TECHNICAL UNIVERSITY OF MOMBASA

A Centre of Excellence

Faculty of Applied & Health Sciences

DEPARTMENT OF MATHEMATICS AND PHYSICS

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

MAY 2016 SERIES EXAMINATION

UNIT CODE: AMA 4432

UNIT TITLE: DESIGN AND ANALYSIS OF SAMPLE SURVEYS

TIME ALLOWED: 2HOURS

INSTRUCTIONTO CANDIDATES:

You should have the following for this examination

- Mathematical tables
- Scientific Calculator

This paper consists of **FIVE** questions

Answer question **ONE (COMPULSORY)** and any other **TWO** questions

Maximum marks for each part of a question are as shown

QUESTION ONE (30 MARKS)

(a) Give three reasons why a sample survey may be preferable to a census. (3 marks)

(b) Suppose a simple random sample of size n is selected from a population of size N. Give the first and second order inclusion probabilities under:

i) Selection without replacement (2 marks)

ii) Selection with replacement (2 marks)

(c) Give the conditions which make stratified random sampling preferable to simple random sampling, and briefly describe the procedure of stratified sampling.

(10 marks)

(d) Out of 600 monkeys in a population, a sample of 12 was selected without replacement and their weights recorded as follows:

8.6,7.7,8.9,9.2,10.7,10.1,12.1,6.5,8.5,9.8,10.5 and 8.8 kilograms.

Determine:

95% and 99% confidence interval for the population means.

(13 marks)

QUESTION TWO (20 MARKS)

a) Obtain the allocation for a stratum in stratified random sampling under:

i) Proportional allocation (1 mark)

ii) Optimum allocation where the cost is fixed to say C. (6 marks)

b) A population has three strata with sizes, mean and strata variance as given below:

	N_i	$\overline{X_i}$	S_i
Stratum 1	28000	4.2	1.1
Stratum 2	45000	3.6	1.7
Stratum 3	37000	3.8	1.5

i) Obtain the population mean and variance

ii) Obtain the proportion and Neyman allocation given that a stratified sample of size 1500 is to be drawn from the sample.

(13 marks)

QUESTION THREE (20 MARKS)

a) Consider simple random sampling without replacement. Suppose A_i is an indicator variable such that

$$A_i = \begin{cases} 1 & if\ unit\ i\ is\ included\ in\ the\ sample \\ 0 & otherwise \end{cases}$$

Show that

i. $E[\overline{y}] = \mu$ (3 marks)

ii. $Var[\overline{y}] = \left(1 - \frac{n}{N}\right)\frac{S^2}{n}$ (6 marks)

b) Consider a small population of size N= 4 having units 1, 2, 3, 4 with respective values 4, 5, 5 and 7. Suppose a simple random sample of size three is drawn without replacement.

i) Calculate the sample means. (2 marks)

ii) Calculate the population variance, $\sigma_{\overline{X}}^2$. (3marks)

iii). Verify that $E(\overline{x}) = \overline{X}$. (3 marks)

iv). Verify that $Var(\overline{x}) = \frac{N-n}{(N-1)n}\ \sigma^2$, where σ^2 is the population variance . (3 marks)

QUESTION FOUR (20 MARKS)

a) What is systematic sampling . (2 marks)

b) Suppose you have a population consisting of N=nk units. Explain how you draw a systematic sample of size n from this population.

(3 marks)

c) Show how the variance of a systematic sample mean is obtained in two different ways.

(15 marks)

QUESTION FIVE (20 MARKS)

a) Show that when using the stratified random sampling scheme, the estimator

$$\overline{y}_{str} = \frac{1}{N}\sum_{I=1}^{K} N_i \overline{y}_i$$

Is unbiased for the population mean. (6 marks)

b). Suppose we want to estimate the average number of hours of TV watched in the previous week for all adults in some county. Suppose also that the population of this county can be grouped naturally into 3 strata (Town A, Town B and Rural). Let the information be summarized as below;

statistic	Town A	Town B	Rural
h	1	2	3
N_h	155	62	93
n_h	20	8	12
$\overline{y}_{st}$	33.90	25.12	19.00
S_h	5.95	15.24	9.36
$N_h\,\overline{y}_h$	5254.5	1557.4	1767.0

Estimate the following;

i. Average TV time in hours per week for all household in the county (2 marks)
ii. Variance of the average TV time in the county. (3 marks)
iii. Total number of hours per week that the household view TV in the county. (1 mark)
iv. Variance of the total number of hours . (2 marks)
v. 95% Confidence interval for both the population mean and total. (4 marks)