

Bayesian Analysis of Sparse Counts Obtained From the Unrelated Question Design

Balgobin Nandram¹, Yuan Yu¹

¹ Department of Mathematical Sciences, Worcester Polytechnic Institute, Worcester, MA, US

Correspondence: Correspondence: Balgobin Nandram, Department of Mathematical Sciences, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609, USA. Tel: 1-508-353-5494. E-mail: balnan@wpi.edu

Received: July 17, 2019 Accepted: August 9, 2019 Online Published: August 22, 2019

doi:10.5539/ijsp.v8n5p66 URL: <https://doi.org/10.5539/ijsp.v8n5p66>

Abstract

In sample surveys with sensitive items, sampled units may not respond or they respond untruthfully. Usually a negative answer is given when it is actually positive, thereby leading to an estimate of the population proportion of positives (sensitive proportion) that is too small. In our study, we have binary data obtained from the unrelated-question design, and both the sensitive proportion and the nonsensitive proportion are of interest. A respondent answers the sensitive item with a known probability, and to avoid non-identifiable parameters, at least two (not necessarily exactly two) different random mechanisms are used, but only one for each cluster of respondents. The key point here is that the counts are sparse (very small sample sizes), and we show how to overcome some of the problems associated with the unrelated question design. A standard approach to this problem is to use the expectation-maximization (EM) algorithm. However, because we consider only small sample sizes (sparse counts), the EM algorithm may not converge and asymptotic theory, which can permit normality assumptions for inference, is not appropriate; so we develop a Bayesian method. To compare the EM algorithm and the Bayesian method, we have presented an example with sparse data on college cheating and a simulation study to illustrate the properties of our procedure. Finally, we discuss two extensions to accommodate finite population sampling and optional responses.

Keywords: data augmentation, EM algorithm, Gibbs sampler, latent variables, non-identifiable parameters, proportion, Rao-Blackwellized estimates

1. Introduction

When people are asked sensitive (stigmatizing) questions, there is a tendency for them not to respond or to tell lies. For example, this occurs when questions are asked about annual income, tax evasion, insurance fraud, abortion and students' cheating. One way to reduce these effects is to use the technique of randomized response. In this approach to survey sampling, the randomization is not only in drawing the sample but also in obtaining the response, and there is an enormous literature. One possible design (Greenberg, Abu-Ela, Simmons and Horvitz 1969) is to ask an unrelated nonsensitive (innocuous) question in addition to the sensitive question. The respondents are asked to give an honest answer to one of the two questions selected according to a random mechanism, the essential features of the random mechanism being known to the investigator. This is an extension of the mirrored question design (Warner 1965) in which the respondents are asked the opposite question instead of the unrelated question. When randomized response techniques are used, a respondent's individual answer is not of interest, rather inference is needed for the population, finite or infinite (superpopulation). One needs to strike a compromise between efficiency and response burden but respondents' protection is paramount.

Direct questioning exposes a respondent's privacy that is obviously unacceptable. Any design, which adds noise to the response, will be less efficient than a direct questioning design. One cannot compromise respondents' privacy, but one can compromise respondents' burden and efficiency. However, it has been argued that socially desirable answers and refusals are expected when sensitive questions are asked directly (e.g., see Tourangeau, Rips and Rasinski 2000 and Tourangeau and Yan 2007). Evidently, as supported by many psychologists, sensitive questions should not be asked directly.

We assume that respondents respond truthfully. It should be obvious that this assumption is more easily attained under indirect questioning than direct questioning. Clearly, in direct questioning, it is more likely that there will be nonresponse that may be nonignorable. One way to proceed is to develop nonignorable nonresponse models (e.g., Nandram and Choi 2002a,b, Nandram and Choi 2005, 2010) to handle them. A related alternative approach is to use Bayesian uncertainty analysis (e.g., Nandram and Woo 2015 and Woo, Nandram and Kim 2018) to accommodate nonidentifiable parameters. Both of these approaches are difficult. So at least from two fronts, indirect questioning is advantageous.

