be the value of the characteristic under study for the i^{th} unit of the population and X_i be the value of the auxiliary characteristic of the i^{th} unit of the population. Show that the ratio estimate (\bar{y}_R) is a biased estimator of population mean \bar{Y} (6 marks)
- b) A daily newspaper conducts a survey of food costs by taking a simple random sample of 48 basic foodstuffs purchased in a large supermarket. Prices (in Kenya shillings) for these items are recorded in two separate occasions, three months apart, the earlier ones being denoted x_i and the latter y_i . The sample ratio r gives an indication of change of these basic food prices over three months period in the form of an estimate of the population ratio R of the mean prices of food.

The following results were obtained:

$$\bar{x} = 11.41, \quad \bar{y} = 12.07, \quad \sum x_i^2 = 8431.7$$
$$\sum x_i y_i = 8564.1 \quad \sum y_i^2 = 9270.6 \quad n = 48$$

- i. Obtain the value of r (2 marks)
- ii. Estimate variance of r (3 marks)
- iii. Obtain the 95% confidence interval for the population ratio R (3 marks)
- c) In studying lung function in a group of 560 workers in a coal mine, an estimate was required of mean value of some relevant measure Y . A simple random sample of 10 workers was chosen and their values, y_i determined by an appropriate test. A note was also made of their heights, x_i . The results were:

y_i	3.0	3.5	3.3	3.1	4.1	3.2	3.7	2.9	3.9	3.4
x_i (cm)	173	183	170	175	160	157	168	180	178	163

From routine medical records the average height for the group of 560 workers is known to be

$\bar{x} = 173.2$ cm. Estimate \bar{Y} from the data. (6 marks)

QUESTION THREE

- a) Research is to be carried out in the informal sector. The sector is divided into three groups, that is, low income earners, medium income earners and high income earners. The total amount allocated to this research is \$1500. The following data was collected after the interviews:

Class	Number of subsectors taken N_i	Variance of money spent S_i	Cost per interview
Low	350	200	5
Medium	500	400	8
High	250	500	12

- i. Design a proportional allocation sample (6 marks)
- ii. Design an optimal allocation sample (7 marks)
- iii. For the two designs in (i) and (ii), determine the estimated variances of the average amount spent hence or otherwise state the most efficient design and explain why (7 marks)

QUESTION FOUR

- a) Suppose we have a population of size $N=5$ whose population units are 1, 2, 3, 4, 5 and that we require a sample of size $n=2$ from this population. Assuming we use simple random sampling without replacement. Find;
 - i. Specify these samples (2 marks)
 - ii. Show that the sample mean is unbiased for the population mean using this data (4½ marks)
 - iii. Show that the sample variance is unbiased for the population variance using this data (5½ marks)
 - iv. Find the variance of the sample mean (2 marks)
- b) In a particular sector of the industry, a survey is conducted to investigate the extent of absenteeism from duty which is not connected to illness or official holidays. A random sample of 1000 employees out of the workforce of 36000 were asked to indicate the number of days they had failed to report to work for the previous six month. They were asked to give reasons. The following were the results

Number of days	0	1	2	3	4	5	6	7	8	9
Number of employees	451	162	187	112	49	21	5	11	2	0

- i. Determine the average number of days that were lost by the industry (3 marks)
- ii. Determine an estimate of population variance, S^2 (3 marks)

QUESTION FIVE

- a) A videotape hire company has shops in each of the 5 regions; three regions have 12 shops and others have 8. To estimate the total number Y of video films hired from the company in a

particular week, the sales manager phones 12 shops chosen by picking 3 regions at random then making a choice of 3, 5 and 4 shops from the chosen regions. The results were as follows:

First region	260	296	182			-12 shops
Second region	156	261	130	302	241	-8 shops
Third region	196	356	264	284		-12 shops

Estimate the total number y (10 marks)

- b) Let $N=1000$ and $n=10$. Assume that the random start r is selected, say, $r=5$. Find a systematic sample composition (3 marks)
- c) Assume that a population of size N is divisible into k groups each of size n so that $N= nk$. Let the units be labeled $1, 2, \dots, N$ in a fixed order of some kind.
- Describe briefly the process of drawing a systematic sample (4 marks)
 - Show that \bar{y}_{sy} is an unbiased estimator of \bar{Y} (3 marks)