2017

STATISTICS

(Major)

Paper : 5.2

(Sample Survey)

Full Marks : 60

Time : 3 hours

The figures in the margin indicate full marks for the questions

1. Answer the following questions/Fill in the blanks/Choose the correct option (no reasoning is necessary) : 1×7=7

(a) The number of possible samples of size n from a population of N units, drawn without replacement is

(i) N^n

(ii) $N!$

(iii) N_{C_n}

(iv) n^2

(b) Name the possible source of bias in the following procedure :

A basket of oranges is sampled by taking some oranges from the top.

(c) Finite population correction factor is

(i) $\frac{n}{N}$

(ii) $1-\frac{n}{N}$

(iii) $1+\frac{n}{N}$

(d) Judgement sampling is

(i) probabilistic

(ii) non-probabilistic

(iii) mixed

(e) Simple random sampling is also known as _____ random sampling.

(f) Efficiency of cluster sampling increases as the cluster size _____.

(g) In two-stage sampling, what sampling design is used to select second-stage units from the selected first-stage units?

2. Answer the following questions in brief : 2×4=8

(a) Name the test of randomness usually adopted to ascertain whether random number tables are really random or not.

(b) Explain sampling unit and sampling frame.

(c) Explain cluster sampling procedure with example.

(d) A population of eight households, say *a*, *b*, *c*, *d*, *e*, *f*, *g* and *h*, write down all possible samples of size 3 according to the technique of circular systematic sample.

3. Answer any *three* from the following questions : 5×3=15

(a) What are the advantages of a sample survey over a complete census?

(b) Prove that in stratified random sampling, the sample mean is an unbiased estimate of population mean. Also find its sampling variance.

(c) Explain the cumulative total methods and the Lahiri's method of selecting a probability proportional to size (pps) sample with replacement.

(d) Show that the regression estimate is always more efficient than the ratio estimate unless the regression equation of y on x is a straight line passing through the origin.

(e) Explain the concept of linear and circular systematic samplings. Give the condition under which a systematic sample is more precise than a simple random sample drawn without replacement.

4. Answer *either (a) or (b)* of the following questions : 10

(a) (i) What are different sources of errors in a sample survey? How can these errors be controlled? 5

Or

Define simple random sampling with and without replacement. 5

(ii) In stratified random sampling (STRS) with given cost function of the form

$$C = a + \sum_{i=1}^{n} c_i n_i$$

where, a is the overhead cost and c_i is the cost per unit in the ith stratum, prove that

$$n_i \propto \frac{N_i S_i}{\sqrt{C_i}}$$

(Symbols have their usual meanings) 5

(b) Describe the two-stage sampling procedure. For a two-stage sampling, where first-stage units are of equal size, obtain the variance of the estimator of the population mean. 4+6=10

5. Answer *either (a) or (b)* of the following : 10

(a) Show that for a simple random sampling without replacement (SRSWOR), the probability of two specified units being selected at any two given draw is $\frac{1}{N(N-1)}$.

Also prove that

$$V(\bar{y}_{st})_{\text{Ney}} \leq V(\bar{y}_{st})_{\text{prop}} \leq V(\bar{y}_n)_R$$

2+8=10

(b) Find an unbiased estimate of the population mean in systematic sampling.

If the population consists of a linear trend of the form

$$Y_i = a + bi,\ i = 1, 2, \ldots, N;\ N = nk$$

then prove that

$$V(\bar{y}_{st}) \leq V(\bar{y}_{sys}) \leq V(\bar{y}_n)_R$$

(Symbols have their usual meanings) 2+8=10

6. Answer *either (a) or (b)* of the following questions : 10

(a) What is the basic difference between simple random sample (SRS) and sampling with probability proportional to size (pps)?

Show that in a pps sampling with replacement, an unbiased estimator of the population total Y is given by

$$\hat{Y}_{pps} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i}{p_i}\right)$$

with its sampling variance

$$V(\hat{y}_{pps}) = \frac{1}{n}\sum_{i=1}^{N} p_i\left(\frac{y_i}{p_i} - Y\right)^2$$ 2+3+5=10

(b) Explain the ratio method of estimation with examples. Show that the first approximation of the expected value of the ratio estimate of the population mean is

$$E_1(\bar{Y}_R) = \bar{Y}_N \left\{1 + \frac{N-n}{N_n}\left(Cx^2 - \rho c_x c_y\right)\right\}$$

(Symbols have their usual meanings)

Find the efficiency of the ratio estimate in comparison to simple random sampling. 3+4+3=10

★ ★ ★