Sample Surveys Assignment for Final Semester Assessment of M1-2020 Students

Total Marks = 30 : Each question carries 6 marks.

Submit your answers at my eamil id kajaldihidar@gmail.com by writing in a file (typed or scanned hand-written) named as yourrollno(SSassign).pdf

1. The following figures relate to a group of 15 households.

Serial No.	HH Size	Expenditure last month (Rs.)
1	8	5470.35
2	6	2716.80
3	5	1873.75
4	4	1693.20
5	3	1393.55
6	6	2398.74
7	2	3153.35
8	5	2708.75
9	7	2873.60
10	6	3775.80
11	8	5027.25
12	3	1175.28
13	4	2952.15
14	2	1032.27
15	2	2075.41

Consider the above data as a population data. Estimate the average last month's h.h. expenditure in this area based on Durbin (1967)'s IPPS sampling scheme of size $n = 2$ h.h.s. with h.h. size as size measure. In addition give the estimates of standard error, CV and 95% CI of $\bar{Y}$.

Determine how the selection probability need to be modified for Lahiri-Midzuno-Sen (1952)'s scheme to make the scheme as IPPS sampling scheme for $n = 6$.

2. Explain how the double sampling approach can be used to overcome the problem in stratified random sampling when the stratum preparation are not completely known.

3. Explain how non-response situation in an SRSWR surveys can be handled by Politz and Simmon's 'at-home-probability' technique without additional attempt of recovering data.

4. Suppose y is a real valued variable relating to a stigmatizing characteristics like expenses on treatment of AIDS, gain or loss through gambling during last month, money earned or spent in dubious means, etc. Let Y_i denote the value of the i_{th} person, $i = 1, 2, \ldots, N$. The problem is to estimate the population mean $\bar{Y} = \sum_{i=1}^{N} Y_i/N$.

   Suppose n persons are selected by using a sampling scheme which has the properties that $\pi_i > 0, \forall i = 1, 2, \ldots, N$ and $\pi_{ij} > 0, \forall i \neq j, i, j \in U = \{1, 2, \ldots, N\}$.

   To gather randomized response from a sampled person, suppose an investigator approaches with a box of cards marked (i) 'True y value' with C as their proportion or (ii) marked $x_1, \ldots, x_j, \ldots, x_M$ with respective proportions $q_1, \ldots, q_j, \ldots, q_M$, such that $C + \sum_{j=1}^{M} q_j = 1$. The device thus produces the RR for i_{th} person as

   $$z_i = y_i \text{ if 'True y value' card appears}$$

   $$= x_j \text{ if } x_j \text{ card appears .}$$

   Based on the RRs gathered using such box,
   (a) obtain an unbiased estimator of $\bar{Y}$.
   (b) Also obtain an unbiased variance estimator of (a).

5. Define protection of privacy measure in estimating a sensitive population proportion based on SRSWR of respondents and by using RRT that produces either yes or no response.

   By using illustration of Warner (1965)'s RR model, show that as the privacy is protected more and more the efficiency in estimation goes on declining.