(*Knowledge for Development*)

KIBABII UNIVERSITY

UNIVERSITY EXAMINATIONS

2020/2021 ACADEMIC YEAR

SECOND YEAR FIRST SEMESTER

MAIN EXAMINATION

FOR THE DEGREE OF BACHELOR OF SCIENCE

COURSE CODE: **STA 212**

COURSE TITLE: **SAMPLE SURVEYS II**

DATE: 15/6/2021 **TIME:** 9:00 A.M - 11:00 A.M

INSTRUCTIONS TO CANDIDATES

Answer Question One and Any other TWO Questions

TIME: 2 Hours

QUESTION 1: (30 marks)

(a)By use of suitable examples, describe the following sampling techniques:

i) Genealogy sampling (2 marks)

ii) Snowball sampling (2 marks)

Iii) Purposive sampling (2 marks)

iv) Quota sampling (2 marks)

(b) Let the sample arithmetic mean $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ be an estimator of the population mean $\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$. Verify that $\bar{y}$ is an unbiased estimator of $\bar{Y}$ under:

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

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

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

(d) Suppose it is desired that the coefficient of variation, CV of $\bar{y}$ 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,

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

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 (6 marks)

(d) Describe Cluster sampling. How does it differ from Stratified sampling? (4 marks)

QUESTION 2: (20 marks)

(a) The sample size is to be determined such that the variance of $\bar{y}$ should not exceed a given value, say V. In this case we find n such that $Var(\bar{y}) \le V$ .Proceed to justify that for large N, the smallest value of the sample size, $n_{smallest}$ is

$n_{smallest} = n_e$ and $n \ge n_e$, where $n_e = \frac{S^2}{v}$ (10 marks)

(b) Suppose we want to determine sample size, n such that the following requirement is satisfied,

$$P\left[I\bar{y} - \bar{Y}I \le e\right] = 1 - \alpha$$

Where e is the absolute estimation error and α the level of significance.

Assume $\bar{y}$ follows $N\left(\bar{Y}, \frac{N-n}{Nn}S^2\right)$. Show that $n = \left(\frac{Z_{\frac{\alpha}{2}}}{e}S\right)$ for large N. (10 marks)

QUESTION 3: (20 marks)

(i) What do you understand by the term, Power of a statistical test? (4 marks)

(ii) An investigator is planning a clinical trial to evaluate the efficacy of a new drug designed to reduce systolic blood pressure. The plan is to enrol participants and to randomly assign them to receive either the new drug or a placebo. Systolic blood pressures will be measured in each participant after 12 weeks on the assigned treatment. Based on prior experience with similar trials, the investigator expects that 10% of all participants will be lost to follow up or will drop out of the study. If the new drug shows a 5 unit reduction in mean systolic blood pressure, this would represent a clinically meaningful reduction. At 5% level of significance, how many patients should be enrolled in the trial to ensure that the power of the test is 80% to detect this difference? (16 marks)

QUESTION 4: (20 marks)

(a) Describe stratified sampling (6 marks)

(b) 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 (4 marks)

(ii) Proportional allocation scheme (4 marks)

(iii) Neyman allocation scheme (4 marks)

QUESTION 5: (20 marks)

(a) State the advantages of systematic over simple random sampling scheme. (4 marks)

(b) Distinguish linear systematic sampling (LSS) from Circular systematic sampling (CSS). (6 marks)

(c) Prove that in LSS the population variance is the sum of the variations within the samples and between the samples (10 marks)