As a concrete example, one tosses a die and if one or six comes up, the respondent must give an honest answer to the

sensitive question, and if two, three, four or five comes up, the respondent must give an honest answer to the nonsensitive question. In this way the respondents should be more comfortable to answer the sensitive question because the investigator can never know which question the respondents are answering. We do not rule out the situations where both questions might be sensitive; one of them may be much less sensitive (most respondents do not care). For example, two possible questions are stated below, where students at a university are asked to circle the true answer to the question selected. The questions are

Question 1: Have you ever cheated on an examination anywhere?

Question 2: Do you spend more than twenty hours per week on all courses studying outside the classroom?

Circle your response. [Yes, No]

Here Question 1 is sensitive and Question 2 may be sensitive to some respondents, but much less sensitive than Question 1. We need inference of the proportions of students answering 'yes' to respectively Question 1 and Question 2, the first being the sensitive proportion of greater interest.

Warner (1965) proposed the randomized response method as a survey technique to reduce potential bias due to nonresponse and social desirability when asking questions about sensitive behaviors and beliefs. The method asks respondents to use a randomization device, such as a coin flip, the outcome being unobserved by the interviewer. By introducing random noise, the randomized response method conceals individual responses and protects respondent privacy. As a result, respondents may be more inclined to answer truthfully. For a survey with large sample sizes, this procedure can be computationally assisted.

The proportions of people in the populations who answered "yes" to respectively the sensitive and unrelated (nonsensitive) questions are of interest in the superpopulation. It is well known that the normal equations (see Section 3) do not always provide estimates in $[0, 1]$. The expectation-maximization (EM) can do so, but it may fail to converge to the true value (e.g., Wu 1983), especially when there are sparse counts from the unrelated question design. Moreover, because we are considering only small sample sizes, asymptotic theory to obtain proper confidence intervals that are based on normality is not appropriate. The Bayesian approach to the current problem requires no analytical approximations, although one needs to be careful to execute a Gibbs sampler.

Our paper demonstrates difficulties in running a Gibbs sampler for any randomized response design. Because there are weakly identified parameters, one needs to be careful to run a blocked Gibbs sampler. This issue has nothing to do with what software (e.g., SAS, R, Fortran, C) is used; it is just the algorithm that matters. Unnikrishnan and Kunte (1999) and Oh (1994) provided Bayesian analysis of the unrelated question design, but presumably they have not encountered any difficulties with running the Gibbs sampler because they did not study the convergence of the Gibbs sampler. The Gibbs sampler is not a black-box procedure, and convergence needs to be assessed very carefully. We demonstrate this issue in our paper.

Generally, the unrelated question design is run with exactly two samples (clusters). We can improve the efficiency of the unrelated question design by taking more clusters with smaller sample sizes. In some surveys, this can be a convenience. In fact, this is an additional contribution in our paper.

Throughout, we use the uniform prior because it leaves the likelihood function unchanged, and in this sense it is noninformative. We do not use general beta priors because the parameters of the beta priors need to be specified; we simply comment on how others have used beta priors. Informative priors require elicitation of additional parameters, and this is useful whenever the information is actually available.

The plan of the rest of the paper is as follows. In Section 2, we present a review of the literature of the unrelated question design. In Section 3, we show some difficulties associated with the analysis of the unrelated question design. In Section 4, we present the Bayesian methodology for at least two samples. In Section 5, we discuss empirical studies, where we use a small illustrative example on college cheating, and we describe a small simulation study to show the frequentist properties of the Bayesian model. In Section 6, we present concluding remarks and a discussion of two possible extensions.

2. Unrelated Question Design and Extensions

Blair, Imai, and Zhou (2015) gave an excellent review paper on randomized response techniques (RRT). They classified many of the designs into four types: mirrored question, forced response, disguised response, and unrelated question. For each type, they provide a brief explanation, an example, and a discussion about identification. All four designs make two key assumptions: (1) the randomization distribution is fully known to researchers, and (2) respondents comply with the instructions and answer the questions truthfully.

