INDIAN STATISTCAL INSTITUTE

CHENNAI CENTRE

M.STAT I YEAR: SECOND SEMESTRAL EXAMINATIONS

Duration: 3 hours Sample Surveys Date: February 20, 2017

Note: Question number 4 is compulsory. The rest of the paper carries 59 marks and you can answer as much as you can from this part. The maximum you can score from this part is 44.

1 a) Describe how you would draw a sample of size n from a population of size N using Lahiri-Midzuno-Sen sampling scheme.

b) Verify whether this scheme yields a πPS design. If not, suggest a modification under which this scheme results in inclusion probabilities π_i's proportional to size.

c) For size measure x taking values X_i on unit U_i, I = 1,2,...,N, let $p_i = X_i / X$, where $X = \sum_1^N X_i$. Define q_i as revised probabilities proportional to $p_i(1-p_i) / (1-2p_i)$. For selecting a sample of size two, let the first unit be drawn with ppq (probability proportional to q_i) and the second unit with probability $p_j / (1 - p_i)$, $j \neq i$, given that unit i is selected first. Show that this scheme provides a non negative variance estimator due to Sen-Yates-Grundy.

3 + (4+3) + 8 = (18)

2 a) Explain how the concept of sufficiency is useful in showing that an estimator based on an unordered sample of distinct units is better than one based on an ordered sample.

b) Demonstrate the admissibility of Horvitz Thompson estimator in the class of all homogeneous linear unbiased estimators for population total.

c) Show that Mahalanobis's IPNS Technique besides controlling the nonsampling errors, also provides an estimate of the sampling error.

d) Consider the sampling design { $(s_1, p(s_1)) = ((1,2,3), 0.2)$; $(s_2, p(s_2)) = ((1,2), 0.4)$;

$((s_3,, p(s_3)) = ((1,2,4), 0.1)$; $((s_4, p(s_4)) = ((3,4), 0.3)$ }. Devise a one-by-one drawing mechanism for obtaining sample s_2 of this design.

e) Defining clearly, give an example of (i) a model -unbiased estimate and (ii) a design as well as model unbiased estimate.

6 + 6 + 4 + 4 + 4 = (24) Contd./-

3. a) Explain what you understand by Warner's Randomized Response Model for unbiasedly estimating the Proportion of respondents belonging to a sensitive character.

   b) Compare Warner's estimator with a direct and truthful response estimator for this case.

   c) What are the modifications in model and the corresponding estimator as well as its variance under Simmon's unrelated model .

   4 +5 +8 = (17)

4. A sampling statistician has selected two districts out of five, growing a certain crop in Tamil Nadu to estimate the total yield Y of the crop in these five districts. The sample districts with serial numbers 2 and 4 are selected by Probability Proportional to Size (x) and With Out Replacement (PPSWOR) design, size being the number of plots(x) in the districts. Based on the data given in the table below:

District sl. no.	1	2	3	4	5
Number of plots x	80	100	70	50	40
Yield y	?	15.7	?	7.5	?

   a) Calculate the statistician's estimate $\hat{Y}$ of the total Y using the Horvitz and Thompson method. Also compute the unbiased estimate of variance of $\hat{Y}$ suggested by Sen-Yates-Grundy and check whether it is nonnegative in this example.

   b) If the same sample is selected by PPS With Replacement design instead of PPSWOR, calculate the Horvitz Thompson estimate of Y.

   c) Using the information that the statistician has selected district number 4 first in the sample followed by district number 2 without replacement , calculate an unbiased estimate t of the total Y and an unbiased estimate v(t) of Var.(t) .

   (7 +3) + 2 + (2 + 2) = (16)