



MACHAKOS UNIVERSITY

University Examinations 2017/2018

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS AND STATISTICS

**FOURTH YEAR SECOND SEMESTER EXAMINATIONS FOR BACHELOR OF
SCIENCE (STATISTICS AND PROGRAMMING)**

SST 408: DESIGN AND ANALYSIS OF SAMPLE SURVEY II

DATE: 14/12/2017

TIME: 2.00-4.00 PM

INSTRUCTIONS: Answer Questions ONE and any other TWO questions

QUESTION ONE (30 MARKS)

- a) Define the following terms as used in surveying (4 marks)
 - i) Study Population
 - ii) Sampling frame
 - iii) Stratified random sampling
 - iv) Frame
- b) Discuss two types of errors that occurs that typically occur with survey studies (4 marks)
- c) Discuss Five advantages of sampling over complete enumeration (5 marks)
- d) Briefly explain Four non-sampling errors in sample survey (4 marks)
- e) Describe the ratio estimator (3 marks)
- f) Signatures to a petition were collected on 676 sheets each sheet had enough space for 42 signatures but on many sheets a smaller number of signatures were collected. The number of signatures per sheet were counted on a simple random sample of 50 sheets and results are shown below.

No of signatures	3	4	5	6	7	9	10	11	14	15	16	19	23	27	29	32	36	41	42
Frequency (f_i)	1	1	2	3	1	1	1	1	1	2	2	1	1	2	1	1	1	4	23

- Estimate the population mean (2 marks)
- Estimate the total number of signatures (2 marks)
- Estimate the population variance (6 marks)

QUESTION TWO (20 MARKS)

- Discuss Four advantages of stratified random sampling (4 marks)
- A researcher is allocated Ksh 2000 to contact a survey of stratified random sample. The population consist of stratum A of size 40,000 stratum B of size 20,000 and stratum C of size 10,000. The setup cost of administering the survey is Ksh 200 and the cost of sampling one unit is Ksh 2.25, Ksh 4.00 and Ksh 1.00 for stratum A, B and C respectively the standard deviation of stratum A is found to be twice that for observation in stratum B and C
 - Estimate the optimum allocation (8 marks)
 - Proportional allocation assuming that all the money is spend (8 marks)

QUESTION THREE (20 MARKS)

- A population of size 800 is divided into 3 strata. Their size and standard deviation are shown below

Strata	Size	S.D
1	200	6
2	300	8
3	300	12

A stratified random sample of size 120 is to be taken from the population. Determine the size of samples from the strata using Neyman allocation. (6 marks)

- b) The average amount of money for a hospital account receivable must be estimated there are $N=1000$ open accounts. Find the sample size needed to estimate $\bar{Y}$ with bound on the error of estimation $\beta = 3$ and $S^2 = 625$. Hence find the 80% confidence interval for mean and population total (14 marks)

QUESTION FOUR (20 MARKS)

- a) In a simple random sampling without replacement, verify $\bar{y}$ that is an unbiased estimator of $\bar{Y}$ and that its sample variance is given as $Var(\bar{y}) = \frac{N-n}{N} \frac{S^2}{n}$ where $S^2 = \frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})^2$ (10 marks)
- b) An advertising firm is interested in determining how much on TV advertisement. The firm decide to conduct a sample survey to estimate the average number of hours per week that household watch television. The household are divided into 3 distinct strata, the simple random samples are selected and interviews conducted and results were as follows

	1	2	3
n_i	20	8	12
$\bar{y}_i$	33.9	25.125	19
S^2	35.358	232.411	87.636
N_i	155	62	93

- i) Estimate the average viewing time for all the household and the total viewing time (2 marks)
- ii) Calculate the estimator variance (4 marks)
- iii) Estimate the total number of hours per week that households in the country view T.V and place a bound on the error of estimation (4 marks)

QUESTION FIVE (20 MARKS)

A nurseryman wants to estimate the average height of seedlings in a large field that is divided into 50 plots that vary slightly in size He believes that the heights are fairly constant throughout each plot but may vary considerably from plot to plot Therefore, he decides to sample 10% of the trees within each of 10 plots using two stage cluster sample. The data are given in the table below.

Estimate the average height of seedlings in the field and place a bound on the error of estimation
(20 marks)

Plot	Number of seedlings	Number of seedlings sampled	Heights of seedlings (inches)
1	61	6	12,11,12,13,12,12
2	60	6	7,10,8,9,9,10
3	45	6	9,10,8,9,9,10
4	49	5	6,7,6,8,7
5	51	5	14,15,13,12,13
6	50	5	10,11,13,12,12
7	56	5	7,8,7,7,6
8	60	6	6,5,7,5,6,4
9	46	6	10,9,7,9,8,10
10	52	5	12,11,12,10,13