

IMPROVED ESTIMATION OF MEAN IN RANDOMIZED RESPONSE MODELS

Zawar Hussain*[†] and Javid Shabbir*

Received 25:07:2009 : Accepted 08:07:2010

Abstract

The present investigation considers the problem of estimating the mean of a sensitive quantitative variable μ_A in a human population survey, using the scrambled response technique suggested by Ryu, Kim, Heo and Park (*On stratified randomized response sampling*, Model Assisted Statistics and Application **1**(1), 31–36, 2005–2006). Specifically, using the prior estimate (or guessed mean) of the mean of a population, a family of estimators $\hat{\mu}_{Ak}$ is presented to estimate the population mean μ_A , and its properties are examined. The optimum value of the degree $k(0 \leq k \leq 1)$ of the belief in the prior estimate depends, besides others, on the unknown population parameters, e.g. mean and variance, so the proposed family of estimators may have limited practical applications. In an attempt to overcome this problem, another estimator based on the estimated optimum value of k has been proposed. The proposed estimator has been compared with the Ryu *et al.* and Hussain and Shabbir (*Improved estimation procedure for the mean of a sensitive variable using randomized response model*, Pakistan Journal of Statistics **25**(2), 205–220, 2009) estimators assuming simple random sampling with replacement.

Keywords: Sensitive question, Estimation of mean, Simple random sampling with replacement, Scrambled response, Mean squared error, Prior estimate.

2000 AMS Classification: 62D05.

*Department of Statistics, Quaid-i-Azam University 45320, Islamabad 44000, Pakistan.
E-mail: (Z. Hussain) zhlangah@yahoo.com (J. Shabbir) jsqau@yahoo.com

[†]Corresponding Author.

1. Introduction

Since the introduction of randomized response models by Warner (1965) a large number of theoretical as well as practical studies have been done. Based on these studies it has been established that deliberate reporting of falsified answers and the refusal to respond are two major sources of non-sampling errors. The bias produced in the estimators due to these non-sampling errors sometimes may seriously mislead, particularly when the survey is about sensitive attributes. The randomized response technique was further enhanced by Greenberg *et al.* [6] to the estimation of the mean of sensitive quantitative variables. Till now, there has been a rich growth in the realm of randomized response models. In fact, randomized response (RR) models are used as a tool to decrease the evasive answer bias and, of course, to provide privacy protection to the respondents in order to get them ready to divulge their response honestly. Some of the recent randomized response models allowing the scrambling of true responses are Eichhorn and Hayre [4], Gupta *et al.* [7], Singh and Mathur [19, 20], Espejo and Singh [5], Singh and Mathur [21, 22], Singh and Mathur [23, 24, 25], Gupta and Shabbir [8], Bar-Lev *et al.* [3], Ryu *et al.* [15], Singh and Mathur [26], Arnab and Dorffner [2], Singh and Mathur [27], Hussain and Shabbir [9], Hussain *et al.* [11], Singh and Mathur [28], and many others.

Using Mangat and Singh [12], Ryu *et al.* [15] suggested a RR model. A sample of size n is taken using simple random sample with replacement. The i^{th} respondent selected in the sample is requested to use the randomization device R_1 , which consists of two statements:

- (i) "Report the true response A of the sensitive question," and
- (ii) "Go to randomization device R_2 in the second stage",

represented with probabilities P and $(1 - P)$, respectively. The randomization device R_2 consists of two statements:

- (i) "Report the true response A of the sensitive question," and
- (ii) "Report the scrambled response AB of the sensitive question,"

represented by probabilities T and $(1 - T)$, respectively. Using the assumption of a known distribution of the scrambling variable B such that $\mu_B = 1$ and $\sigma_B^2 = \gamma^2$, the response Y_i of the i^{th} respondent can be written as

$$(1.1) \quad Y_i = \alpha_i A_i + (1 - \alpha_i)[\beta_i A_i + (1 - \beta_i)A_i B_i],$$

where $\alpha_i = 1$ if the i^{th} respondent is randomly assigned to the statement (i) in R_1 , and $\alpha_i = 0$ if a respondent is randomly assigned to the statement (ii) in R_1 . Further, $\beta_i = 1$ if the i^{th} respondent is randomly assigned to the statement (i) in R_2 , and $\beta_i = 0$ if the i^{th} respondent is randomly assigned to the statement (ii) in R_2 . The expected value of the observed response is given by

$$(1.2) \quad E(Y_i) = P\mu_A + (1 - P)\{T\mu_A + (1 - T)\mu_A\mu_B\} = \mu_A,$$

where α_i and β_i are Bernoulli random variables with means P and T respectively. Ryu *et al.* [15] proposed an unbiased estimator of the mean μ_A as

$$(1.3) \quad \hat{\mu}_A = \frac{1}{n} \sum_{i=1}^n Y_i.$$

The variance of $\hat{\mu}_A$ is given by

$$(1.4) \quad \text{Var}(\hat{\mu}_A) = \frac{1}{n} \{\sigma_A^2 + (\mu_A^2 + \sigma_A^2)(1 - P)(1 - T)\gamma^2\} = \frac{\mu_A^2 U_A}{n},$$

where σ_A^2 (which may or may not be known) is the population variance of the sensitive variable under study, $U_A = C_A^2 + (1 + C_A^2)(1 - P)(1 - T)\gamma^2$ and $C_A^2 = \frac{\sigma_A^2}{\mu_A^2}$.

Based on knowledge of the conditions of the experiment, nature of the study or from past experience, sometimes investigators may be able to give some prior estimate (or initial guess), say μ_0 , of the true parameter μ . While studying the shrinkage of an unbiased estimator, $\hat{\mu}$, towards the prior estimate, μ_0 , Thompson [30] suggested an estimator

$$t(\underline{y}; \mu) = \{k\hat{\mu} + (1 - k)\mu_0\}, \quad 0 \leq k \leq 1,$$

where k is the strength of the belief in the prior estimate. Specifying k closer to 1 depicts a weak belief in μ_0 . It was established by Thompson [30] that the closer the true mean μ is to the initial prior estimate, the higher the efficiency of the shrinking estimator $t(\underline{y}; \mu)$, which is also the case in our study. Based on his observations he concluded that if a point estimator is available as a prior estimate, it can be used in a better way to produce a better estimate. Some of the relevant work in this direction is found in Tse and Tse [32], Ahmed and Rohatgi [1], Tracy *et al.* [31], Singh and Shukla [29], Shirke and Nalawade [18] and Saxana [16]. Keeping in view the suggestion made by Thompson [30], we propose a family of estimators of μ_A assuming the availability of the prior estimate μ_{A0} . Also, we discuss its properties and give a comparison of the proposed estimator with the usual Ryu *et al.* [15] estimator and the Hussain and Shabbir [10] estimator.