In the mirrored question design (Warner 1965), a respondent is asked to perform a Bernoulli trial. If a success occurs, the respondent is asked to answer the sensitive question, otherwise the respondent is asked to answer the opposite of the sensitive question. Because a respondent has to answer 'yes' or 'no' to the question and its complement, both are sensitive, a natural drawback of the mirrored question design. In the unrelated question design (Greenberg, Abul-Ela, Simmons and Horvitz 1969), a Bernoulli trial is performed, and if it is a success, the respondent is asked to answer the sensitive question, otherwise the respondent is asked to answer the unrelated question. The unrelated question design is a natural extension of the mirrored question design; the complement question is replaced by an unrelated (nonsensitive) question. The forced response design (Boruch 1971, Fox and Tracy 1986), is like the unrelated question design, but there is no unrelated question. Two Bernoulli trials are performed. In the first trial, if a success occurs, the respondent is asked to answer the sensitive question ('yes' or 'no'); otherwise a different Bernoulli trial is performed. If it is a success, the respondent must answer 'yes'; otherwise the respondent must answer 'no'. In the disguised design (Kuk 1990), the respondent performs two Bernoulli trials. If the answer to the sensitive question is 'yes', report the answer from the first Bernoulli trial, otherwise report the answer from the second Bernoulli trial. This is like the mirrored question design but the answer to the sensitive question is not requested directly.

Researchers have worked on many extensions of Warner's randomized response techniques (RRT). Greenberg et al. (1969), Folsom et al. (1973), Christofides (2005), Odumade and Singh (2009), Mangat (1992), Perri (2008), Mahmood, Singh and Horn (1998), Kim and Warde (2004) are some of them. Chaudhuri and Mukerjee (1988), Fox and Tracy (1986) and Tracy and Mangat (1996) and more recently Chaudhuri (2016) and Chaudhuri and Christofides (2013) have comprehensive discussions on RRT. Finally, see Johnson, Sedory and Singh (2016).

It is important to use an optimal design. When both the probabilities of 'yes' of the sensitive question and the unrelated question are unknown (the case of interest here), we need at least two independent samples (it is standard practice to use exactly two samples). In one sample, a coin is tossed and the probability of "success" is p_1 and in the second sample a coin is tossed and the probability of "success" is p_2 . Greenberg, Abul-Ela, Simmons and Horvitz (1969) used a heuristic argument to suggest a choice of p_1 as $.2 \pm .1$ or $.8 \pm .1$ and $p_2 = 1 - p_1$. Our data on college cheating were collected under this optimality criterion. Moors (1971) provided a more systematic study of optimality and showed that $p_2 = 0$. That is, the randomized experiment should be performed on a sample of n_1 individuals and a sample of n_2 individuals should only be asked the nonsensitive question and none of these should be asked the sensitive question; see also Lanke (1975). This optimal design may not be convenient because the unrelated question might be a bit sensitive to some individuals.

Moors (1971) also showed that the unrelated question design is more efficient than the mirrored question design, everything being equal. There are further increases in efficiency with mild addition of response burden. Mangat, Singh and Singh (1992), Mangat and Singh (1990), Mangat (1992) and Mangat (1994) introduced the two-stage designs. One design is an extension of the mirrored question design and another is an extension of the unrelated question design. For example, two different Bernoulli trials are performed. If the first Bernoulli trial is a success, the respondent answers the sensitive question. If the first trial is a failure, the respondent performs the second Bernoulli trial. If the second Bernoulli trial is a success, the respondents answer the sensitive question; otherwise the respondent answers the unrelated or opposite question depending on whether the mirrored question design or the unrelated question design is used. The two-stage design with the unrelated question is more efficient than the corresponding one for the mirrored question design and each is more efficient than the corresponding one-stage design (Mangat, Singh and Singh 1992, Mangat and Singh 1990, Mangat 1992, Mangat 1994).

We also mention nonrandomized designs that do not use random mechanisms as in the randomized designs. The crosswise and triangular designs (e.g., Tan, Tian and Tang 2009, Tian, Yuen, Tang and Tan 2009) can be viewed as extensions of the unrelated question design because each of them has a sensitive (stigmatizing) question and an unrelated nonsensitive question. Let X be a binary sensitive variable and W denote a nonsensitive variable $X = 0$ or $W = 0$ are 'nos' and $X = 1$ or $W = 1$ are 'yeses'. In the crosswise design, a respondent is asked to answer 'yes' to $X = 1, W = 1$ or $X = 0, W = 0$ and 'no' otherwise. In the triangular design, a respondent is asked to answer 'no' to $X = 0, W = 0$ and 'yes' to $X = 0, W = 1$ or $X = 1, W = 1$ or $X = 1, W = 0$. As in the unrelated question design, if both probabilities are unknown, two samples are needed. There are similar difficulties for estimation and multiple answers are proposed (e.g., Groenitz 2017); the key gain is a random mechanism is not needed.

