



THE CATHOLIC UNIVERSITY OF EASTERN AFRICA
A. M. E. C. E. A

P.O. Box 62157
00200 Nairobi - Kenya
Telephone: 891601-6
Fax: 254-20-891084
e-mail:academics@cuea.edu

GABA CAMPUS - ELDORET
MAIN EXAMINATION
SEPTEMBER – DECEMBER 2021 TRIMESTER
FACULTY OF SCIENCE
BACHELOR OF SCIENCE
DEPARTMENT OF MATHEMATICS AND ACTUARIAL SCIENCE
MAT 364: DESIGN AND ANALYSIS OF SAMPLE SURVEYS

Date: December 2021	Duration: 2 Hours
----------------------------	--------------------------

Instructions: Answer Question ONE and any other TWO Questions
--

QUESTION ONE

- a) Define the following terms.
- i) Sampling theory (1 mark)
 - ii) Sample survey (1 mark)
 - iii) Sample unit (1 mark)
 - iv) Sample frame (1 mark)
 - v) Sample error (1 mark)
 - vi) Unbiased estimate (1 mark)
 - vii) Purposive sampling (1 mark)
- b) Give two Disadvantages of Sample Surveys compared with Censuses (2 marks)
- c) Give two disadvantages of personal interviews as a method of survey administration (2 marks)
- d) Give three measures of Reducing Errors in Surveys (3 marks)
- e) Differentiate between one-stage and two-stage cluster sampling (2 marks)
- f) Suppose that we have a population of size $N = 4$ whose population units are $\{1, 2, 3, 4\}$ and that we require a sample of size $n = 2$ from the population. Assuming we use simple random sampling without replacement (SRSWOR).

Find

- i) The number of possible samples and the probability of selecting each sample **(2 marks)**
 - ii) Specify these samples **(2 marks)**
 - iii) Show that the sample mean is an unbiased for the population mean using this data **(5 marks)**
- g) 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. There are $N = 1000$ open accounts. Find the sample size needed to estimate μ with a bound on the error of estimation $B = \$3$. **(5 marks)**

QUESTION TWO

- a) Discuss the steps of a survey **(10 marks)**
- i) Give two advantages and one disadvantage of systematic sampling **(3 marks)**
 - ii) A 1-in-6 systematic sample is obtained from a voter registration list to estimate the proportion of voters in favor of the proposed bond issue. Several different random starting points are used to ensure that the results of the sample are not affected by periodic variation in the population. The coded results of this pre-election survey are as shown in the accompanying table. Estimate p , the proportion of the 5775 registered voters in favor of the proposed bond issue ($N = 5775$) Place a bound on the error of estimation. **(7 marks)**

Voter	Response
4	1
10	0
16	1
-	-
-	-
-	-
5760	0
5766	0
5772	1
$\sum_{i=1}^{962} y_i = 652$	

QUESTION THREE

- a) A manufacturing company wishes to estimate the ratio of change from last year to this year in the number of worker hours lost due to sickness. A preliminary study of $n = 10$ employee records is made, and the results are given in the table below. The company records show that the total number of worker-hours lost due to sickness for the previous year was $\tau x = 16300$. Use the data to determine the sample size required to estimate R , the rate of change for the company, with a bound on the error of estimation of magnitude $B = 0.01$. Assume the company has $N = 1000$ employees. **(10 marks)**

Employee	Worker-hour lost in previous year, x	Worker-hour lost in current year, y
1	12	13
2	24	25
3	15	15
4	30	32
5	32	36
6	26	24
7	10	12
8	15	16
9	0	2
10	14	12

- b) The U.S. government's American Housing Survey keeps tabs on many aspects of the characteristics of housing in America, including monthly costs for home ownership and the value of houses. One aspect of the survey tracks 47 metropolitan statistical areas (MSAs) over time by sampling a subset of them every four years or so. The survey for 2002 sampled the 13 MSAs and the results are listed in the table below. Given $sr = 67.8$
- i) Estimate R , the ratio of mean typical monthly costs for 2002 as compared to those of 1994 for all 47 MSAs and calculate an appropriate margin of error. **(7 marks)**

	n	Mean
Y= 2002 monthly	13	901.5
X= 1994 monthly	13	695.8

- ii) Determine the coefficient of variation for x . **(3 marks)**

QUESTION FOUR

- i) Define a simple random sample without replacement
(2 marks)
- ii) Show that in simple random sampling without replacement, the sample mean is an unbiased estimator of the population mean
(5 marks)
- iii) A simple random sample of $n = 9$ hospital records is drawn to estimate the average amount of money due on $N = 484$ open accounts. The sample values for these nine records are listed in the table below. Estimate μ , the average amount outstanding, and place a bound on your error of estimation.
(6 marks)

Amount of money owed	
y_1	33.50
y_2	32.00
y_3	52.00
y_4	43.00
y_5	40.00
y_6	41.00
y_7	45.00
y_8	42.50
y_9	39.00

- iv) Student government leaders at a college want to conduct a survey to determine the proportion of students who favor a proposed honor code. Because interviewing $N = 2000$ students in a reasonable length of time is almost impossible, determine the sample size (number of students to be interviewed) needed to estimate p with a bound on the error of estimation of magnitude $B = 0.05$. Assume that no prior information is available to estimate p .
(7 marks)

QUESTION FIVE

- i) Define stratified sampling
(2 marks)
- ii) Give three reasons of using stratified sampling rather than simple random sampling
(3 marks)

- a) An advertising firm, interested in determining how much to emphasize television advertising in a certain county, decides to conduct a sample survey to estimate the average number of hours each week that households within the county watch television. The county contains two towns, A and B, and a rural area. Town A is built around a factory, and most households contain factory workers with school-age children. Town B is an exclusive suburb of a city in a neighboring county and contains older residents with few children at home. There are 155 households in town A, 62 in town B, and 93 in the rural area. The results are given in the following table:

	N	N	Mean	Median	SD
Town A	155	20	33.90	34.50	5.95
Town B	62	8	25.12	26.00	15.25
Rural	93	12	19.00	17.50	9.36

- i) Estimate the average television-viewing time in hours per week for
- All households in the county and place a bound on the error of estimation **(8 marks)**
 - All households in town B and place a bound on the error of estimation **(4 marks)**
- ii) The advertising firm wants to estimate the proportion of households in the county that view show X. Interviews are conducted in the 40 sampled households; results are shown in the table below. Estimate the proportion of households viewing show X, and place a bound on the error of estimation. **(3 marks)**

stratum	Sample size n_1	Number of households viewing show x	P_i
1	20	16	0.80
2	8	2	0.25
3	12	8	0.50

END