

**NATIONAL UNIVERSITY OF SCIENCE AND
TECHNOLOGY**

DEPARTMENT OF APPLIED MATHEMATICS

SMA 2205 SURVEY METHODS

MAY/JUNE 2004

Time: 3 hours

Answer **ALL** questions in Section A and **ANY THREE** questions from section B

SECTION A [40 marks]

Answer **ALL** questions from this section

A1. Briefly explain the meaning of each of the following:

- (a) Analytic survey.
- (b) Sampling error.
- (c) Overcoverage in a sampling frame.
- (d) Sampling frame.
- (e) Probability sampling.

[10 marks]

A2. Suppose y_1, \dots, y_n denotes a simple random sample from a population of size N with mean μ and variance σ^2 . Show that

$$\hat{Var}(\hat{T}) = \frac{s^2}{n} N(N-n)$$

is an unbiased estimator of $Var(\hat{T})$.

[4 marks]

A3. With reference to a survey on students social life at a certain university, criticize the following question proposed for inclusion in a questionnaire.
Where did you sleep last week?

- (i) on campus residence
- (ii) off campus residence

[2 marks]

A4. In a study to estimate the total sugar content of a truckload of oranges, a random sample of $n = 10$ oranges was juiced and weighted. The details are as shown in the table below

Orange	Sugar Content in kg (y_i)	Weight of Orange in kg (x_i)
1	0.021	0.40
2	0.030	0.48
3	0.025	0.43
4	0.022	0.42
5	0.033	0.50
6	0.027	0.46
7	0.019	0.39
8	0.021	0.41
9	0.023	0.42
10	0.025	0.44

The total weight of all the oranges obtained by first weighing the truck loaded and then unloaded, was found to be 1800kg. Estimate T , the total sugar content for the oranges and find a 95% confidence interval. **[12marks]**

A5. Let X_1, \dots, X_n be a random sample obtained from an exponential distribution with parameter θ .

(a) Find the Cramer-Rao lower bound of the variance of unbiased estimators of

$$\tau(\theta) = \frac{1}{\theta} = E(X).$$

(b) Show that \bar{X} is a minimum variance unbiased estimator of $E(X)$. **[8 marks]**

SECTION B [60marks]

Candidates may attempt any **THREE** questions from this section

B6. (a) What is a cluster sample? Under what circumstances are cluster samples advantageous?

(b) A certain type of circuit board manufactured for installation in computers has 12 microchips per board. During the quality control inspection of 10 of these boards was done. Assume there is a shipment of 50 such boards. The number of defective microchips per board were as follows:

2, 0, 1, 3, 2, 0, 0, 1, 3, 4

- (i) Estimate the proportion of defective microchips in the population from which this sample was drawn and find a 95% confidence interval.
(ii) Estimate the total number of defective microchips in the shipment and find a 95% confidence interval. **[20 marks]**

B7. (a) Define stratified random sampling. Under what circumstances are stratified random samples advantageous?

- (b) A random sample of 40 workers is selected from a work force of 1200 workers. The mean force vital capacity y_i , based on age, sex and height, was calculated for each worker. Given that $\sum y_i = 3240$ and $\sum y_i^2 = 268026$ estimate the population mean force vital capacity and place a bound on the error of estimation. Use 10% level of significance. Suppose in addition the workers were categorized as high, medium or low with respect to exposure to pulmonary stress giving the following results:

		GROUP		
		Low	Medium	High
Group Size	N_i	100	100	1000
Sample Size	n_i	6	6	28
	$\sum y_i$	500	501	2239
	$\sum y_i^2$	42640	42447	182939

Using exposure levels as strata, estimate the stratified population mean force vital capacity and place a bound on the error of estimation. Use 10% level of significance. [20 marks]

- B8.** (a) With reference to systematic sampling, explain what is meant by each of the following terms:
- Random population.
 - Ordered population.
 - Periodic population.
- (b) Outline a procedure for selecting a repeated systematic random sample.
- (c) In a survey, a 1-in-10 systematic sample is drawn from an alphabetic list of NUST students. It is found that 45.2% of the students in the sample agree to paying \$5 for photocopies. Given a sample size of 600 students estimate T the total number of students who agree to this price and find a 95% confidence interval for T [20 marks]
- B9.** (a) Suppose y_1, \dots, y_n denotes a simple random sample from a population of size N with mean μ and variance σ^2 . Show that $Var(\bar{y}) = \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right)$.
- (b) In order to count the number of baobab trees, above a specified height, in Zimbabwe, a survey was carried out. The survey consisted of counting the number of all baobab trees in a sample of blocks of about 4km^2 area.
- Describe how you could perform a simple random sampling procedure in this case.

(ii) Chipinge area in Manicaland province consist of 500 blocks, 50 of these are included in the sample. Let y_i be the number of observed baobab trees in block i .

In the sampled blocks $\sum_{i=1}^{50} y_i = 350$, $\sum_{i=1}^{50} y_i^2 = 3600$

Estimate the total number of baobab trees in Chipinge area and give a 99% confidence interval for that number. [20 marks]

END OF EXAMINATION