



(Knowledge for Development)

KIBABII UNIVERSITY
UNIVERSITY EXAMINATIONS
2021/2022 ACADEMIC YEAR
FIRST YEAR FIRST SEMESTER
MAIN EXAMINATION
FOR THE DEGREE OF MASTER OF SCIENCE IN STATISTICS

COURSE CODE: STA 811

COURSE TITLE: SAMPLE SURVEY

DATE: 27/05/2021

TIME: 9:00 AM - 11:00 AM

INSTRUCTIONS TO CANDIDATES

Answer Question One any other Two Questions

TIME: 2 Hours

QUESTION ONE (30 MARKS)

(a) Discuss the sources and types of errors in sample surveys (10mks)

(b) Consider sample values $y_1, y_2, y_3, \dots, y_n$ from the population

(i) Show that the sample mean \bar{y} is unbiased estimate of the population mean \bar{Y} (4mks)

(ii) Obtain the variance of the sample mean (5mks)

(c) Consider a relatively large sample of size n . Let the sample be randomly divided into K groups each of size m units such that $n = mk$

Let \hat{S}^2 be the estimator of population variance S^2 and be defined as

$$\hat{S}^2 = \frac{m}{k-1} \sum_{i=1}^K (y_i - \bar{y})^2$$

Show that $E(\hat{S}^2) = S^2$. Comment on the result. (6 mks)

(d) The variance S^2 of a sample of size n may be given by

$S^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$ verify that the sample variance, S^2 is an unbiased estimator of the population variance, S^2 (5 mks)

QUESTION TWO (20 MARKS)

(a) state the main differences between a parameter and a statistic. Discuss at least three properties required for a good estimator. (9mks)

(b) An agricultural company has 45 farms of different sizes totaling one million hectares. the company wishes to determine the amount of land on its farms which is not being adequately used by using a PPS (probability proportion to size) sample size of 15 farms. The size of the farm X_i and the corresponding hectares of land Y_i which is not being optimally used are given below in thousands of hectares.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X_i	54	90	75	60	45	38	66	22	42	32	46	27	18	15	29
Y_i	33	15	31	10	8	11	18	2	6	9	12	3	3	2	4

Obtain an estimate of the total size of land which is not being optimally utilized. Estimate its corresponding standard error. (7mks)

(c) Why is necessary to carry out a pilot survey first? (4mks)

QUESTION THREE (20 MARKS)

a) Prove that the variance of the sample proportion p is given by

$$\text{Var}(p) = \frac{S^2}{n} \left(1 - \frac{n}{N}\right) = \frac{PQ}{n} \cdot \frac{N-n}{N-1}$$

$$\text{Where } S^2 = \frac{\sum_{i=1}^N y_i^2 - N\bar{y}^2}{N-1}$$

And P is the probability that a unity has an attribute that is required; $Q = 1-P$. (9 mks)

b) To estimate the proportion of diseased citrus plants in a certain farm, a random sample of $n = 75$ plants selected by the SRSWR was examined and 55% of them were found to be diseased, obtain an estimate of the proportion of the diseased plants and an estimate of the standard error. (6 mks)

c) State the advantages of using stratified sampling techniques (5 mks)

QUESTION FOUR (20 MARKS)

Discuss the Bootstrapping method in sample survey (20mks)

QUESTION FIVE (20 MARKS)

Let $(x_1, y_1) (x_2, y_2) \dots \dots \dots (x_n, y_n)$ be a random sample of size n on a paired variable (X, Y) drawn preferably by simple random sampling without replacement (SRSWOR) from a population of size N . The regression estimate of population means \bar{Y}_{reg}

is defined as $\hat{Y}_{reg} = \bar{Y} + B_0 (X - \bar{x})$

Where B_0 is known

Required:

- a) Find the Bias of \bar{Y}_{reg}
- b) Variance of \bar{Y}_{reg}
- c) Optimal value of B_0
- d) Comparison with SRSWOR