Kasarani Campus
Off Thika Road
Tel. 2042692 / 3
P. O. Box 49274, 00100
NAIROBI
Westlands Campus
Pamstech House
Woodvale Grove
Tel. 4442212
Fax: 4444175

KIRIRI WOMENS' UNIVERSITY OF SCIENCE AND TECHNOLOGY

UNIVERSITY EXAMINATION, 2023/2024 ACADEMIC YEAR
THIRD YEAR, SECOND SEMESTER EXAMINATION
FOR THE DEGREE OF BACHELOR OF EDUCATION (ARTS)

Date: 13th December, 2023
Time: 8.30am –10.30am

KMA 2314 - SAMPLE SURVEYS

INSTRUCTIONS TO CANDIDATES

ANSWER **QUESTION ONE (COMPULSORY)** AND **ANY OTHER TWO** QUESTIONS

QUESTION ONE (30 MARKS)

a) The following are observations of a simple random sample without replacement from a population of size 10: 5, 4, 3, and 8.

i) Calculate the estimate of population mean [2 marks]

ii) Fine the estimate of the variance of the estimate obtained in (i) above [4 marks]

b) In SRSWOR, show that:

i) Sample mean is an unbiased estimator for population mean [3 marks]

ii) Sample variance is an unbiased estimator for population variance [4marks]

c) A sample survey for the study of yield and cultivation practices of coffee trees was conducted in certain division. Out of a total of 40 coffee growing villages, 8 villages were selected by simple random sampling. Data for the total number of coffee trees and area under coffee for the 8 selected villages are given below:

Serial No. of villages	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Total no. of coffee trees	492	108	714	1265	1889	784	294	798
Area under coffee trees (acres)	4.8	5.99	4.27	8.43	14.39	6.53	1.88	6.35

Given that the total number of trees of the 40 villages is 48,000, use ratio estimation to estimate the total area under coffee trees [5 marks]

d) Suppose in an experimental station, there 49 primary sampling units (clusters) of unequal sizes with a total of 400 secondary sampling units. Suppose five clusters are selected at random with replacement and used to estimate the population mean and variance of the estimate. The selected cluster values are as follows

Cluster id	**1**				**2**		**3**	**4**		**5**			
Observations	20	15	19	18	32	34	49	34	41	20	15	19	18
	15	23	19	26	27	30	41	39	33	15	23	19	26
	25	26	22	23	29	29	48	33	33	25	26	22	23
	21	23	22	27	34	37	43	33	34	21	23	22	27

Estimate the population mean using

i) The method of weighted mean of cluster means where the estimator is given as

$$\bar{y}_c^{*} = \frac{1}{n\,\bar{m}} \sum_{i=1}^{n} m_i\, \bar{y}_i$$ [4 marks]

ii) An unbiased ratio type estimation given that the estimator is

$$\bar{\bar{y}}_c^{**} = \bar{\bar{y}}_c + \left(\frac{N-1}{N\,\bar{m}}\right) s_{m\bar{y}} \quad \text{where } \bar{\bar{y}}_c = \frac{1}{n}\sum_{i=1}^{n} \bar{y}_i, \quad s_{m\bar{y}} = \frac{1}{n-1}\sum_{i=1}^{n} \left(m_i - \bar{m}\right)\left(\bar{y}_i - \bar{\bar{y}}_c\right)$$

[5 marks]

e) Describe any three circumstances under which census surveys are preferred to sample surveys [3 marks]

QUESTION TWO (20 MARKS)

a) In a population of N units, the number of units possessing a certain attribute is **A** and in a simple random sample of size **n** from it, the number of units possessing that attribute is **a**.

If $P = \frac{A}{N}$, $p_s = \frac{a}{n}$, $Q = 1 - P$ and $q = 1 - p_s$

i) Find the unbiased estimator of A and its variance [6 marks]

ii) Give an unbiased estimate of the variance obtained in (i) above [6 marks]

b) A random sample of 170 boys from 8502 boys in an area showed that 21 had some nutritional deficiency.

i) Estimate the proportion of nutritionally deficient boys in the population. Hence, find the estimated total number of nutritionally deficient boys in the population [3 marks]

ii) Estimate the standard error of the estimate of A [5 marks]

QUESTION THREE (20 MARKS)

a) In stratified random sampling, show that in optimal allocation for a fixed sample size, $n \propto N_i S_i$ [6 marks]

b) A population of size 140 is divided into three strata. The corresponding stratum sizes, means and variances are given below

Stratum No.	**Size** N_i	**Mean,** $\overline{Y_i}$	**Variance,** S_i^2
1	30	25	30
2	50	30	35
3	60	45	54

i) Determine the sample sizes required for the three strata in case of proportional allocation and Neyman's Optimum Allocation if a stratified sample of size 50 units was to be drawn [8 marks]

ii) By considering proportional allocation, obtain the estimate of the population mean, $\overline{y}_{st}$, and its variance [6 marks]

QUESTION FOUR (20 MARKS)

a) Derive the sample size required for estimating population mean in SRSWOR using α-level of significance and a margin error *d* when S^2 is unknown [6 marks]

b) A population having 400 sampling units is known to have a mean, $\overline{Y} = 25$ and variance $S^2 = 55.6$. In SRSWOR, how many sampling units should be chosen to estimate $\overline{Y}$ with a permissible error of 5% and at 98% confidence coefficient [4 marks]

c) Consider a population consisting of 12 units with the following values of a variable Y: 29, 87, 94, 10, 13, 43, 59, 63, 45, 98, 29, 31. A random sample of size 3 using a systematic procedure is required from this population.

i) Write down all the possible random samples [2 marks]

ii) Verify that the sample mean is unbiased for the population mean [3 marks]

iii) Use samples obtained in c) i) above to calculate $Var(\overline{y}_{sys})$ [5 marks]

QUESTION FIVE (20 MARKS)

a) Consider N identifiable population units $Y_1, Y_2 \ldots\ldots Y_N$. Associated with these units are auxiliary variable $X_1, X_2 \ldots\ldots X_N$. Let Y_i and X_i ($i = 1, 2, \ldots\ldots N$) be the values of the two characteristics for the i-th unit of a simple random sample without replacement from the population. Consider the ratio estimator $\bar{y}_R = r\bar{X}$ where $r = \frac{\bar{y}}{\bar{x}}$ as an estimator of $\bar{Y}$, the population mean. Show that $\bar{y}_R$ is biased for $\bar{Y}$

[5 marks]

b) The following data shows arable land (in acres) for 30 holdings in a certain district

Serial no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Total available land	60	50	14	10	1	20	40	70	0	0	100	0	20	30	15

Cont.	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
	25	30	95	0	13	30	0	10	0	85	30	35	36	10	20

Using systematic sampling with an interval of 5, estimate the average arable land in the district

[8 marks]

c) At an experimental station, there were 40 fields sown with wheat. Each field was divided into 16 plots of equal size. Out of the 40 fields, 5 were selected by SRSWOR. From each selected field, 4 plots were chosen by SRSWOR. The yields in kg/plot were as shown below;

Selected field	**1**	**2**	**3**	**4**
1	3.80	4.64	4.32	3.96
2	4.12	4.22	3.82	3.64
3	5.16	4.43	5.04	4.14
4	4.34	4.76	3.94	5.09
5	4.06	4.54	4.76	4.12

Estimate the mean yield per plot for the experimental station along with its standard error

[7 marks]