2. Proposed estimation method

If some prior information is available about the mean of the study variable it may be used together with sample information. One of the methods using prior knowledge is the Bayesian method of estimation, where the prior knowledge is used in the form of a prior distribution. When prior information is available in the form of a point guess, it can also be used in shrinking the estimator towards the prior point estimate. Motivated by Thompson [30] and Mathur and Singh [13], we present a family of estimators to estimate the population mean μ_A of a sensitive quantitative variable as follows

$$(2.1) \quad \hat{\mu}_{Ak} = k\hat{\mu}_A + (1 - k)\mu_{A0},$$

where $0 \leq k \leq 1$ and μ_{A0} is the prior estimate of μ_A (a prior estimate may be available from past study or simply be an intelligent guess). The value of k depends upon the degree of the investigator's belief in the prior estimate μ_{A0} . The estimator given in (2.1) has a bias given by

$$(2.2) \quad \text{Bias}(\hat{\mu}_{Ak}) = d(1 - k)\mu_A,$$

where $d = \frac{(\mu_{A0} - \mu_A)}{\mu_A}$.

The mean squared error (MSE) of the estimator $\hat{\mu}_{Ak}$ is given by

$$(2.3) \quad \text{MSE}(\hat{\mu}_{Ak}) = E(\hat{\mu}_{Ak} - \mu_A)^2 = \frac{k^2\mu_A^2 U_A}{n} + d^2\mu_A^2(1 - k)^2.$$

The proposed estimator $\hat{\mu}_{Ak}$ is more efficient than the estimator $\hat{\mu}_A$ if

$$\text{MSE}(\hat{\mu}_{Ak}) - \text{Var}(\hat{\mu}_A) \leq 0.$$

From (1.4) and (2.3), we can show easily that $\text{MSE}(\hat{\mu}_{Ak}) - \text{Var}(\hat{\mu}_A) \leq 0$, when

$$(2.4) \quad \frac{U_A(1 + k)}{n(1 - k)} > d^2, \quad 0 \leq k \leq 1,$$

or

$$\frac{d^2n - U_A}{d^2n + U_A} < k \leq 1.$$

Using different values of P , T , μ_A , γ^2 , σ_A^2 and n , we have computed the ranges of values of k in which the proposed estimator is more efficient than the Ryu *et al.* [15] estimator.

The ranges of k for different values of the other parameters are given in Tables 1–3 (see Appendix).

We have computed the relative efficiency (RE) of the proposed estimator $\hat{\mu}_{Ak}$ with respect to the Ryu *et al.* [15] estimator $\hat{\mu}_A$ for different values of the parameters and selection probabilities, as $\text{RE}(\hat{\mu}_{Ak}) = \frac{\text{Var}(\hat{\mu}_A)}{\text{MSE}(\hat{\mu}_{Ak})}$. The REs are given in Tables 4–6 (see Appendix).

From Tables 1–3, we observe that:

- (1) For all values of k , our proposed estimator $\hat{\mu}_{Ak}$ is more efficient than the Ryu *et al.* [15] estimator when n , d and the true value of the population mean are small.
- (2) For fixed P , the range of k squeezes towards 1 when T increases, for all values of d and n .
- (3) For fixed T , the range of k squeezes towards 1 when P increases, for all values of d and n .
- (4) The range of values of k depends on the sample size and the true value of the population mean. As the population mean increases the range of k squeezes towards 1 for the other fixed parametric values. The same is the case when the sample size increases from moderate to large.

From Tables 4–6, we observe that for all values of k , the RE decreases as P and/or T increases, but it becomes stable when k and d are closer to one. A large reduction in the MSE of $\hat{\mu}_{Ak}$ can be gained when the sample size is smaller, by setting P and T closer to zero, or by setting k closer to 1. When the sample size is increased, then we have to set k larger if d is larger (closer to 1). This means that the greater the relative difference between the prior estimated mean and the actual mean, the larger is the weight we must attach to sample information. Otherwise, we may attach greater weight to the prior information. It seems a bit more natural to attach a heavy weight to the prior information if one expects the prior estimate of the mean to be almost accurate, and this is case with the estimator proposed above.

3. Optimum estimators amongst the family of estimators $\hat{\mu}_{Ak}$

We can also find an optimum estimator in the family of estimators $\hat{\mu}_{Ak}$ by differentiating (2.3) with respect to k and setting it equal to zero. By doing so, the optimum value of k is given by

$$(3.1) \quad k_{opt} = \frac{\mu_A^2 d^2}{d^2 \mu_A^2 + \text{Var}(\hat{\mu}_A)} = \frac{d^2}{d^2 + \frac{U_A}{n}}.$$

Thus the optimum estimator is given by

$$(3.2) \quad \hat{\mu}_{Ak_{opt}} = k_{opt} \hat{\mu}_A + (1 - k_{opt}) \mu_{A0}.$$

The MSE of $\hat{\mu}_{Ak_{opt}}$ is given by

$$(3.3) \quad \text{MSE}(\hat{\mu}_{Ak_{opt}}) = \frac{d^2 \mu_A^2 U_A}{nd^2 + U_A}.$$

Singh and Mathur [19, 27] and Hussain and Shabbir [10] have used the idea that if it is difficult to guess the value of k or the unknown population parameters, then these parameters can be replaced by their consistent estimates from the sample. Using the same idea we can obtain the estimated optimum value of k as

$$(3.4) \quad \widehat{k}_{opt} = \frac{\hat{d}^2}{\hat{d}^2 + \frac{\hat{U}_A}{n}},$$

where $\hat{d} = \frac{(\mu_{A0} - \bar{y})}{y}$, $\hat{U}_A = \hat{C}_A^2 + (1 + \hat{C}_A^2)(1 - P)(1 - T)\gamma^2$, $\hat{C}_A^2 = \frac{s_A^2}{\bar{y}^2}$.

