

NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY

DEPARTMENT OF APPLIED MATHEMATICS

SMA 2205 SURVEY METHODS

MAY 2005

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) Descriptive survey.
- (b) nonsampling error.
- (c) Element.
- (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

$$E(s^2) = \frac{N}{N-1} \sigma^2.$$

[7 marks]

A3. A survey was conducted to gather information concerning food supplies for undergraduate students at a certain university. Resources are available for selecting a sample of 1000 students. Suggest and describe how to implement a suitable sampling method for this survey.

[6 marks]

A4. The table below contains raw data collected in a survey on HIV status of patients at hospitals in two cities in Zimbabwe. The research objectives included an investigation into the percentage of patients who were HIV-positive. Summarize the information provided in a form suitable for presentation in a report and comment on your findings.

Gender	CITY A			CITY B		
	HIV STATUS			HIV STATUS		
	Positive	Negative	Total	Positive	Negative	Total
Males	90	86	176	118	120	238
Females	32	19	51	81	67	148
Total	122	105	227	199	187	386

[10marks]

A5. Let X_1, \dots, X_n be a random sample from the Bernoulli distribution

$$P(X = x) = p^x(1-p)^{1-x}, x = 0, 1 \quad 0 < p < 1.$$

Show that $\hat{p} = \frac{1}{n} \sum_{i=1}^n X_i$ is an unbiased and consistent estimator of p . [7 marks]

SECTION B [60marks]

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

B6. (a) With reference to systematic sampling, explain what is meant by each of the following terms:

- (i) Random population,
- (ii) Ordered population,
- (iii) Periodic population.

(b) Outline a procedure for selecting a repeated systematic random sample.

(c) The headmaster of a certain school is interested in estimating the proportion of students favoring civics day. A 1-in-10 systematic sample of students selected as they leave the school at the end of the day is interviewed. The information is shown in the table below.

Student No. Sampled	Response $y_i = 1$ if in favour, 0 not in favour
6	0
16	1
26	1
.	.
.	.
.	.
1996	0
	$\sum y_i = 164$

- (i) Estimate p the proportion of students in favour of civics day and find a 95% confidence interval,

- (ii) Determine the sample size required to estimate p to within a margin of error of 0.01 units. [20 marks]

- B7.** (a) What is a cluster sample? Under what circumstances are cluster samples advantageous?
 (b) A statistician developed a test designed to assess the attitudes of students towards mathematics at a certain school with 100 classes. He sampled 20 classes and tested every member of the sampled class. The results are shown in the table below.

Class (i)	No. of students (m_i)	Total score (y_i)	Class (i)	No. of students (m_i)	Total score (y_i)
1	30	1400	11	35	1700
2	25	1000	12	40	1800
3	35	1500	13	20	1000
4	20	1400	14	25	1000
5	28	1400	15	38	2100
6	30	1200	16	32	1200
7	32	1800	17	26	1000
8	34	1400	18	28	1500
9	36	1600	19	25	1000
10	32	1200	20	40	2000

- (i) Use the data to estimate the average score at the school,
 (ii) Estimate the variance of the average score,
 (iii) Hence find a 95% confidence interval for the average score. [20 marks]

- B8.** (a) Suppose y_1, \dots, y_N denotes a population of size N with mean μ and variance

$$\sigma^2. \text{ Show that } \text{Cov}(y_i, y_j) = \frac{-\sigma^2}{N-1}.$$

- (b) An investigator is interested in estimating the total cost of water consumed in a particular month in a certain town of $N = 2000$ houses. A simple random sample of $n = 200$ houses was selected. The sample average cost for the n houses was $\bar{y} = \$800.00$ with a sample standard deviation of $s = 50$.
 (i) Estimate, T , the total cost of water in the town,
 (ii) Place a bound on the error of estimation using 2% level of significance,
 (iii) Determine the sample size required to estimate T , the total cost of water consumed in the town, with a bound on the error of estimation of magnitude $d = 20000$ at 10% level of significance. [20 marks]

- B9.** (a) Define stratified random sampling. Under what circumstances are stratified random samples advantageous?
 (b) A certain university wishes to estimate the average score obtained on a statistics

examination for first year students doing a Bachelor of Commerce (BCom) degree programme. The 200 students are stratified into three groups depending on a specific programme they are studying [Accounts, Banking and Marketing]. A random sample of 50 is proportionally allocated and simple random samples of sizes 14, 20 and 16 respectively are selected from the three groups. The results are summarized below.

		GROUP		
		I	II	III
Group Size	N_i	55	80	65
Sample Size	n_i	14	20	16
Scores	$\sum y_i$	1116	1295	599
Standard deviation	s_i	10.254	12.578	13.645

- (i) Estimate the average score for the BCom students and place a bound on the error of estimation using $\alpha = 0.05$,
- (ii) Estimate the total score for the BCom students and place a bound on the error of estimation using $\alpha = 0.05$. **[20 marks]**

END OF EXAMINATION