



MACHAKOS UNIVERSITY

University Examinations 2020/2021 Academic Year

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS AND STATISTICS

THIRD YEAR SPECIAL /SUPPLEMENTARY EXAMINATION FOR
BACHELOR OF SCIENCE IN MATHEMATICS AND COMPUTER SCIENCE
BACHELOR OF SCIENCE IN STATISTICS AND PROGRAMMING
BACHELOR OF SCIENCE IN MATHEMATICS
BACHELOR OF EDUCATION (ARTS, SCIENCE)

SMA 365: DESIGN AND ANALYSIS OF SAMPLE SURVEYS

DATE: 25/3/2021

TIME: 8.30-10.00 AM

INSTRUCTIONS

1. Answer **Question 1** and any other **two** questions.
2. Out of the **three** questions answered, each question must start on a new page.
3. You need a Scientific Calculator and Statistical Tables for this paper

QUESTION ONE (30 MARKS)

- a) Explain the importance of a *pre-test* in the collection of statistical data in a survey. (6 marks)
- b) With the aid of an example in each case, differentiate between each of the following terms as used in the collection of statistical data.
- (i) Observational unit and sampling unit;
 - (ii) Census and sample survey. (6 marks)
- c) Given an SRS-WOR with a sample mean $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ and a population mean $\bar{X} = \frac{1}{N} \sum_{i=1}^N x_i$. Show that the sample mean $\bar{x}$ is an unbiased estimator of the population mean $\bar{X}$. (4 marks)

- d) Given a stratified random sampling with a stratified mean given by $\bar{X}_{st} = \frac{1}{N} \sum_{i=1}^m N_i \bar{x}_i$ and a population mean given by $\bar{X} = \frac{1}{N} \sum_{i=1}^m X_i$, show that the stratified mean $\bar{X}_{st}$ is an unbiased estimator of the population mean $\bar{X}$. (4 marks)
- e) Using a population of size $N = 5$ comprising the values 24, 21, 18, 27, 15, from which a random sample of size $n = 3$ is to be selected, determine all the possible samples of size 3 which can be drawn from the population.
- Hence using the selected samples, prove that the sample mean $\bar{x}$ is an unbiased estimator of the population mean $\bar{X}$. (10 marks)

QUESTION TWO (20 MARKS)

- a) Explain the term *multi-stage sampling* as used in the collection of statistical data. (2 marks)
- b) Explain the circumstance under which each of the following sampling techniques is suitable in the collection of statistical data.
- Simple random sampling;
 - Stratified random sampling. (5 marks)
- c) A survey was conducted to determine the number of mobile phone SIM cards used by households among Kenyan urban residents. A simple random sample of 25 households was drawn from a residential estate comprising 800 households. The number of phone SIM cards owned in a household in the sample was as follows:
- 8, 6, 5, 7, 9, 4, 8, 5, 4, 6, 9, 3, 8, 9, 6, 5, 7, 9, 6, 4, 8, 4, 6, 5, 9.
- Estimate the total number of phone SIM cards owned by the households in the residential estate; (3 marks)
 - Determine the 95% confidence limits for the population mean of the number of phone SIM cards owned by residents in the residential estate; (5 marks)
 - Determine the 95% confidence limits for the population total of the number of phone SIM cards owned by residents in the residential estate. (5 marks)

QUESTION THREE (20 MARKS)

- a) Explain the term *questionnaire* as used in the collection of statistical data. (2 marks)
- b) Explain the **three** errors which may occur in a sampling in a sample survey, illustrating each with an example from a real life situation. (9 marks)
- c) Given an SRS-WOR with a sample variance given by $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ and its corresponding population variance given by $S^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{X})^2$, show that the sample variance s^2 is an unbiased estimator of the population variance S^2 . (9 marks)

QUESTION FOUR (20 MARKS)

- a) Explain the term *sampling frame* as used in a sample survey, giving *two* examples from real life situations. (4 marks)
- b) Given a two-stage sample in which the sample mean is given by $\bar{\bar{x}} = \frac{1}{n m} \sum_{i=1}^n \sum_{j=1}^m x_{ij}$ and its corresponding population mean is given by $\bar{\bar{X}} = \frac{1}{N M} \sum_{i=1}^N \sum_{j=1}^M x_{ij}$, show that the sample mean $\bar{\bar{x}}$ is an unbiased estimator of the population mean $\bar{\bar{X}}$. (6 marks)
- c) A survey to be conducted among patients in a county hospital, with the patients categorised as children, teenagers, middle age and old age. A sample of size $n = 20$ is to be selected from a population of size $N = 800$. Select an *actual sample* procedurally stating clearly the range of your sampling frame and the selected sampling units using each of the following sampling techniques:
- (i) Simple random sampling; (6 marks)
 - (ii) Systematic random sampling; (4 marks)

QUESTION FIVE (20 MARKS)

A survey was conducted to determine the number of cattle kept in a homestead among a pastoralist community in Kenyan. A stratified random sample of size 140 homesteads was drawn from a target population of size 2,800 homesteads. This target population is divided into five strata as shown below:

Stratum	A	B	C	D
Stratum size	800	900	600	500
Stratum sample mean	64	75	80	74
Stratum standard deviation	4	5	8	3

- a) Determine the sample size for each stratum if the sample is to be drawn using stratified random sampling technique with:
- (i) proportional allocation;
 - (ii) optimum allocation with fixed cost. (10 marks)
- b) Using the sample sizes obtained under proportional allocation, determine the 95% confidence limits for each of the following measures about the pastoralists:
- (i) the population mean of the number of cattle kept in a homestead; (5 marks)
 - (ii) the population total of the number of cattle kept in a homestead. (5 marks)