Substituting (3.4) in (3.2), we have another estimator of μ_A as

$$(3.5) \quad \hat{\mu}_{A\hat{k}_{opt}} = \left(\frac{n\hat{d}^2}{n\hat{d}^2 + \hat{U}_A} \right) \hat{\mu}_A + \left(\frac{\hat{U}_A}{n\hat{d}^2 + \hat{U}_A} \right) \mu_{A0} = \left(\frac{n\hat{d}^2 \hat{\mu}_A + \hat{U}_A \mu_{A0}}{n\hat{d}^2 + \hat{U}_A} \right).$$

Using the Ryu *et al.* [15] model, Hussain and Shabbir [10] proposed an estimator of the population mean μ based on Searls' [17] technique. Symbolically, the Hussain and Shabbir [10] estimator is given by

$$(3.6) \quad \hat{\mu}_{A\lambda} = \lambda \hat{\mu}_A, \quad 0 < \lambda \leq 1.$$

The bias and mean squared error of their estimator are given by

$$(3.7) \quad \text{Bias}(\hat{\mu}_{A\lambda}) = (\lambda - 1) \mu_A$$

$$(3.8) \quad \text{MSE}(\hat{\mu}_{A\lambda}) = \lambda^2 \{ \mu_A^2 + \text{Var}(\hat{\mu}_A) \} + \mu_A^2 (1 - 2\lambda).$$

Hussain and Shabbir [10] showed that the optimum estimator value of $\hat{\mu}_{A\lambda_{opt}}$ is given by

$$(3.9) \quad \hat{\mu}_{A\lambda_{opt}} = \frac{\hat{\mu}_A^3}{\hat{\mu}_A^2 + n^{-1}s_Y^2},$$

where

$$(3.10) \quad \hat{\lambda}_{opt} = \frac{\hat{\mu}_A^2}{\hat{\mu}_A^2 + n^{-1}s_Y^2}.$$

4. Efficiency comparisons

The (RE) of the optimum estimator relative to the Ryu *et al.* [15] estimator is given by

$$(4.1) \quad RE = \frac{\text{Var}(\hat{\mu}_A)}{\text{MSE}(\hat{\mu}_{A\hat{k}_{opt}})},$$

$$RE = 1 + \frac{U_A}{nd^2}.$$

From (4.1), it is obvious that the proposed optimum estimator is always more efficient than the Ryu *et al.* [15] estimator in terms of the variability, but it has a limitation because the optimum value of k depends on the unknown population mean and population variance. So its practicability is limited.

We now compare our estimator with the one given in (3.6). On comparing (2.3) and (3.8), we see that $\hat{\mu}_{Ak}$ will be more efficient than $\hat{\mu}_{A\lambda}$ if $\hat{\mu}_{A\lambda} - \hat{\mu}_{Ak} \geq 0$. That is, if

$$[\lambda^2 \{ \mu_A^2 + \text{Var}(\hat{\mu}_A) \} + \mu_A^2 (1 - 2\lambda)] - \left[\frac{k^2 \mu_A^2 U_A}{n} + d^2 \mu_A^2 (1 - k)^2 \right] \geq 0,$$

or if

$$d^2 \leq \frac{2\lambda^2 (1 - \lambda)}{(1 - k)^2} - \frac{U_A (k^2 - \lambda^2)}{n(1 - k)^2}.$$

In Tables 7–9, we have calculated the ranges of k in which the proposed estimator will be more precise than the Hussain and Shabbir [10] estimator for different values of the other parameters. From Tables 7–9, it is observed that when λ and n are smaller, we have wider range of k in which proposed estimator is more precise, and this range shrinks towards 1 as d increases. For fixed λ and d , the ranges of values of k also shrinks towards 1. Tables 10–12 contain the RE of the proposed estimator relative to the Hussain and Shabbir [10] estimator. For fixed λ and k , the RE of the proposed model decreases as d increases, which is expected too. Greater efficiency is achieved when n as well as λ is

smaller, which suggests that when the population can be divided into strata, the proposed estimator should be preferably used.

Further, if we compare their optimum MSEs, then the efficiency condition reduces to $d^2 \leq 1$. This shows that the RE of the proposed estimator is dependent on the closeness of the prior estimate μ_{A0} to the true mean μ_A . The closer the prior estimate is to the true mean, the higher the RE of the proposed estimator.

Derivation of the expressions for the bias and MSE of the estimator $\hat{\mu}_{A\hat{k}_{opt}}$ is much too difficult, even to the first order of approximation, and to the first order of approximation the bias and MSER of $\hat{\mu}_{A\hat{\lambda}_{opt}}$ are derived by Hussain and Shabbir [10]. To compare them, we have worked out the simulated RE of $\hat{\mu}_{A\hat{k}_{opt}}$ compared to $\hat{\mu}_{A\hat{\lambda}_{opt}}$, and the results are given in Table 13.

5. Conclusions

When the investigators have some prior knowledge about the true mean of a study variable, then using Thompson's [30] proposal, a shrinkage estimator may be used to give a better estimate. What we have observed is that in the situations where it is impossible/difficult to study large samples, and it is suspected that the actual mean of the study variable may be small, the proposed approach to estimation may be fruitfully used to get precise estimates. To achieve a maximum gain in efficiency, the selection probabilities P and T should preferably be chosen small. The mean and variance of the scrambling variable should also be smaller. Compared to the Hussain and Shabbir [10] estimator, the proposed estimator is more efficient over a wide range of design parameters. Since the Ryu *et al.* [15] estimator is more efficient than the Greenberg *et al.* [6], Eichhorn and Hayre [4], and Gupta *et al.* [7] estimators, in sensitive surveys we recommend the application of the proposed estimator when an initial guess can be obtained (as a point estimate) from previous experience, and we infer that choosing k such that $\hat{\mu}_{Ak}$ is better than $\hat{\mu}_{A\lambda}$ and $\hat{\mu}_A$ has a great scope, even when μ_{A0} is far away from the true population mean μ_A . Further, using the sample values to estimate the optimum values of k and λ results in a superiority of the proposed estimators.

6. Appendix

6.1. Tables 1–13.

Table 1. Range of values of k for different values of selection probabilities when $\mu_S = 1, \sigma_S^2 = 0.5, C_y^2 = 2.0, n = 15$

$d = 0.15$					
P	T				
	0.1	0.3	0.5	0.7	0.9
0.1	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.3	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.5	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.7	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.9	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
$d = 0.35$					
0.1	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.3	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.5	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.7	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.9	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
$d = 0.50$					
0.1	0.08 ⁻¹	0.12 ⁻¹	0.16 ⁻¹	0.21 ⁻¹	0.25 ⁻¹
0.3	0.12 ⁻¹	0.15 ⁻¹	0.19 ⁻¹	0.22 ⁻¹	0.26 ⁻¹
0.5	0.16 ⁻¹	0.19 ⁻¹	0.21 ⁻¹	0.24 ⁻¹	0.26 ⁻¹
0.7	0.20 ⁻¹	0.22 ⁻¹	0.24 ⁻¹	0.25 ⁻¹	0.27 ⁻¹
0.9	0.25 ⁻¹	0.26 ⁻¹	0.26 ⁻¹	0.27 ⁻¹	0.27 ⁻¹
$d = 1.0$					
0.1	0.55 ⁻¹	0.57 ⁻¹	0.59 ⁻¹	0.60 ⁻¹	0.63 ⁻¹
0.3	0.57 ⁻¹	0.58 ⁻¹	0.60 ⁻¹	0.62 ⁻¹	0.64 ⁻¹
0.5	0.59 ⁻¹	0.60 ⁻¹	0.61 ⁻¹	0.62 ⁻¹	0.64 ⁻¹
0.7	0.61 ⁻¹	0.62 ⁻¹	0.62 ⁻¹	0.63 ⁻¹	0.64 ⁻¹
0.9	0.63 ⁻¹	0.64 ⁻¹	0.64 ⁻¹	0.64 ⁻¹	0.64 ⁻¹

