SOUTH EASTERN KENYA UNIVERSITY

UNIVERSITY EXAMINATIONS 2016/2017

FIRST SEMESTER EXAMINATION FOR THE DEGREE OF BACHELOR OF BACHELOR OF SCIENCE IN STATISTICS.

STA 308: SAMPLE SURVEY THEORY AND METHODS 1

DATE: 14TH DECEMBER, 2016   **TIME: 10.30-12.30 PM**

INSTRUCTIONS:

Answer question 1 and any other 2 questions

Question 1 (30 marks)

a) Define the following terms;
      (i)   Simple random sampling.
      (ii)   Stratified sampling.
      (iii)   Cluster sampling.   (3 marks)

b) Give four advantages of using cluster sampling in surveys.   (4 marks)

c) Show that $S^2 = \frac{np(1-p)}{n-1}$

    Starting from $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_1 - \bar{y})^2$

    where y_1 is an indicator variable.   (5 marks)

d) Show that $var(\bar{y}) = \frac{\sigma^2}{n}$ under simple random sampling with replacement(SRSWR) procedure.   (5 marks)

e) (i) Give the difference between ratio separate and ratio combined estimators in stratified sampling.   (2 marks)

    (ii) Under which conditions should the two estimators (ratio separate and ratio combined) be used?   (4 marks)

f) A random sample of 11 unpaid medical claims was drawn from records at a certain university which has a total of 300 unpaid medical claims. The sample values for the amount owed was obtained in Kenya Shillings as follows:

690, 360, 1560, 1220, 5850, 3460, 280, 760, 140, 2680, 2480

   (i) Estimate the mean amount owed by the university on each claim and the total amount of money in medical claim owed by the university. (2 marks)
   (ii) Estimate the standard error of the estimate for the mean amount owed on each claim under SRSWOR procedure. (5 marks)

Question 2 (20 marks)

a) Give five disadvantages of using secondary data in any survey. (5 marks)
b) Derive the equation for determining the sample size under SRSWOR(Simple random sampling without replacement) and under SRSWR(Simple random sampling with replacement) procedures based on coefficient of variation . (7 marks)

c) The number of primary school children enrolled in Kenya in 1988 was 5123581. To estimate the expected total enrollment in 1989, a random sample of 58 schools was selected from a total of 14691 primary schools in the country. A summary of the sample information is given as follows,

Total enrolment in sampled schools in 1989 was 21886.

Total enrolment in sampled schools in 1988 was 20880. $\sum x_i^2 = 8075098$ $\sum y_i^2 = 8804869$

$\sum x_i y_i = 8246915$ $s_y^2 = 9584.2$

$s_x^2 = 9794.0$ $s_{xy} = 6455.0$

   (i)   Give an estimate for the population ratio. (2 marks)
   (ii) Find the population total in1989. (2 marks)
   (iii) Find the standard error of the estimator for the population total. (4 marks)

Question 3 (20 marks)

a) Give two advantages and two disadvantages of multi-stage sampling. (4 marks)
b) Show that the ratio estimator $\hat{R}$ is asymptotically unbiased. (4 marks)

c) The sales manager of a certain bread manufacturing company wishes to estimate the mean number of loaves of bread purchased by each household in a certain town in a month. Due to the diversity of the households the town is stratified in to three strata. Sample data on the number of loaves and the annual income for each household was obtained and summarized as follows:

Stratum	N_n	n_n	$\bar{y}$	$\bar{x}$	Sx_hy_h	Sy_h	Sx_h
1	685	50	30	Sh.10000	425	2	250
2	245	22	90	Sh.100000	4592	10	820
3	420	35	55	Sh.60000	2142	6	525

The total reported annual earnings in the town is sh. 15 million.

  (i)   Estimate the mean number of loaves of bread purchased by each household.
                                                                              (2 marks)
  (ii)   Find the variance and standard error of the combined estimator for the mean number of loaves per month.   (6 marks)
  (iii) If the total sample size taken is 107 households and the cost of surveying each household is sh. 100. Obtain the optimum allocation for minimum variance.
                                                                              (4 marks)

Question 4 (20marks)

In a study to investigate the effect of a new food of supposedly higher nutritional value for chicken were weighed before the new food was introduced. A few weeks after the food was introduced, the farm manager wanted to check if there was any gain in weight before continuing the study. A random sample of 12 chicken was taken and weighed as shown in the table below.

Pre X	1.54	1.78	2.0	1.64	1.51	1.79	1.92	1. 73	1.52	1.82	1.67	1.52
Post Y	2.81	2.20	3.42	3.14	2.68	2.79	3.11	3.61	3.00	3.52	2.95	2.48

The total weight of the 150 chicken at the beginning of the study was 255kg.

  a) Evaluate the ratio estimate of the total current weight and obtain its standard error.
                                                                              (7 marks
  b) Obtain the regression estimate of the total current weight and obtain its standard error, hence compare this with the estimate in a).   (13 marks)

Question 5 (20 marks)

a) A company has branches in 15 major towns. Each branch has a certain number of cars. The company wants to estimate the mean annual number of miles covered by each car. To do this, five branches were randomly selected and the total annual mileage obtained for all the cars. The data is given below with the mileage being in thousands of kilometers.

Town	m_i	y_i	Si^2
Nairobi	15	521	6.28
Mombasa	5	250	6.10
Kisumu	7	338	7.31
Nakuru	6	290	5.22
Kakamega	8	381	5.82

(i)   Find the unweighted estimate for the mean annual mileage for each car and its standard error.   (6 marks)

(ii)   Estimate the ratio estimator and its standard error.   (6 marks)

b) Consider the following data obtained from 8 clusters out of a total of 110 each having 6 elements.

Cluster	Elements
1	15, 5, 4, 3, 9, 10
2	5, 5, 3, 6, 8, 4
3	13, 5, 4, 2, 3, 7
4	10, 2, 3, 3, 4, 6
5	5, 10, 5, 11, 7, 9
6	3, 4, 4, 9, 11, 12
7	8, 9, 10, 6, 5, 14
8	4, 3,5, 6, 6, 16

Estimate the standard error of the population mean and calculate the intra-class correlation coefficient.   (8 marks)