Greenberg, Kuebler, Abernathy and Horvitz (1971) and Eriksson (1973) extended the unrelated question model of Greenberg, Abul-Ela, Simmons and Horvitz (1969) to the case in which the response is quantitative. There is also a large literature on quantitative (continuous) responses but they are not appropriate for our study.

Prior information about the unknown parameters is sometimes obtainable and can be used along with the sample information for estimation of that unknown parameter. This is the Bayesian approach of estimation. There are not many works within the Bayesian paradigm of randomized response models. Nonetheless, attempts have been made on the Bayesian

analysis of randomized response techniques.

Winkler and Franklin (1979) gave an approximate Bayesian analysis of Warner’s mirrored design, O’Hagan (1987) derived Bayesian linear estimators for the unrelated question design, and Oh (1994) used data augmentation to introduce latent variables to Gibbs sampling of the mirrored design, the unrelated question design and the two-stage design with binary and polychotomous responses. Unnikrishnan and Kunte (1999) performed a similar Bayesian analysis for the unrelated question design and provided a unified Bayesian analysis for the mirrored question design and the unrelated question design. These papers did not discuss the performance of the Gibbs sampler for the unrelated question design perhaps because the sample sizes they considered are much larger. Groenitz (2015) obtained a Bayesian analysis for polychotomous and categorical sensitive items; our interest here is on the unrelated question design for a single sensitive item. Also, Tian, Yuen, Tang and Tan (2009) proposed Bayesian approaches to non-randomized response models without using randomized mechanisms; in non-randomized designs (crosswise and triangular), random mechanisms are not used and they may be more efficient than their randomized response counterparts; see, for example, Tan, Tian and Tang (2009). Most recently, Song and Kim (2017) gave a Bayesian analysis of two rare unrelated questions (i.e., Poisson modeling rather than Binomial modeling). Jayraj, Odumade and Singh (2017) obtained a new quasi empirical Bayes approach; this method is a full Bayesian approach except that the parameters of their beta prior distributions must be specified. However, these researchers were not generally concerned about sparse counts that are of particular interest in our work.

Finally, we note that in 2017, there was a special issue on randomized response techniques in the Journal of Statistics and Applications in honor of Warner. Many of these fifteen papers cover extensions of the mirrored question design and the unrelated question design. Randomized designs for both qualitative and quantitative data were discussed. There are extensions of the unrelated question design to optional response (e.g., Dass and Chhabra, 2017) and to inverse unrelated question design used for additional privacy protection (Dihidar and Basu 2017). But most of the papers are on design-based analyses, generally true in randomized response analyses. In at least one paper, Shaw and Chaudhuri (2017), an empirical Bayes approach, not a Bayesian approach, was used. It is worth noting here that the empirical Bayes approach, like the EM algorithm, generally underestimates variability though.

Our interest is different from the above mentioned works in three ways. First, we are considering very sparse counts obtained from the unrelated question design where we cannot put any reliance on asymptotic normality and the EM algorithm might fail to converge. Second, to help alleviate the problem of sparse counts, we are extending the unrelated question design to accommodate at least two samples with at least two different random mechanisms. Third, for sparse counts, we are providing a full Bayesian analysis by using uniform prior distributions for the proportions of respondents having the sensitive and nonsensitive traits respectively. This is convenient because the likelihood function and the posterior density are the same, thereby providing a fair comparison of the Bayesian method and the frequentist method based on the EM algorithm. Of course, one can improve inference using informative priors (e.g., $\text{Beta}(\alpha, \beta)$ with α and β specified). This prior is equivalent to adding a number of observations, and in effect reducing the sparseness of the counts.

3. Difficulties of the Unrelated Question Design

We discuss some difficulties arising in the analysis of the unrelated question design that has two unknown parameters. We consider the situation where two Bernoulli trials are performed with success probabilities p_1 and p_2 . Both proportions, π_1 and π_2 , of people with the sensitive characteristic and the nonsensitive characteristic respectively are of interest.