Table 2. Range of values of k for different values of selection probabilities when $\mu_S = 1, \sigma_S^2 = 0.5, C_y^2 = 2.0, n = 30$

$d = 0.15$					
P	T				
	0.1	0.3	0.5	0.7	0.9
0.1	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.3	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.5	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.7	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.9	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
$d = 0.35$					
0.1	0.07 ⁻¹	0.11 ⁻¹	0.15 ⁻¹	0.20 ⁻¹	0.24 ⁻¹
0.3	0.11 ⁻¹	0.14 ⁻¹	0.18 ⁻¹	0.21 ⁻¹	0.25 ⁻¹
0.5	0.15 ⁻¹	0.18 ⁻¹	0.20 ⁻¹	0.23 ⁻¹	0.25 ⁻¹
0.7	0.20 ⁻¹	0.21 ⁻¹	0.23 ⁻¹	0.24 ⁻¹	0.26 ⁻¹
0.9	0.24 ⁻¹	0.25 ⁻¹	0.25 ⁻¹	0.26 ⁻¹	0.27 ⁻¹
$d = 0.50$					
0.1	0.35 ⁻¹	0.38 ⁻¹	0.41 ⁻¹	0.44 ⁻¹	0.47 ⁻¹
0.3	0.38 ⁻¹	0.40 ⁻¹	0.43 ⁻¹	0.45 ⁻¹	0.48 ⁻¹
0.5	0.41 ⁻¹	0.43 ⁻¹	0.44 ⁻¹	0.46 ⁻¹	0.48 ⁻¹
0.7	0.44 ⁻¹	0.45 ⁻¹	0.46 ⁻¹	0.47 ⁻¹	0.48 ⁻¹
0.9	0.47 ⁻¹	0.48 ⁻¹	0.48 ⁻¹	0.48 ⁻¹	0.49 ⁻¹
$d = 1.0$					
0.1	0.68 ⁻¹	0.69 ⁻¹	0.71 ⁻¹	0.72 ⁻¹	0.74 ⁻¹
0.3	0.69 ⁻¹	0.71 ⁻¹	0.72 ⁻¹	0.73 ⁻¹	0.74 ⁻¹
0.5	0.71 ⁻¹	0.72 ⁻¹	0.73 ⁻¹	0.74 ⁻¹	0.74 ⁻¹
0.7	0.72 ⁻¹	0.73 ⁻¹	0.74 ⁻¹	0.74 ⁻¹	0.75 ⁻¹
0.9	0.74 ⁻¹	0.74 ⁻¹	0.74 ⁻¹	0.75 ⁻¹	0.75 ⁻¹

Table 3. Range of values of k for different values of selection probabilities when $\mu_S = 1, \sigma_S^2 = 0.5, C_y^2 = 2.0, n = 50$

$d = 0.15$					
P	T				
	0.1	0.3	0.5	0.7	0.9
0.1	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.3	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.5	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.7	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.9	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
$d = 0.35$					
0.1	0.28 ⁻¹	0.31 ⁻¹	0.34 ⁻¹	0.38 ⁻¹	0.42 ⁻¹
0.3	0.31 ⁻¹	0.34 ⁻¹	0.36 ⁻¹	0.39 ⁻¹	0.41 ⁻¹
0.5	0.34 ⁻¹	0.36 ⁻¹	0.38 ⁻¹	0.40 ⁻¹	0.42 ⁻¹
0.7	0.38 ⁻¹	0.39 ⁻¹	0.40 ⁻¹	0.42 ⁻¹	0.43 ⁻¹
0.9	0.42 ⁻¹	0.42 ⁻¹	0.42 ⁻¹	0.43 ⁻¹	0.43 ⁻¹
$d = 0.50$					
0.1	0.50 ⁻¹	0.52 ⁻¹	0.54 ⁻¹	0.57 ⁻¹	0.59 ⁻¹
0.3	0.52 ⁻¹	0.54 ⁻¹	0.56 ⁻¹	0.58 ⁻¹	0.60 ⁻¹
0.5	0.54 ⁻¹	0.56 ⁻¹	0.57 ⁻¹	0.58 ⁻¹	0.60 ⁻¹
0.7	0.57 ⁻¹	0.58 ⁻¹	0.58 ⁻¹	0.59 ⁻¹	0.60 ⁻¹
0.9	0.60 ⁻¹	0.60 ⁻¹	0.60 ⁻¹	0.60 ⁻¹	0.60 ⁻¹
$d = 1.0$					
0.1	0.75 ⁻¹	0.76 ⁻¹	0.78 ⁻¹	0.79 ⁻¹	0.80 ⁻¹
0.3	0.76 ⁻¹	0.77 ⁻¹	0.78 ⁻¹	0.79 ⁻¹	0.80 ⁻¹
0.5	0.78 ⁻¹	0.78 ⁻¹	0.79 ⁻¹	0.80 ⁻¹	0.80 ⁻¹
0.7	0.79 ⁻¹	0.79 ⁻¹	0.80 ⁻¹	0.80 ⁻¹	0.80 ⁻¹
0.9	0.80 ⁻¹	0.80 ⁻¹	0.80 ⁻¹	0.80 ⁻¹	0.81 ⁻¹

Table 4. RE of the $\hat{\mu}_{Ak}$ relative $\hat{\mu}_A$ when $\mu_S = 1, \sigma_S^2 = 0.5, C_y^2 = 2.0, n = 15$

$k = 0.15, d = 0.15$					
T					
P	0.1	0.3	0.5	0.7	0.9
0.1	12.928976	11.862504	10.792546	9.719086	8.642106
0.3	11.862504	11.030617	10.196613	9.360485	8.522224
0.5	10.792546	10.196613	9.599596	9.001492	8.402298
0.7	9.719086	9.360485	9.001492	8.642106	8.282328
0.9	8.642106	8.522224	8.402298	8.282328	8.162315
$k = 0.15, d = 0.35$					
0.1	2.412910	2.210942	2.008852	1.806639	1.604303
0.3	2.210942	2.053772	1.896527	1.739207	1.581813
0.5	2.008852	1.896527	1.784163	1.671762	1.559322
0.7	1.806639	1.739207	1.671762	1.604303	1.536830
0.9	1.604303	1.581813	1.559322	1.536830	1.514336
$k = 0.65, d = 0.50$					
0.1	5.846190	5.430315	5.002610	4.562565	4.109635
0.3	5.430315	5.098704	4.759698	4.413047	4.058490
0.5	5.002610	4.759698	4.512885	4.262078	4.007178
0.7	4.562565	4.413047	4.262078	4.109635	3.955697
0.9	4.109635	4.058490	4.007178	3.955697	3.904049
$k = 0.85, d = 1.0$					
0.1	6.529830	6.143726	5.736086	5.305059	4.848572
0.3	6.143726	5.828626	5.499642	5.155835	4.796181
0.5	5.736086	5.499642	5.255636	5.003701	4.743444
0.7	5.305059	5.155835	5.003701	4.848572	4.690358
0.9	4.848572	4.796181	4.743444	4.690358	4.636919

