

Fixing Size of a Varying Probability Sample in a Direct and a Randomized Response Survey

Arijit Chaudhuri¹ and Dipika Patra²

¹Indian Statistical Institute, Kolkata

²Seth Anandram Jaipuria College, Kolkata

Received 06 March 2023; Accepted 27 March 2023

SUMMARY

For a Direct Survey on innocuous characteristics Chebyshev's inequality is helpful in prescribing the size of a sample in a survey. An extension of the same to cover stigmatizing features in Randomized Response (RR) survey is not smooth enough. Different situations are illustrated and solutions proposed.

AMS subject classification: 62D05.

Keywords: Randomized response survey; Sample-size fixation; Varying probability sampling.

1. INTRODUCTION

Our main concern here is to unbiasedly estimate the proportion of people in a community bearing a specific stigmatizing characteristic A , say, like criminal propensities, alcoholism, intoxicating drug habits and similar qualitative features or to estimate total or average expenses incurred because of such sensitive experiences like costs of treatment of AIDS, loss in gambling, paying fines for fraudulent conviction, income loss due to confinement in jail etc. A stigmatizing variable \mathcal{Y} will take real value \mathcal{Y}_i which may be simply 1 or 0 for a person i in a population $U = (1, 2, \dots, i, \dots, N)$ bearing a sensitive feature A or its complement A^c . The total $Y = \sum_{i=1}^N \mathcal{Y}_i$ or mean $\bar{Y} = \frac{Y}{N}$ is our estimated parameter of interest. A sample s from U of a 'suitable size n ' is to be chosen according to a design P assigning a value $p(s)$ to s . It is to be surveyed gathering directly (called a Direct Response or DR survey) or by a Randomized Response (RR) Technique (RRT). Simplest design is SRSWR (Simple Random Sampling With Replacement) with

its variant SRSWOR (Simple Random Sampling Without Replacement). Here we shall deal with more complex sampling designs, namely, PPSWR (Probability Proportional to Size With Replacement), IPPS (Inclusion Probability Proportional to Size) and RHC (Rao, Hartley and Cochran's) sampling scheme. Corresponding estimation procedures given by Hansen and Hurwitz (HH), Horvitz and Thompson (HT) and by RHC themselves will be described in Section 3 below. In Section 2, we describe a few RRTs we choose to deal with in this paper. Our main concern is of course to discuss how to prescribe sample size in respective sampling designs to be followed in DR and RR surveys.

2. A FEW ILLUSTRATIVE RR DEVICES

2.1 Warner's RR Device

Warner (1965) as the pioneer concerning RRT's prescribed essentially that an interviewer is to obtain an RR from a sampled person i of U as

$I_i = 1$ if a 'match' results in his/her feature A or A^c when he/she draws randomly from a pack of cards