



**UNIVERSITY EXAMINATIONS  
SPECIAL EXAMINATION**

**THIRD YEAR EXAMINATION FOR THE AWARD OF THE  
DEGREE OF BACHELOR OF SCIENCE MATHEMATICS /BACHELOR OF  
SCIENCE APPLIED STATISTICS  
SECOND SEMESTER 2021/2022  
(JULY, 2022)**

**MATH 341/STAT 341: SAMPLING METHODS1**

**STREAM: Y3 S2**

**TIME: 2 HOURS**

**DAY: FRIDAY, 11.30 AM – 1.30 PM**

**DATE: 29/07/2022**

---

**INSTRUCTIONS:**

- 1. Do not write anything on this question paper.**
- 2. Answer Question ONE (Compulsory) and any other TWO Questions.**

**QUESTION ONE (30 MARKS)**

- a) Differentiate between the following terms as used in sampling theory
- i. Population and sampling frame (2marks)
  - ii. Probability sampling and nonrandom sampling (2 marks)
  - iii. Stratified sampling and systematic sampling (2 marks)
- b) A large construction company has 120 houses in various stages of completion. For estimation of the total amount spent in the construction process, a simple random sample of 12 houses is selected without replacement and the accumulated costs determined on each. The following were the costs in million Kshs obtained:
- 35.5 30.2 28.9 36.4 29.8 34.1 32.6 26.4 38.0 38.2 32.2 27.5
- i. Calculate the sample mean and variance (3 marks)

- ii. Estimate the total construction cost and its approximate 95% confidence limits  
(4 marks)
- c) A dentist is interested in the effectiveness of a new tooth paste. A group of 1000 school children in Kisii County are identified as the sampling frame for the study. It is established that the variance of the number of cavities is 2.12.
- i. The sample size if the  $\text{Var}(\bar{y})$  should not exceed 0.125 (3 marks)
- ii. The sample size for an estimation error of 0.0125 at 5% level of significance.  
(3 marks)
- iii. Estimate the total number of cavities in the whole group of students and the 99% confidence limits.  
(5 marks)
- d) a population of size 500 is divided into two colleges A and B. their sizes, mean and standard deviations are as follows:

College	NO. of Students	Mean	Standard deviation
A	330	60	12
B	170	45	20

A stratified random sample of size 80 is to be drawn from the population. Determine the sizes of samples from the strata based on,

- i. Neyman allocation (3 marks)
- ii. proportional allocation (3 marks)

### QUESTION TWO (20 MARKS)

- a) A survey seeks to establish the level of infection with Covid-19 in 6 villages A, B, C, D, E and F. The total number of people infected in the villages were found to be 3, 2, 1, 7, 4 and 5 respectively.

If a sample of 4 villages is drawn using simple random sampling without replacement, verify that  $E(\bar{y}) = \bar{Y}$  (8 Marks)

- b) For simple random sampling show that:

- i.  $\text{Var}(\bar{y}) = \frac{N-n}{Nn} S^2$  under simple random sampling without replacement  
(5 Marks)
- ii. simple random sampling without replacement is more efficient than simple random sampling with replacement (3 Marks)

iii.  $E(s^2) = S^2$ , where  $s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$  and  $S^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{Y})^2$

(5 Marks)

**QUESTION THREE (30 MARKS)**

- a) If  $V_{SRS}(\bar{y})$ ,  $V_{prop}(\bar{y})$  and  $V_{opt}(\bar{y})$  are the variances of the sample mean under simple random sampling without replacement, stratified sampling with proportional allocation and stratified sampling with optimum allocation respectively.

Show that

i.  $V_{prop}(\bar{y}) \leq V_{SRS}(\bar{y})$  (4 marks)

ii.  $V_{opt}(\bar{y}) \leq V_{prop}(\bar{y})$  (4 marks)

- b) A total number of coffee tree in all 64 plantations is to be estimated from a sample size of 24. The farms are categorized into two strata: Large scale plantations which were 16 (Stratum 1) and small scale plantations which were 48 (Stratum 2). The number of coffee trees in the farms based on each stratum are:

Stratum 1	Stratum 2		
900	364	209	113
822	317	183	115
781	328	163	123
805	302	253	154
670	288	232	140
1238	291	260	119
573	253	201	130
634	291	147	127
578	308	292	100
487	272	164	107
442	284	143	114
451	255	169	111
459	270	139	163
464	214	170	115
400	195	150	122
366	260	143	134

Find the standard error of the estimated total for

- i. A simple random sample (4 marks)
- ii. A stratified random sample with proportional allocation (4 marks)
- iii. A stratified random sample with optimum allocation (4 marks)

**QUESTION FOUR (30 MARKS)**

a) Show that:

- i. the variance of the mean of a systematic sample is:

$$V(\bar{y}_{sy}) = \frac{N-1}{N} S^2 - \frac{k(n-1)}{N} S_{wsy}^2$$

where,  $S_{wsy}^2 = \frac{1}{k(n-1)} \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2$  is the variance among units that lie within the same systematic sample. (7 marks)

- ii. Systematic sampling is more efficient than the simple random sampling without replacement if and only if  $S_{wsy}^2 > S^2$  (5 marks)

b) Suppose we take a 1 in every 5 systematic sample from a population of 25 physicians whose registration numbers are 1 to 25 and that the first random number chosen is 3. The following table gives the physicians in the sample along with their household visits.

Physician in the sample	Number of visits ( $y_i$ )
3	1
8	0
13	6
18	0
23	0

- i. Estimate the variance of the mean number of visits (5 Marks)
- ii. Find the 99% confidence interval for the total number of visits made by the 25 physicians. (3 Marks)

**QUESTION FIVE (20 MARKS)**

A. A group of Actuarial Science students of Kisii university are planning to conduct a survey on students' opinions on dining facilities.

- I. outline the steps they will follow in developing a sampling plan (4 marks)

- II. if the sampling frame was 4000 students and the estimated proportion of students in the sample who identified the dining facilities in the university as adequate (P) was 0.465.

Find

- i. the sample size if a margin of error (d) of 0.05 is allowed at  $\alpha = 5\%$ . (3 marks)
- ii. the 95% confidence limits for the number of students in the population who considered the dining facilities as adequate. (2 marks)

B. Table below shows the stratification of some factories in Kenya for the year 2021 by industry group (according to the National Industrial and Occupational classification)

Industry group	Number of Factories	Average Number of Factories $\bar{Y}_h$	Standard Deviation $S_h$
A	327	92.1	200.8
B	139	58.7	140.3
C	203	19.5	36.7
D	170	49.5	94.5
E	177	142.2	372.6
F	89	45.8	53.4

For a sample of 500 factories, verify that  $V_{opt}(\bar{y}) \leq V_{prop}(\bar{y}) \leq V_{SRS}(\bar{y})$  (11 marks)