Table 5. RE of the $\hat{\mu}_{Ak}$ relative $\hat{\mu}_A$ when $\mu_S = 1, \sigma_S^2 = 0.5, C_y^2 = 2.0, n = 30$

$k = 0.15, d = 0.15$					
T					
P	0.1	0.3	0.5	0.7	0.9
0.1	6.559902	6.011477	5.462598	4.913265	4.363477
0.3	6.011477	5.584610	5.157469	4.730052	4.302361
0.5	5.462598	5.157469	4.852199	4.546790	4.241240
0.7	4.913265	4.730052	4.546790	4.363477	4.180113
0.9	4.363477	4.302361	4.241240	4.180113	4.118980
$k = 0.30, d = 0.35$					
0.1	1.775856	1.627445	1.478901	1.330225	1.181415
0.3	1.627445	1.511923	1.396320	1.280636	1.164872
0.5	1.478901	1.396320	1.313697	1.231033	1.148328
0.7	1.330225	1.280636	1.231033	1.181415	1.131782
0.9	1.181415	1.164872	1.148328	1.131782	1.115234
$k = 0.65, d = 0.50$					
0.1	3.334966	3.066988	2.796880	2.524615	2.250168
0.3	3.066988	2.857089	2.645890	2.433377	2.219538
0.5	2.796880	2.645890	2.494229	2.341895	2.188881
0.7	2.524615	2.433377	2.341895	2.250168	2.158197
0.9	2.250168	2.219538	2.188881	2.158197	2.127485
$k = 0.85, d = 1.0$					
0.1	4.272835	3.948116	3.617686	3.281393	2.939079
0.3	3.948116	3.691616	3.431590	3.167965	2.900666
0.5	3.617686	3.431590	3.243659	3.053863	2.862176
0.7	3.281393	3.167965	3.053863	2.939079	2.823609
0.9	2.939079	2.900666	2.862176	2.823609	2.784965

Table 6. RE of the $\hat{\mu}_{Ak}$ relative to $\hat{\mu}_A$ when $\mu_S = 1, \sigma_S^2 = 0.5, C_y^2 = 2.0, n = 50$

$k = 0.15, d = 0.15$					
P	T				
	0.1	0.3	0.5	0.7	0.9
0.1	44.79634	43.39322	41.81893	40.04012	38.01418
0.3	43.39322	42.18514	40.85843	39.39470	37.77157
0.5	41.81893	40.85843	39.82811	38.72006	37.52515
0.7	40.04012	39.39470	38.72006	38.01418	37.27484
0.9	38.01418	37.77157	37.52515	37.27484	37.02054
$k = 0.65, d = 0.35$					
0.1	4.135159	3.799031	3.460918	3.120804	2.778670
0.3	3.799031	3.536227	3.272214	3.006985	2.740529
0.5	3.460918	3.272214	3.082889	2.892940	2.702364
0.7	3.120804	3.006985	2.892940	2.778670	2.664173
0.9	2.778670	2.740529	2.702364	2.664173	2.625957
$k = 0.65, d = 0.50$					
0.1	3.933113	3.616242	3.297000	2.975357	2.651289
0.3	3.616242	3.368149	3.118607	2.867606	2.615130
0.5	3.297000	3.118607	2.939470	2.759583	2.578941
0.7	2.975357	2.867606	2.759583	2.651289	2.542722
0.9	2.651289	2.615130	2.578941	2.542722	2.506472
$k = 0.85, d = 1.0$					
0.1	5.823400	5.376943	4.923298	4.462289	3.993737
0.3	5.376943	5.024738	4.668103	4.306953	3.941202
0.5	4.923298	4.668103	4.410604	4.150771	3.888571
0.7	4.462289	4.306953	4.150771	3.993737	3.835845
0.9	3.993737	3.941202	3.888571	3.835845	3.783022

Table 7. Range of values of k for different values of the selection probabilities when $\mu_S = 1, \sigma_S^2 = 0.5, C_y^2 = 2.0, n = 15, \text{ and } \lambda = 0.1$

P	$d = 0.15$				
	0.1	0.3	0.5	0.7	0.9
0.1	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.3	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.5	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.7	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.9	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
$d = 0.35$					
0.1	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.3	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.5	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.7	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.9	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
$d = 0.50$					
0.1	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.3	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.5	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.7	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.9	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
$d = 1.0$					
0.1	0.09 ⁻¹	0.10 ⁻¹	0.10 ⁻¹	0.10 ⁻¹	0.10 ⁻¹
0.3	0.10 ⁻¹	0.10 ⁻¹	0.10 ⁻¹	0.10 ⁻¹	0.10 ⁻¹
0.5	0.10 ⁻¹	0.10 ⁻¹	0.10 ⁻¹	0.10 ⁻¹	0.10 ⁻¹
0.7	0.10 ⁻¹	0.10 ⁻¹	0.10 ⁻¹	0.10 ⁻¹	0.10 ⁻¹
0.9	0.10 ⁻¹	0.10 ⁻¹	0.10 ⁻¹	0.10 ⁻¹	0.10 ⁻¹

Table 8. Ranges of values of k for different values of the selection probabilities when $\mu_S = 1, \sigma_S^2 = 0.5, C_y^2 = 2.0, n = 15, \text{ and } \lambda = 0.5$

P	$d = 0.15$				
	0.1	0.3	0.5	0.7	0.9
0.1	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.3	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.5	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.7	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.9	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
$d = 0.35$					
0.1	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.3	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.5	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.7	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.9	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
$d = 0.50$					
0.1	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.3	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.5	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.7	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.9	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
$d = 1.0$					
0.1	0.46 ⁻¹	0.46 ⁻¹	0.46 ⁻¹	0.47 ⁻¹	0.47 ⁻¹
0.3	0.46 ⁻¹	0.46 ⁻¹	0.47 ⁻¹	0.47 ⁻¹	0.47 ⁻¹
0.5	0.46 ⁻¹	0.47 ⁻¹	0.47 ⁻¹	0.47 ⁻¹	0.47 ⁻¹
0.7	0.47 ⁻¹	0.47 ⁻¹	0.47 ⁻¹	0.47 ⁻¹	0.47 ⁻¹
0.9	0.47 ⁻¹	0.47 ⁻¹	0.47 ⁻¹	0.47 ⁻¹	0.47 ⁻¹

