Journal of Mathematics

WILEY

RESEARCH ARTICLE OPEN ACCESS

Stratified Randomized Response Sampling for Estimating Sensitive Proportions in Populations: An Enhanced Approach to Privacy and Precision

Tanveer Ahmad Tarray[1] | Zahoor Ahmad Ganie[2] | Gazala Salam[1] | Eid Sadun Alotaibi[3]

[1]Department of Mathematical Sciences, Islamic University of Science and Technology, Kashmir, India | [2]Department of Electrical Engineering, Islamic University of Science and Technology, Kashmir, India | [3]Department of Mathematics and Statistics, Taif University, Taif, Makkah, Saudi Arabia

Correspondence: Gazala Salam (gazala.salam@iust.ac.in)

Received: 17 December 2025 | **Revised:** 24 April 2026 | **Accepted:** 4 May 2026

Academic Editor: Smritijit Sen

Keywords: estimation accuracy | privacy protection | randomized response technique (RRT) | sensitive data collection | stratified sampling | survey bias reduction

ABSTRACT

In surveys, social desirability bias and privacy concerns often hinder honest responses to sensitive questions. The randomized response technique (RRT) mitigates these issues by preserving respondent anonymity, but its conventional application under simple random sampling ignores population heterogeneity. This paper proposes a novel methodology that integrates stratified sampling with RRT to simultaneously improve estimation precision and protect privacy. By partitioning the population into homogeneous strata and applying RRT within each stratum, the proposed approach reduces response bias and yields more accurate estimates of sensitive proportions. Explicit formulas for estimating population proportions and variances are derived, and an optimal sample allocation rule is provided to minimize overall variance. A hypothetical survey on tax evasion illustrates the methodology, and a simulation study demonstrates that the proposed stratified RRT estimator achieves a relative efficiency of 2.34 (a 57% variance reduction) compared to conventional RRT under simple random sampling. The procedure is versatile and applicable in social sciences, public health, and other fields requiring confidential data collection.

MSC 2010 Classification: 62D05; 68T37; 93E24

1 | Introduction

Estimating the prevalence of sensitive or stigmatized attributes—such as tax evasion, illicit drug use, corruption, or HIV status—is a fundamental challenge in the social and behavioral sciences. Respondents often refuse to answer or provide untruthful responses due to fear of judgment, legal repercussions, or privacy violations, leading to nonresponse bias and response distortion [1]. To overcome these difficulties, Warner [1] introduced the randomized response technique (RRT), which allows respondents to answer sensitive questions indirectly while maintaining confidentiality. In its classical form, RRT is typically implemented under simple random sampling (SRS). However, SRS ignores population heterogeneity: when the prevalence of the sensitive attribute varies across subgroups defined by characteristics such as age, income, or region, SRS-based RRT yields estimates with unnecessarily high variance and fails to provide subgroup-specific insights [2]. Stratified sampling offers a natural remedy by partitioning the population into internally homogeneous strata and drawing independent samples from each. Stratified designs are known to improve precision when the stratification variable is correlated with the characteristic of