MURANG'A UNIVERSITY OF TECHNOLOGY

SCHOOL OF PURE AND APPLIED SCIENCE

DEPARTMENT OF APPLIED SCIENCE

UNIVERSITY ORDINARY EXAMINATION

2017/2018 ACADEMIC YEAR

RE-SIT/SPECIAL EXAMINATION FOR BACHELOR OF SCIENCE IN APPLIED STATISTICS WITH PROGRAMMING

AMS2302 – SAMPLE SURVEYS II

DURATION: 2 HOURS

DATE: 20TH APRIL 2018

TIME: 9.00AM – 11.00AM

Instructions to Candidates:

1. Answer **Section A** and **Any Other Two** questions in **Section B**.
2. Mobile phones are not allowed in the examination room.
3. You are not allowed to write on this examination question paper.

SECTION A – ANSWER ALL QUESTIONS IN THIS SECTION (30 Marks)

QUESTION ONE

a) In a national enquiry into the agricultural needs of the country, it is required that a study through a survey be conducted in the general agricultural practices. The research and extension staff impact is to be studied over one year period and this requires information on the number of farmers who go to researchers for consulting, those who need more extension services, those who go to private farming consultants etc. Two possible sampling schemes could be used in carrying out this study. Either a complete enumeration could be carried out on the small group of extension officers, research staff and private consultants or a sample of individual farmers across the country could be used. Discuss the advantages and disadvantages of the two sampling schemes in relation to the problem at hand. (4 Marks)

b) i) Briefly describe the stratified sampling scheme. (2 Marks)

ii) Show that the estimator $\overline{y}_{st} = \frac{1}{N}\sum_{h=1}^{N} N_h \overline{y}_h$ for the population mean is unbiased. (3 Marks)

c) Differentiate between a ratio estimator and a regression estimator and give two importance of each. (6 Marks)

d) In studying lung function in a group of 560 students in a Kenyan University, an estimate was required of the mean value of some relevant measure Y. A SRS of 10 students was chosen and their Y values,y_i, determined by appropriate test. A note was made of their heights, x_i. The results were;

y_i	3.0	3.5	3.3	3.1	4.1	3.2	3.7	2.9	3.9	3.4
x_i	173	183	170	175	160	157	168	180	178	163

From routine medical records the average height for the group of 560 students in known to be $\overline{X} = 173.2cm$. Estimate $\overline{Y}$ and Y from the data and calculate an appropriate standard error for your estimator. (7 Marks)

e) The values of Y and *x* are measured for each unit in a simple random sample from a population. If $\overline{X}$, the population mean of X, is known, which of the following procedures do you recommend for estimating $\overline{Y}/\overline{X}$?

i. Always use $\overline{y}/\overline{x}$

ii. Sometimes use $\overline{y}/\overline{x}$ and sometimes $\overline{Y}/\overline{X}$

iii. Always use $\overline{Y}$

Give reasons for your answer. (3 Marks)

f) A circulation manager of a certain newspaper wishes to estimate the average number of newspapers that are purchased in a household in a given community. From the study 4,000 households in the community are listed in 400 geographical clusters of 10 households. Each. Under simple random sampling, 4 clusters are selected and interviews conducted. The results are shown below

Cluster	No. of Newspapers

1	1,2,1,3,3,2,1,4,1,1
2	1,3,2,2,3,1,4,1,1,2
3	2,1,1,1,1,3,2,1,3,1
4	1,1,3,2,1,5,1,2,3,1

Estimate the average number of newspapers per household for the community. (5 Marks)

SECTION B – ANSWER ANY TWO QUESTIONS IN THIS SECTION

QUESTION TWO (20 MARKS)

a) If the n units and m subunits from each chosen unit are selected by simple random sampling, show that (10 Marks)

$$Var\left(\bar{\bar{y}}\right) = \left(\frac{N-n}{n}\right)\frac{S_1^2}{n} + \left(\frac{M-m}{M}\right)\frac{S_2^2}{mn}$$

b) A mathematics achievement test was given to 486 students prior to their entering a certain college. From those students a SRS of n=10 students was selected and their progress in calculus I observed. Final Calculus I grades were then reported as follows;

Student	**Achievement test score X**	**Final calculus I score Grade Y**
1	39	65
2	43	78
3	21	52
4	64	82
5	57	92
6	47	89
7	28	73
8	75	98
9	34	56
10	52	75

It is known that $\overline{X} = 52$ for all the 486 students taking the achievement test.

Estimate $\overline{Y}$ for this population and obtain its standard error. (10 Marks)

QUESTION THREE (20 MARKS)

a) An experienced farmer made an eye estimate of the weight of peaches X_i on each tree in an orchard of N=200 trees. He finds a total weight of X=11,600kg. The peaches are picked and weighed on a SRS of ten trees with the following results

No.	1	2	3	4	5	6	7	8	9	10
Actual weight	61	42	50	58	67	45	39	57	71	53
Estimated weight	59	47	52	60	67	48	44	58	76	58

As an estimate the total weight Y, we take $\hat{Y} = N[\overline{y} + (\overline{X} - \overline{x})]$. Compute the estimate and find the standard error. (11 Marks)

b) Suppose that $\overline{X}$ is known and $\overline{Y}_R$ denotes the ratio estimator. Show that $var(\overline{Y}_R)$ is smaller than the variance of the mean of a simple random sample of n y-values of $p > {}^1/_2 \, R \frac{S_x}{S_y}$ (9 Marks)

QUESTION FOUR (20 MARKS)

a) Define the Horvitz Thomson estimator of the mean in stratified random sampling and show that (10 Marks)

$$V(\hat{Y}_{HT}) = \sum_{h=1}^{H} \left(\frac{N_h - n_h}{N_h}\right) \frac{N_h^2 S_h^2}{n_h} = \sum_{h=1}^{H} \left(1 - \frac{n_h}{N_h}\right) \frac{N_h^2 S_h^2}{n_h}$$

b) In a survey to examine trends of real estate an investigator is interested in the relative change of two year period in the assessed value of Homes in a particular community. A simple random sample of n=20 homes is selected from N=1000 homes in the community.

From tax records the investigator obtained the assessed value for this year 1994 (y_i) and the corresponding values for two years ago (x_i) for each of the n=20 homes included in the sample. Estimate R, the relative change in assessed values for N=1000 homes using the information contained in the sample below. (10 Marks)

Home	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Assessed Value 2yrs ago	6.7	8.2	7.9	6.4	8.3	7.2	6.0	7.4	8.1	9.3	8.2	6.8	7.4	7.5	8.3	9.1	8.6	7.9	6.3	8.9
Current Value	7.1	8.4	8.2	6.9	8.4	7.9	6.5	7.6	8.9	9.9	9.1	7.3	7.8	8.3	8.9	9.6	8.7	8.8	7.0	9.4