Table 9. Ranges of values of k for different values of the selection probabilities when $\mu_S = 1, \sigma_S^2 = 0.5, C_y^2 = 2.0, n = 15,$ and $\lambda = 0.9$

$d = 0.15$					
P	T				
	0.1	0.3	0.5	0.7	0.9
0.1	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.3	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.5	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.7	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.9	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
$d = 0.35$					
0.1	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.3	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹
0.5	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0.01 ⁻¹
0.7	0 ⁻¹	0 ⁻¹	0 ⁻¹	0 ⁻¹	0.01 ⁻¹
0.9	0 ⁻¹	0 ⁻¹	0.01 ⁻¹	0.01 ⁻¹	0.02 ⁻¹
$d = 0.50$					
0.1	0.15 ⁻¹	0.18 ⁻¹	0.22 ⁻¹	0.26 ⁻¹	0.30 ⁻¹
0.3	0.18 ⁻¹	0.21 ⁻¹	0.24 ⁻¹	0.27 ⁻¹	0.30 ⁻¹
0.5	0.22 ⁻¹	0.24 ⁻¹	0.26 ⁻¹	0.29 ⁻¹	0.31 ⁻¹
0.7	0.26 ⁻¹	0.27 ⁻¹	0.29 ⁻¹	0.30 ⁻¹	0.31 ⁻¹
0.9	0.30 ⁻¹	0.30 ⁻¹	0.31 ⁻¹	0.31 ⁻¹	0.32 ⁻¹
$d = 1.0$					
0.1	0.58 ⁻¹	0.60 ⁻¹	0.62 ⁻¹	0.64 ⁻¹	0.66 ⁻¹
0.3	0.60 ⁻¹	0.61 ⁻¹	0.63 ⁻¹	0.64 ⁻¹	0.66 ⁻¹
0.5	0.62 ⁻¹	0.63 ⁻¹	0.64 ⁻¹	0.65 ⁻¹	0.66 ⁻¹
0.7	0.64 ⁻¹	0.64 ⁻¹	0.65 ⁻¹	0.66 ⁻¹	0.66 ⁻¹
0.9	0.66 ⁻¹	0.66 ⁻¹	0.66 ⁻¹	0.66 ⁻¹	0.67 ⁻¹

Table 10. RE of the $\hat{\mu}_{Ak}$ relative to $\hat{\mu}_{A\lambda}$ when $\mu_S = 1, \sigma_S^2 = 0.5, C_y^2 = 2.0, n = 15$ and $\lambda = 0.1$

$k = 0.15, d = 0.15$					
P	T				
	0.1	0.3	0.5	0.7	0.9
0.1	48.98996	49.05901	49.12827	49.19777	49.26750
0.3	49.05901	49.05901	49.16686	49.22099	49.27526
0.5	49.12827	49.16686	49.20551	49.24423	49.28302
0.7	49.19777	49.22099	49.24423	49.26750	49.29079
0.9	49.26750	49.27526	49.28302	49.29079	49.29856
$k = 0.25, d = 0.35$					
0.1	11.63541	11.64534	11.65530	11.66527	11.67527
0.3	11.64534	11.65308	11.66084	11.66860	11.67638
0.5	11.65530	11.66084	11.66638	11.67194	11.67750
0.7	11.66527	11.66860	11.67194	11.67527	11.67861
0.9	11.67527	11.67638	11.67750	11.67861	11.67972
$k = 0.35, d = 0.50$					
0.1	7.563589	7.572280	7.580994	7.589732	7.598494
0.3	7.572280	7.579055	7.585845	7.592650	7.599469
0.5	7.580994	7.585845	7.590704	7.595571	7.600444
0.7	7.589732	7.592650	7.595571	7.598494	7.601420
0.9	7.598494	7.599469	7.600444	7.601420	7.602396
$k = 0.50, d = 1.0$					
0.1	3.202809	3.205892	3.208982	3.212080	3.215185
0.3	3.205892	3.208295	3.210702	3.213114	3.215530
0.5	3.208982	3.210702	3.212424	3.214149	3.215876
0.7	3.212080	3.213114	3.214149	3.215185	3.216221
0.9	3.215185	3.215530	3.215876	3.216221	3.216567

Table 11. RE of the $\hat{\mu}_{Ak}$ relative to $\hat{\mu}_{A\lambda}$ when $\mu_S = 1, \sigma_S^2 = 0.5, C_y^2 = 2.0, n = 15$ and $\lambda = 0.5$

$k = 0.15, d = 0.15$					
T					
P	0.1	0.3	0.5	0.7	0.9
0.1	18.31270	18.07068	17.82787	17.58427	17.33987
0.3	18.07068	17.88190	17.69264	17.50289	17.31267
0.5	17.82787	17.69264	17.55716	17.42143	17.28545
0.7	17.58427	17.50289	17.42143	17.33987	17.25823
0.9	17.33987	17.31267	17.28545	17.25823	17.23099
$k = 0.45, d = 0.35$					
0.1	7.599130	7.532305	7.464656	7.396170	7.326830
0.3	7.532305	7.479761	7.426713	7.373152	7.319072
0.5	7.464656	7.426713	7.388508	7.350039	7.311303
0.7	7.396170	7.373152	7.350039	7.326830	7.303524
0.9	7.326830	7.319072	7.311303	7.303524	7.295733
$k = 0.55, d = 0.50$					
0.1	5.524982	5.479283	5.432973	5.386038	5.338466
0.3	5.479283	5.443318	5.406976	5.370252	5.333141
0.5	5.432973	5.406976	5.380784	5.354395	5.327807
0.7	5.386038	5.370252	5.354395	5.338466	5.322465
0.9	5.338466	5.333141	5.327807	5.322465	5.317115
$k = 0.65, d = 1.0$					
0.1	2.361835	2.336040	2.310040	2.283832	2.257414
0.3	2.336040	2.315835	2.295506	2.275050	2.254466
0.5	2.310040	2.295506	2.280907	2.266244	2.251515
0.7	2.283832	2.275050	2.266244	2.257414	2.248561
0.9	2.257414	2.254466	2.251515	2.248561	2.245605

Table 12. RE of the $\hat{\mu}_{Ak}$ relative to $\hat{\mu}_{A\lambda}$ when $\mu_S = 1, \sigma_S^2 = 0.5, C_y^2 = 2.0, n = 15$ and $\lambda = 0.9$

