EASTERN UNIVERSITY, SRI LANKA
DEPARTMENT OF MATHEMATICS
THIRD EXAMINATION IN SCIENCE – 2008/2009
SECOND SEMESTER (JANUARY, 2012)

ST 302 – SAMPLING THEORY
(SPECIAL REPEAT)

Answer **all** questions Time: **Three** Hours

(01) (a) Explain the advantages and disadvantages of sampling method.

(b) Distinguish between target population and study population.

(c) Show that:

i. Sample mean is an unbiased estimator for population mean.

ii. Sample variance is an unbiased estimator for population variance.

(02) (a) Show that: $n \geq \frac{N}{1+NV}$, where, N is the population size, $V = \left(\frac{d}{Z_\alpha S}\right)^2$, d is the marginal of error, S is the population standard deviation and α is the level of significance.

(b) A survey is conducted to investigate the extent of absenteeism not connected with illness or official holidays. A random sample of 500 men out of a total workforce of 36000 were asked how many days they have taken off from work in the previous six (6) months as 'casual holidays' and the results were as follows.

Days off	1	2	3	4	5	6	7	8	9	10
No of men	157	192	90	31	18	5	2	4	0	1

(P. T. O)

The objective is to estimate the average number of casual holidays taken by workm in the industry.

i.) Find an estimate for the population variance.

ii.) What is the margin of error, if sample size 500? (Consider 5% as level significance)

iii.) How large a sample is needed to estimate the average number of casual holida taken to within 10% of the correct figure with 95% assurance?

(c) In a private library, the books are kept on 150 shelves of similar size. The numbers (C books on shelves picked at random were found to be 28, 23, 25, 33, 18, 22, 29, 29, 3 22, 26, 20, 21, 28 and 25. Compute confidence interval for total number of books the library at 5% significance level.

(03) (a) The values of two random variables X and Y are observed on a simple random samp of size n from a population of size N. Let $\bar{x}$ and $\bar{y}$ be the sample means and $\overline{X}$ a $\overline{Y}$ be the population means respectively. For sufficiency large n, prove each of th following:

i.) Ratio of the sample means r is an unbiased estimator for that ratio of the populatio means R. Where $r=\frac{\bar{y}}{\bar{x}}$ and $R=\frac{\overline{Y}}{\overline{X}}$

ii.) $V(r)=\frac{(1-f)}{n\overline{Y}^2}\sum_{i=1}^{N}\frac{(Y_i-RX_i)^2}{N-1}$

(b) The following figures give the information on weekly expenditure of food(Y), th number of persons(X_1) and weekly family income (X_2) in a simple random sample o 33 middle income families.

$\sum y = 907.2$ $\sum x_1 = 123$ $\sum x_2 = 2394$

$\sum y^2 = 28{,}224$ $\sum x_1^2 = 533$ $\sum x_2^2 = 177\,254$

$\sum y x_1 = 3595.5$ $\sum y x_2 = 66\,678$

Neglecting sample fraction, estimate each of the following:

i.) The mean weekly expenditure on food per family.

ii.) The mean weekly expenditure on food per person.

iii.) The expenditure of the income of that spent on food.

In each case, compute the standard error of the estimator.

(04) (a) Define an estimator for the population mean in a Stratified Random Sampling Scheme and derive its variance.

(b) Distinguish between Bowley's and Neyman's methods of constructing a stratified random sample. Deduce the expression of the variance in each of these cases using the results you obtained in part (a).

(c) In a survey on the area under a crop, a total of 186 villages in a district were divided into 4 strata according to the area of the villages. From each stratum, simple random samples were selected by proportional method and the areas under crop were noted. The following data is obtained from the survey.

Stratum	Size (N_i)	Sample size(n_i)	Area under the crop
1	72	8	14,12,8,11,12,10,13,16
2	53	5	27,20,21,22,30
3	35	4	36,47,52,61
4	26	3	92,105,82

Obtain an estimator of the total area under the crop in the district and estimate the variance of this estimate.

(P. T. O)

(05) (a) Define a " Linear Systematic Sample" and show that its sample mean is an unbias estimator of the population mean.

(b) Show that the variance of the estimated mean $V(\bar{y}_{sys})$ is given by,

$$V(\bar{y}_{sys}) = \left(\frac{N-1}{N}\right)S^2 + \left(\frac{(n-1)k}{N}\right)S^2_{way},$$

Where $S^2_{way} = \frac{1}{k(n-1)}\sum_{i=1}^{k}\sum_{j=1}^{n}\left(y_{ij} - \bar{y}_i\right)^2$ is the sum of squares among units which l within the same systematic sample, and $\bar{Y}$ is the population mean.

(06) A government wants to determine whether the taxpayers support increasing local taxes provide more public funding to schools. For this purpose, they randomly select 500 scho children from a list of all children in public school and then survey the parents of the children.

(a) What is the target population?

(b) What is the sampling frame?

(c) Explain why "using the wrong sampling frame" might lead to a biased estimate taxpayer support for increasing taxes.

(d) Explain a method for collecting each of the following types of samples for this surv and discuss the advantages and disadvantages:

i.) A simple random sample

ii.) A stratified random sample

iii.) A systematic sample

iv.) A cluster sample

----------END----------