UNIVERSITY EXAMINATIONS

SECOND SEMESTER 2023/2024 ACADEMIC YEAR

FIRST YEAR EXAMINATION FOR THE DEGREE OF BACHELOR OF SCIENCE (STATISTICS)

STAT 122: PRINCIPLES OF SAMPLE SURVEY THEORY

STREAM: R   **TIME: 2 HRS**

DAY: MONDAY [11.30A.M – 1.30 P.M]   **DATE: 08/04/2024**

THIS QUESTION PAPER CONSISTS OF FIVE (5) PAGES

PLEASE DO NOT OPEN UNTIL THE INVIGILATOR SAYS SO.

Vision   :   A University for Valued Transformation of Society   ***Page 1 of 5***
Mission: To serve students and society through research, education, scholarship, training, innovation, outreach and consultancy

Laikipia University is ISO 9001:2015 and ISO/IEC 27001:2013 Certified

INSTRUCTIONS

      (i)   Answer question one and any other two questions

      (ii)   Show all the workings clearly

      (iii)   Do not write on the question paper

QUESTION ONE (30 MARKS)

a) What is the difference between sample survey and census?   (**2 Marks**)

b) Describe two procedures of selecting units in a population.   (**4 Marks**)

c) In simple random sampling the mean of each possible random sample can be denoted by

$\bar{y}_k \{k=1,2,\ldots\ldots,N_{C_n}\}$ with associated probability $P_k = \dfrac{1}{N_{C_n}}$ for each.

Show that $E(\bar{y}) = \bar{Y}$. Comment on biasedness? (**5 Marks**)

d) A simple random sample of $n = 9$ hospital records is drawn to estimate the average amount due on N = 484 open accounts. The sample values for these nine records are listed below.

accounts	1	2	3	4	5	6	7	8	9
Amounts in dollars	33.5	32.0	52.0	43.0	40.0	41.0	45.0	42.5	39.0

    i)   Estimate the sample mean.   (**2 Marks**)

    ii)   The total cost of completion.   (**2 Marks**)

    iii)   The sample variance.   (**3 Marks**)

    iv)   Calculate $V(\bar{y})$ (**2 Marks**)

    v)   Calculate $V(\hat{Y})$.   (**2 Marks**)

    vi)   Standard error of the mean.   (**2 Marks**)

    vii)   Standard error of the total.   (**1 Mark**)

    viii) The confidence interval of the mean.   (**2 Marks**)

e) The average amount of money μ for a hospital's accounts receivable must be estimated. Although no prior data are available to estimate the population variance, it is known that most accounts lie within a $100 range. Find the sample size needed to estimate μ with a bound on the error of estimation B =$ 3 at 99% confidence level   .   (**3 Marks**)

Vision : *A University for Valued Transformation of Society*   *Page 2 of 5*
Mission: *To serve students and society through research, education, scholarship, training, innovation, outreach and consultancy*

Laikipia University is ISO 9001:2015 and ISO/IEC 27001:2013 Certified

QUESTION TWO (20 MARKS)

a) A study to assess the attitudes of accountants toward advertising their services involved sending questionnaires to 200 accountants selected from a list of 1400 names. A total of 82 usable questionnaires were returned.

    i) Estimate the proportion of accountants who returned the questionnaires and give its standard error of the estimate. **(3 Marks)**

    ii) Give the 95% confidence interval for the proportion of accountants who returned the questionnaires. **(2 Marks)**

b) The data below shows the number of persons (x), the weekly family income (z) and the weekly expenditure on food (y) in a sample random sample of 5 low-income families.

Families	1	2	3	4	5
X	5	4	7	2	4
Z	65	58	92	88	79
Y	30.5	41.2	28.2	24.2	30

Estimate from the data;

    i)   The mean weekly expenditure on food per family and its standard error. **(5 Marks)**
    ii)   The mean weekly expenditure on food per person and its standard error **(5 Marks)**
    iii)   The percentage of the income that is spent on food and its standard error. **(5 Marks)**

QUESTION THREE (20 MARKS)

a) Describe the following methods of determining the sample sizes in stratified random sampling.
    i) Proportional to size **(3 Marks)**

    ii) Optimum allocation (Neyman allocation) **(3 Marks)**

b)

STATUM	1	2	3
N_i	45	20	64
S_i	10	19	5

Vision : A University for Valued Transformation of Society   *Page 3 of 5*
Mission: *To serve students and society through research, education, scholarship, training, innovation, outreach and consultancy*

Laikipia University is ISO 9001:2015 and ISO/IEC 27001:2013 Certified

For a fixed sample size of n=50, find n_i under

 i) Proportional allocation **(3 Marks)**
 ii) Optimum allocation (Neyman allocation) **(3 Marks)**

b) What is the difference between stratified and cluster sampling. **(4 Marks)**
c) Describe the two main sources of errors in sample survey **(4 Marks)**

QUESTION FOUR (20 MARKS)

a) i) Describe stratified random sampling technique. **(3 Marks)**

 ii) State **three** reasons for stratification in sample survey. **(3 Marks)**

b) The advertising firm wants to estimate the proportion of households in the county that view show X. The county is divided into three strata, town A. town B. and the rural area. The strata contain $N_1 = 155$, $N_2 = 62$, and $N_3 = 93$ households, respectively. A stratified random sample of $n = 40$ households is chosen with proportional allocation. In other words, a simple random sample is taken from each stratum; the sires of the samples are $n_1 = 20$, $n_2 = 8$, and $n_3 = 12$. Interviews are conducted in the 40 sampled households; results are shown in Table below.

Number of households

Stratum	Sample size	viewing show X	$\hat{p}_i$
1	$n_1 = 20$	I6	0.80
2	$n_2 = 8$	2	0.25
3	$n_3 = 12$	6	0.50

 i. Estimate the proportion of households viewing show X. **(3 Marks)**
 ii. Estimate the proportion variance for each stratum. **(9 Marks)**
 iii. Compute the estimated variance of the stratified proportion. **(2 Marks)**

Vision : A University for Valued Transformation of Society   *Page 4 of 5*
Mission: To serve students and society through research, education, scholarship, training, innovation, outreach and consultancy

Laikipia University is ISO 9001:2015 and ISO/IEC 27001:2013 Certified

QUESTION FIVE (20 MARKS)

a) The director of a computer support department plans to sample three divisions of a large firm that has ten divisions with varying numbers of employees per division. Since the 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 probability sampling with replacement with p_i proportional to number of employees in that division.

Division	No. of employees
1	1000
2	650
3	2100
4	860
5	2840
6	1910
7	390
8	3200
9	1500
10	1200

Three random values were generated as follows, 1200, 7001 and 12787. We decide to sample the three divisions; For division 2, $y = 420$, for division 5, $y = 1785$, and for division 8, $y = 2198$,

i) Show how the values 1200, 7001 and 12787 fall in divisions 2, 5 and 8 respectively

(4 Marks)

ii) Compute the mean and variance of the sample using Hansen-Hurwitz estimator. Use it to find population variance estimate. **(8 Marks)**

b) Describe the following terms as used in sample survey
   i. Population
   ii. Sample
   iii. Sampling frame
   iv. Sampling unit **(8 Marks)**

Vision : A University for Valued Transformation of Society   *Page 5 of 5*
Mission: To serve students and society through research, education, scholarship, training, innovation, outreach and consultancy

Laikipia University is ISO 9001:2015 and ISO/IEC 27001:2013 Certified