$k = 0.15, d = 0.15$					
T					
P	0.1	0.3	0.5	0.7	0.9
0.1	11.075689	10.212830	9.347152	8.478640	7.607280
0.3	10.212830	9.539769	8.864996	8.188504	7.510286
0.5	9.347152	8.864996	8.381963	7.898051	7.413256
0.7	8.478640	8.188504	7.898051	7.607280	7.316192
0.9	7.607280	7.510286	7.413256	7.316192	7.219091
$k = 0.35, d = 0.35$					
0.1	3.431537	3.167750	2.902506	2.635792	2.367597
0.3	3.167750	2.961576	2.754514	2.546559	2.337705
0.5	2.902506	2.754514	2.606066	2.457161	2.307795
0.7	2.635792	2.546559	2.457161	2.367597	2.277867
0.9	2.367597	2.337705	2.307795	2.277867	2.247920
$k = 0.55, d = 0.50$					
0.1	3.341560	3.096673	2.848507	2.596996	2.342071
0.3	3.096673	2.903942	2.709196	2.512403	2.313532
0.5	2.848507	2.709196	2.568840	2.427429	2.284951
0.7	2.596996	2.512403	2.427429	2.342071	2.256326
0.9	2.342071	2.313532	2.284951	2.256326	2.227657
$k = 0.75, d = 1.0$					
0.1	2.603013	2.419467	2.232340	2.041525	1.846914
0.3	2.419467	2.274238	2.126794	1.977082	1.825051
0.5	2.310040	2.232340	2.020091	1.912213	1.803139
0.7	2.041525	1.977082	1.912213	1.846914	1.781180
0.9	1.846914	1.825051	1.803139	1.781180	1.759171

Table 13. Simulated RE of $\hat{\mu}_{A \hat{k}_{opt}}$ compared to $\hat{\mu}_{A \hat{\lambda}_{opt}}$ for $\mu_S = 1, \sigma_S^2 = 0.5, C_y^2 = 2.0, n = 15$

$\mu_{A0} = 1.1$					
T					
P	0.1	0.3	0.5	0.7	0.9
0.1	1.703554	2.019900	2.107988	2.044867	1.987658
0.3	2.205389	2.116791	2.020169	2.154635	2.074227
0.5	2.017404	1.783298	2.178234	1.940251	2.155637
0.7	2.015288	2.271955	1.995428	1.634276	1.803476
0.9	1.913121	1.972407	2.068359	2.009539	2.055349
$\mu_{A0} = 1.2$					
0.1	3.431537	3.167750	2.902506	2.635792	2.367597
0.3	3.167750	2.961576	2.754514	2.546559	2.337705
0.5	2.902506	2.754514	2.606066	2.457161	2.307795
0.7	2.635792	2.546559	2.457161	2.367597	2.277867
0.9	2.367597	2.337705	2.307795	2.277867	2.247920
$k = 0.55, d = 0.50$					
0.1	3.341560	3.096673	2.848507	2.596996	2.342071
0.3	3.096673	2.903942	2.709196	2.512403	2.313532
0.5	2.848507	2.709196	2.568840	2.427429	2.284951
0.7	2.596996	2.512403	2.427429	2.342071	2.256326
0.9	2.342071	2.313532	2.284951	2.256326	2.227657
$k = 0.75, d = 1.0$					
0.1	2.603013	2.419467	2.232340	2.041525	1.846914
0.3	2.419467	2.274238	2.126794	1.977082	1.825051
0.5	2.310040	2.232340	2.020091	1.912213	1.803139
0.7	2.041525	1.977082	1.912213	1.846914	1.781180
0.9	1.846914	1.825051	1.803139	1.781180	1.759171

6.2. Derivation of Equation (1.4). Applying variance on (1.3), we get

$$\text{Var}(\hat{\mu}_A) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^n Y_i\right) = \frac{1}{n} \text{Var}(Y_i). \quad (\text{A.1})$$

We know that

$$\text{Var}(Y_i) = \text{E}(Y_i^2) - \{\text{E}(Y_i)\}^2. \quad (\text{A.2})$$

Now

$$\begin{aligned} \text{E}(Y_i^2) &= \text{E}(\alpha_i^2)\text{E}(A_i^2) + \text{E}(1 + \alpha_i^2 - 2\alpha_i)\{\text{E}(\beta_i^2)\text{E}(A_i^2) + \text{E}(1 - \beta_i)^2\text{E}(A_i^2 B_i^2) \\ &\quad + 2\text{E}(\beta_i - \beta_i^2)\text{E}(A_i^2)\text{E}(B_i)\} \\ &\quad + 2\text{E}\{\alpha_i(1 - \alpha_i)\}\text{E}(A_i^2)\text{E}\{\beta_i + (1 - \beta_i)B_i\} \\ &= P(\mu_A^2 + \delta_A^2) + (1 - P)\{T(\mu_A^2 + \delta_A^2) \\ &\quad + (1 - T)(\mu_A^2 + \delta_A^2)(\mu_B^2 + \delta_B^2) + 2(T - T)(\mu_A^2 + \delta_A^2)\} \\ &= P(\mu_A^2 + \delta_A^2) + (1 - P)\{T(\mu_A^2 + \delta_A^2) + (1 - T)(\mu_A^2 + \delta_A^2)(1 + \gamma^2)\} \end{aligned}$$

Using the above equation in (A.2), we get

$$\begin{aligned} \text{Var}(Y_i) &= P(\mu_A^2 + \delta_A^2) + (1 - P)\{T(\mu_A^2 + \delta_A^2) + (1 - T)(\mu_A^2 + \delta_A^2)(1 + \gamma^2)\} - \mu_A^2 \\ &= (\mu_A^2 + \delta_A^2)\{P + (1 - P)T + (1 - P)(1 - T)(1 + \gamma^2)\} - \mu_A^2 \\ &= (\mu_A^2 + \delta_A^2)\{1 + (1 - P)(1 - T)\gamma^2\} - \mu_A^2 \\ &= (\mu_A^2 + \delta_A^2) + (\mu_A^2 + \delta_A^2)(1 - P)(1 - T)\gamma^2 - \mu_A^2, \end{aligned}$$

so

$$\text{Var}(Y_i) = \delta_A^2 + (\mu_A^2 + \delta_A^2)(1 - P)(1 - T)\gamma^2. \quad (\text{A.3})$$

Substituting (A.3) in (A.1), we get

$$\begin{aligned} \text{Var}(\hat{\mu}_A) &= \frac{1}{n} \{ \delta_A^2 + (\mu_A^2 + \delta_A^2)(1 - P)(1 - T)\gamma^2 \} \\ &= \frac{1}{n} \mu_A^2 \left\{ \frac{\delta_A^2}{\mu_A^2} + \left(1 + \frac{\delta_A^2}{\mu_A^2}\right)(1 - P)(1 - T)\gamma^2 \right\} \\ &= \frac{\mu_A^2 U_A}{n}. \end{aligned}$$

Hence we have Equation (1.4).

