A randomized response model for estimating a rare sensitive attribute in stratified sampling using poisson distribution

Tanveer A. Tarray[a,*] and Housila P. Singh[b]

[a]*Department of Computer Science and Engineering, Islamic University of Science and Technology, Kashmir, India*
[b]*School of Studies in Statistics, Vikram University, Ujjain, India*

Abstract. The crux of this paper is to estimate the mean of the number of persons possessing a rare sensitive attribute based on the Singh et al. [1] randomization device by utilizing the Poisson distribution in stratified sampling. This study also deals with the extension of the estimation reported by Singh and Tarray [2] using a Poisson distribution and an unrelated question randomized response model reported in Singh et al. [1]. In stratified sampling, the estimators are proposed when the parameter of the rare unrelated attribute is known and also when it is unknown. It is shown that the proposed models are more efficient than the model given by Lee et al. [3] in both cases, that is, when the proportion of persons possessing a rare unrelated attribute is known and that when it is unknown. When the sizes of the stratified populations are not given, other estimators are suggested using stratified double sampling. Properties of the proposed randomized response model are studied and recommendations are made.

Keywords: Randomized response technique, estimation of proportion, rare sensitive characteristics, relative efficiency

AMS Subject Classification: 62D05

1. Introduction

In surveys concerning sensitive questions such as gambling, alcoholism, sexual behavior, drug taking, tax evasion, illegal income and else, direct techniques for collecting information may induce interviewed people to refuse answering or to give untruthful or misleading responses. To reduce non respondent's rates and biased responses arising from sensitive, embarrassing, threatening, or even incriminating questions, some special statistical techniques may be employed to ensure interviewee anonymity or, at least, a higher degree of confidence. Such techniques, known as randomized response methods, use a randomization device, such as a die or a deck of cards, rather than a direct response to collect reliable information on sensitive issues. Depending on the result produced by the randomization device, the interviewee gives an answer concerning his/her true status. Since the interviewer is unaware of the result of the device, the use of these methods ensures that respondents cannot be identified on the basis of their answers. Randomized response (RR) survey techniques introduced by Warner [4] provide such an alternative to meeting the twin objectives of generating enough reliable data to yield fruitful inference and creating a feeling among respondents that their privacy is protected despite their truthful replies to cleverly designed questions. Since the work by Warner [4], a huge literature has emerged on the use and construction of different randomization devices to estimate

---

*Corresponding author: Tanveer A. Tarray, Department of Computer Science and Engineering, Islamic University of Science and Technology, Awantipora, Pulwama, Kashmir 192122, India. E-mail: tanveerstat@gmail.com.