CHUKA **UNIVERSITY**

UNIVERSITY EXAMINATIONS

FOURTH YEAR EXAMINATION FOR THE AWARD OF DEGREE OF BACHELOR OF SCIENCE IN MATHEMATICS

MATH 441: SAMPLING METHODS II

STREAMS: BSC MATHS

TIME: 2 HOURS

DAY/DATE : WEDNESDAY 22 /09/ 2021 **2.30 PM – 4.30 PM**

INSTRUCTIONS TO CANDIDATES:

- Answer question **ONE** and **TWO** other questions
- All your working must be clearly shown

QUESTION ONE: (30 MARKS)

a) State three problems that can be caused by missing data. (3marks)

b) Consider the data below and differentiate between an estimate and estimator for population variance; 260, 296, 182 (4 marks)

c) Discuss two advantages of cluster sampling (4 marks)

d) Given the following sample: (5, 6, 0, 7, 3), obtain 12 Bootstrap resamples. From the resamples, obtain the:
i) Resample means
ii) Bootstrap 95% confidence interval. (6 marks)

e) Explain the types of non-response that may be encountered in sample surveys (6 marks)

f) Discuss the relationship between Jackknife and Boot strap resampling procedure. (6 marks)

g) Use the information given below and compute the 99 % confidence interval for the Ratio estimator (8 marks)

N=25, N=275, $\bar{x}$ =9.2, $\bar{y} = 2.6$, $\sum_{i=1}^{25} x_i^2 = 2200$, $\sum_{i=1}^{25} x_i y_i = 500$, $\sum_{i=1}^{25} y_i^2 = 170$

QUESTION TWO: (20 MARKS)

a) Discuss the bootstrap resampling procedure (6 marks)

b) Consider a sample {4,5,6,7,8}, compute the jackknife standard error and construct a 99% confidence interval for mean (14 marks)

QUESTION THREE: (20 MARKS)

a) Explain the condition under which the regression estimator is preferred over the ratio estimator and state the assumptions of the regression estimator. (6 marks)

b) In a survey for estimating the impact of a comprehensive development project for the fish FED (project) in Chuka, 8 villages were selected out of 47 in a blocking the district by srswor and a sample of households was selected by srswor from each selected village. The table below gives the total number of households with pisciculture as principal occupation (M_i) and water village area of the sampled household in each selected village.

Sample Village No.	M_i	Water Area
3	8	40,33,57
7	6	66, 92
13	9	66, 50, 33
17	11	33, 45, 60, 55
21	13	99, 66, 50, 67
31	13	60, 155, 77, 80
37	10	78, 33, 230
41	12	78, 66, 60, 77

Using two stage sampling, estimate the total water area of the block along with its standard error (14 marks)

QUESTION FOUR: (20 MARKS)

a) A population consists of MN elements grouped into N first stage units (fsu's) of M second stage units (ssu's) each. Let n be the number of fsu's in the sample and m be the number of ssu's to be selected from each of the fsu. Show that

i)The sample mean is an unbiased estimator of population mean (3 marks)

ii) $var(\bar{y}) = \frac{N-n}{N}\frac{S_b^2}{n} + \frac{M-m}{M}\frac{S_w^2}{mn}$ where

$S_b^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$ and

$S_w^2 = \frac{1}{N(M-1)} \sum_{i=1}^{N} \sum_{j=1}^{M} (Y_{ij} - Y_i)^2$ (4 marks)

b) The director of computer support department plans to sample 3 divisions of a large firm that has 10 divisions, with varying numbers of employees per division. Since number of computer support requests within each division should be highly correlated with the number of employees in that division, the director decides to use unequal probability sampling with replacement with p_i proportional to number of employees in that division.

Division	Number of Employees
1	1000
2	650
3	2100
4	860
5	2840
6	1910
7	390
8	3200
9	1500
10	1200

i) Implement unequal probability sampling according to the given $p_i s$

ii) With the divisions selected by probability proportional to size, how do we construct the Hansen-Hurwitz estimator for τ and its standard error. (13 marks)

QUESTION FIVE: (20 MARKS)

a) Explain the occurrence of bias in ratio estimators and how bias can be offset. (5 marks)

b) A mathematics achievement test was given to 486 students prior to entering a certain college who then took a calculus class. A simple random sampling of 10 students is selected and their calculus score recorded. It is known that the average achievement test score for the 486 students was 52.

Achievement test X	39	43	21	64	57	47	28	75	34	52
Calculus core Y	65	78	52	82	92	89	73	98	56	75

i) Obtain the $\hat{\bar{Y}}_{reg}$

ii) Calculate var($\hat{\bar{Y}}_{reg}$)

iii) What is the approximate 95% CI for μ using regression estimation (15 marks)

……………………………………………………………………………………………………