6.3. Derivation of Equation (2.3). Applying variance on (2.1), we get

$$\text{Var}(\hat{\mu}_{Ak}) = k^2 \text{Var}(\hat{\mu}_A). \quad (\text{A.4})$$

As we know that

$$\text{MSE}(\hat{\mu}_{Ak}) = \text{Var}(\hat{\mu}_{Ak}) + \{\text{Bias}(\hat{\mu}_{Ak})\}^2, \quad (\text{A.5})$$

then using (2.2) and (A.4) in (A.5), we get

$$\text{MSE}(\hat{\mu}_{Ak}) = \frac{k^2 \mu_A^2 U_A}{n} + d^2 \mu_A^2 (1 - k)^2.$$

Hence we have Equation (2.3).

References

- [1] Ahmed, S. E. and Rohatgi, V. K. *Shrinkage estimation of proportion in randomized response*, *Metrika* **43**, 17–30, 1996.
- [2] Arnab, R. and Dorffner, G. *Randomized response technique for complex survey designs*, *Statistical Papers* **48**, 131–141, 2006.
- [3] Bar-Lev, S. K., Bobovitch, E. and Boukai, B. *A note on randomized response models*, *Metrika* **60**, 255–260, 2004.
- [4] Eichhorn, B. H. and Hayre, L. S. *Scrambled randomized response methods for obtaining sensitive quantitative data*, *Journal of Statistical Planning and inference* **7**, 307–316, 1983.
- [5] Espejo R, M. and Singh, H. P. *Protection of privacy with objective prior distribution in randomized response*, *Statistica Anno LXIII* (4), 697–701, 2003.
- [6] Greenberg, B. G., Kuebler, R. R. Jr., Abernathy, J. R. and Hovertiz, D. G. *Application of the randomized response techniques in obtaining quantitative data*, *Journal of the American Statistical Associations* **66**, 243–250, 1971.
- [7] Gupta S., Gupta, B. and Singh, S. *Estimation of sensitivity level of personal interview survey questions*, *Journal of Statistical Planning and Inference* **100**, 239–247, 2002.
- [8] Gupta S. and Shabbir J. *Sensitivity estimation for personal interview survey question*, *Statistica Anno LXIV* (4), 643–653, 2004.
- [9] Hussain, Z. and Shabbir, J. *On estimation of mean of a sensitive quantitative variable*, *InterStat*. July (6), 1–14, 2007.
- [10] Hussain, Z. and Shabbir, J. *Improved estimation procedure for the mean of a sensitive variable using randomized response model*, *Pakistan Journal of Statistics* **25** (2), 205–220, 2009.
- [11] Hussain, Z., Shabbir, J. and Gupta, S. *An alternative to Ryu randomized response model*, *Journal of Statistics & Management Systems* **10** (4), 511–517, 2007.
- [12] Mangat, N. S. and Singh, R. *An alternative randomized response procedure*, *Biometrika* **77**, 439–442, 1990.
- [13] Mathur, N. and Singh, H. P. *Estimation of population mean with prior information using scrambled response technique*, *Brazilian Journal of Probability and Statistics* **22** (2), 165–181, 2008.

- [14] Mehta, J. S. and Srinivasan, R. *Estimation of mean by shrinkage to a point*, Journal of the American Statistical Association **66**, 86–90, 1971.
- [15] Ryu, J.-B., Kim, J.-M., Heo, T.-Y. and Park, C. G. *On stratified randomized response sampling*, Model Assisted Statistics and Application **1** (1), 31–36, 2005–2006.
- [16] Saxana, S. *Estimation of exponential mean life in complete and failure censored samples with prior information*, The Philippine Statistician **53** (1-4), 35–45, 2004.
- [17] Searls, D. T. *The utilization of a known coefficient of variation in the estimation procedure*, Journal of the American Statistical Association **59**, 1225–1226, 1964.
- [18] Shirke, D. T. and Nalawade, K. T. *Estimation of the parameter of the binomial distribution in the presence of prior point information*, Journal of the Indian Statistical Association **41** (1), 117–128, 2003.
- [19] Singh, H. P. and Mathur, N. *A revisit to alternative estimator for randomized response technique*, Journal of Indian Society of Agricultural Statistics **55** (1), 79–87, 2002.
- [20] Singh, H. P. and Mathur, N. *An alternative to an improved randomized response strategy*, Statistics in Transition **5** (5), 873–886, 2002.
- [21] Singh, H. P. and Mathur, N. *An optimally randomized response technique*, Aligarh Journal of Statistics **23**, 1–5, 2003.
- [22] Singh, H. P. and Mathur, N. *Modified optimal randomized response sampling technique*, Journal of Indian Society of Agricultural Statistics **56** (2), 199–206, 2003.
- [23] Singh, H. P. and Mathur, N. *Unknown repeated trials in the unrelated question randomized response model*, Biometrical Journal **46** (3), 375–378, 2004.
- [24] Singh, H. P. and Mathur, N. *Estimation of population mean with known coefficient of variation under optimal randomized response model using scrambled response technique*. Statistics in Transition **6** (7), 1079–1093, 2004.
- [25] Singh, H. P. and Mathur, N. *Improved estimation of population proportion possessing sensitive attribute with unknown repeated trial in randomized response sampling*, Statistica Anno LXIV (3), 537–544, 2004.
- [26] Singh, H. P. and Mathur, N. *Estimation of population mean when coefficient of variation is known using scrambled randomized response technique*, Journal of Statistical Planning & Inference **131** (1), 135–144, 2005.
- [27] Singh, H. P. and Mathur, N. *An improved estimation procedure for estimating the proportion of a population possessing sensitive attribute in unrelated question randomized response technique*, Brazilian journal of Probability and Statistics **20**, 93–110, 2006.
- [28] Singh, H. P. and Mathur, N. *An improved estimator for proportion of sensitive group of population using optional randomized response technique*, Assam Statistical Review **21** (1), 64–72, 2007.
- [29] Singh, H. P. and Shukla, S. K. *A class of shrinkage estimators for the variance of exponential distribution with type-I censoring*, IAPQR Transitions **27**, 119–141, 2002.
- [30] Thompson, J. R. *Some shrinkage techniques for estimating the mean*, Journal of the American Statistical Association **63**, 113–123, 1968.
- [31] Tracy, D. S., Singh, H. P. and Raghuvansh, H. S. *Some shrinkage estimators for the variance of exponential density*, Microelectronics Reliability **36** (5), 651–655, 1996.
- [32] Tse, S. and Tse, G. *Shrinkage estimation of reliability of reliability for exponentially distributed lifetimes*, Communication in Statistics-Simulation and Computation **25** (2), 415–430, 1996.
- [33] Warner, S. L. *Randomized response: a survey technique for eliminating evasive answer bias*, Journal of the American Statistical Association **60**, 63–69, 1965.