(Knowledge for Development)

KIBABII UNIVERSITY

UNIVERSITY EXAMINATIONS

2020/2021 ACADEMIC YEAR

FIRST YEAR SECOND SEMESTER

MAIN EXAMINATION

FOR THE DEGREE OF BACHELOR OF EDUCATION AND BACHELOR OF SCIENCE

COURSE CODE: **STA 121**

COURSE TITLE: **SAMPLE SURVEYS I**

DATE: 13/07/21 **TIME: 2 PM -4 PM**

INSTRUCTIONS TO CANDIDATES

Answer Question One and Any other TWO Questions

TIME: 2 Hours

QUESTION 1:

(a)state three advantages of sampling over complete enumeration (3 marks)

(b) Let the sample arithmetic mean $\overline{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$ be an estimator of the population mean $\overline{Y}=\frac{1}{N}\sum_{i=1}^{N}Y_i$.

Verify that $\overline{y}$ is an unbiased estimator of $\overline{Y}$ under:

i) Simple random sampling without replacement (SRSWOR), (4 marks)

ii) Simple random sampling with replacement (SRSWR). (4 marks)

(c) Consider the estimation of $\overline{y}$ under SRSWOR and SRSWR. Which of these two sampling schemes is more efficient in carrying out the estimation? (4 marks)

(d) (i) Describe stratified sampling (6 marks)

(ii) Given the following data

Stratum, h	N_h	S_h
1	45	10
2	20	19
3	65	5

For a fixed sample size, n= 60, obtain n_h under the ,

(i) Optimum allocation scheme (3 marks)

(ii) Proportional allocation scheme (3 marks)

(iii) Neyman allocation scheme (3 marks)

QUESTION 2:

(a) Distinguish Cluster from Stratified sampling scheme (4 marks)

(b) Suppose it is desired that the coefficient of variation, CV of $\overline{y}$ should not exceed a given or pre-specified value of coefficient of variation, say C_0,then the required sample size n is to be determined such that,

$$CV(\overline{y}) \leq C_0 \quad \text{or} \quad \frac{\sqrt{\text{var}(\overline{y})}}{\overline{Y}} \leq C_0$$

Under these conditions, show that the smallest possible sample size $n_{smallest}$ is given by

$n_{smallest}=\frac{C^2}{C_0^2}$, where C is the population coefficient of variation (16 marks)

QUESTION 3:

(a) Given that p, a sample proportion is an unbiased estimator of a population proportion P, use the knowledge of $Var(\bar{y})$ to derive an expression for the Var(p). (7 marks)

(b) Assuming both N and n are large then $\frac{p-P}{\sqrt{Var(p)}}$ is approximately standard normal, N(0,1).Use this idea to write down the confidence interval of P at α level of significance. (4 marks)

(c) Illustrate how you would obtain sample size by fixing the confidence interval length (9 marks)

QUESTION 4:

(a) Describe Cluster Sampling procedure (4 marks)

(b) Distinguish Cluster from Systematic sampling scheme (3 marks)

(c) Suppose the number of words in a certain book is to be estimated. It is known that the book has 8 chapters and a total of 450 pages. A random sample of 4 chapters is selected using the simple random sampling procedure and the number of pages in the selected chapters is obtained. The data is given below.

Chapter	No. of pages(M_i)	Total no. of words	S_i
1	36	9650	252.96
2	52	12191	265.49
3	98	20845	311.74
4	66	16022	294.65

Obtain

(i) The mean number of words per page (6 marks)

(ii) The total number of words in the book (7 marks)

QUESTION 5:

(a) Consider a relatively large sample of size n. Let the sample be randomly divided into k groups each of size m units such that n=mk.

Let $\hat{S}^2$ be the estimator of population variance S^2 and be defined as

$$\hat{S}^2 = \frac{m}{k-1}\sum_{i=1}^{k}(\bar{y}_i - \bar{y})^2$$

Show that $E(\hat{S}^2) = S^2$ .Comment on the result. (10 marks)

(b) The variance, s^2 of a sample of size n may be given by

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$$

Verify that the sample variance, s^2 is an unbiased estimator of the population variance, S^2. (10 marks)