

Adaptive strategies in unrelated question randomized response techniques for sensitive data collection

Khalid Ul Islam Rather^a , Osama Abdulaziz Alamri^b  and Tanveer Ahmad Tarry^c 

^aDepartment of Endocrinology, Sher-i-Kashmir Institute of Medical, Sciences, Srinagar, India; ^bDepartment of Statistics, University of Tabuk, Tabuk, Saudi Arabia; ^cDepartment of Mathematical Sciences, Islamic University of Science and Technology, Pulwama, India

ABSTRACT

When surveys involve sensitive issues—such as gambling habits, drug addiction, alcoholism, tax evasion, induced abortions, drunk driving, criminal history, or homosexuality—non-response and response biases frequently occur, as respondents may be unwilling to disclose truthful information. To address this issue, Warner introduced the Randomized Response Technique (RRT), which protects respondent privacy while facilitating accurate data collection. Building on this idea, the present study proposes a novel unrelated question randomized response model designed to reduce both response and non-response biases associated with stigmatized attributes. The model provides an unbiased estimator for the population proportion possessing the sensitive characteristic. Furthermore, a cost function is incorporated into the analysis to evaluate the trade-offs between efficiency and resource utilization, offering a practical decision-making tool for survey planners. Theoretical properties of the model are thoroughly investigated, and a small-scale simulation study is conducted to validate its performance. Numerical illustrations and R-based simulation codes are also provided to support the applicability of the proposed method.

ARTICLE HISTORY

Received 6 June 2023
Accepted 28 August 2025

KEYWORDS

Randomized response sampling; Stratified random sampling; Estimation of proportion; Respondents protection; Sensitive quantitative variable

AMS SUBJECT CLASSIFICATION

62D05

1. Introduction

The social survey is a key tool for gathering information on human communities. The social survey has gained a reputation for being incredibly useful for assessing opinions, attitudes, and behaviours that span a broad range of interests. The surveys are undertaken for a variety of reasons, with the most obvious and understandable being the lack of particular facts or information in the archives. If one wants to know the crime rate, for example, formal records on crimes do not contain information regarding unseen crimes or unreported victimisation experiences. Investigators may occasionally be unable to access information on specific individuals (in a population) due to legal constraints. In general, questionnaires—and particularly social surveys have a lot of items. Due of the societal stigma they carry, some of the things may be concerning sensitive or high-risk behaviour. A common challenge in studies focusing on high-risk behaviors is the potential for participants to provide inaccurate information, whether knowingly or unknowingly. This issue is often attributed to social desirability bias, which significantly distorts responses in standardized personality tests within psychological research. Survey researchers express concerns about the reliability of survey results pertaining to various sensitive topics such as drunk driving, marijuana use, tax evasion, illicit drug consumption, induced abortion, shoplifting, child abuse, family matters, exam cheating, HIV/AIDS, and sexual behaviors. A randomised response model was first used by Warner (1965) to determine the population proportion for a sensitive attribute. The unrelated question format was initially proposed by Greenberg et al. (1969), establishing a theoretical base for privacy-protected estimation. In RRT, such as models by Gupta et al. (2020),

CONTACT Khalid Ul Islam Rather  khalidstat34@gmail.com  Department of Endocrinology, Sher-i-Kashmir Institute of Medical, Sciences, Srinagar, India.

© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.