We consider the first problem of the design-based estimator. The standard model is

$$y_s \stackrel{ind}{\sim} \text{Binomial}\{n_s, p_s\pi_1 + (1 - p_s)\pi_2\}, s = 1, 2.$$

Let $a_s = y_s/n_s, s = 1, 2$, be the MLE of $p_s\pi_1 + (1 - p_s)\pi_2$. Then, one can find the MLEs of π_1 and π_2 by solving the two equations,

$$a_s = p_s\hat{\pi}_1 + (1 - p_s)\hat{\pi}_2, s = 1, 2,$$

called the normal equations, where $\hat{\pi}_1$ and $\hat{\pi}_2$ are respectively MLEs of π_1 and π_2 provided that $\hat{\pi}_1$ and $\hat{\pi}_2$ lie in $(0, 1)$ (may not happen). We call this the design-based approach and the estimators, $\hat{\pi}_1$ and $\hat{\pi}_2$, design-based estimators.

It can be shown that for $p_1 < p_2$ that $\hat{\pi}_1$ and $\hat{\pi}_2$ lie in $(0, 1)$ provided that

$$\frac{1 - p_2}{1 - p_1} < \min\left\{\frac{a_2}{a_1}, \frac{1 - a_2}{1 - a_1}\right\}, \quad \max\left\{\frac{a_2}{a_1}, \frac{1 - a_2}{1 - a_1}\right\} < \frac{p_2}{p_1}.$$

This is a tight condition especially for small sample sizes that are of interest here. Therefore, $\hat{\pi}_1$ and $\hat{\pi}_2$ may not lie in $(0, 1)$, and this is, indeed, disturbing. Lee, Sedory and Singh (2013) considered the problem of developing minimum sample size requirements for different randomized response models based on guessed values of the parameters of interest

and on the randomization device parameters being used in collecting the datasets. They found that very large sample sizes are needed to get an admissible estimate of the required proportion of a sensitive attribute. Our interest is in sparse counts (small sample sizes).

Of course, one can find the MLEs (possibly local maxima) of π_1 and π_2 by minimizing the negative of the log-likelihood function subject to the constraint that they must lie in $[0, 1]$ and this can be done using the Nelder-Mead algorithm (e.g., Erhardt, Nandram and Choi 2012), which is a versatile iterative procedure and it has been used for many statistical problems of this kind. Yet another approach, in the spirit of the design-based estimators, one can minimize $\sum_{s=1}^2 |a_s - \{p_s \hat{\pi}_1 + (1 - p_s) \hat{\pi}_2\}|$ over $\hat{\pi}_s, s = 1, 2$ subject to the constraint $0 < \hat{\pi}_s < 1, s = 1, 2$ again using the Nelder-Mead algorithm.

The more popular way, the expectation-maximization algorithm (Dempster, Laird and Rubin 1977) can be used to obtain values in $[0, 1]$, possibly MLEs. A book-length discussion of the EM algorithm and its extensions is presented by McLachlan and Krishnan (2008). As is well-known, if one uses the inverse Hessian matrix, evaluated at the mode, to estimate the standard error, there is a possible underestimation of the standard error as we will see empirically. This is a difficulty that the Bayesian method can overcome. We note that the EM algorithm was applied to randomized response (e.g., Bourke and Moran 1988).

The second problem is in getting sensible interval estimators. Because we consider small sample sizes, asymptotic results (e.g., normality assumption) are not appropriate or can be inaccurate. Therefore, the estimates of π_1 and π_2 will not be normally distributed and so interval estimators are not available. Of course, we can write down interval estimators, and these intervals may be too narrow and they may not have the nominal coverage. Again, the Bayesian method can fix this problem as it takes into account all sources of variability and the uncertain assumption of normality to construct confidence intervals is not needed.

The third problem is that $\hat{\pi}_1$ and $\hat{\pi}_2$ can be highly correlated, and this can cause computational instability. This is also disturbing because π_1 and π_2 are completely unrelated. It is true that

$$\begin{aligned} \text{var}(\hat{\pi}_1) &= \frac{(1 - p_2)^2 v_1 + (1 - p_1)^2 v_2}{(p_2 - p_1)^2}, \\ \text{var}(\hat{\pi}_2) &= \frac{p_2^2 v_1 + p_1^2 v_2}{(p_2 - p_1)^2}, \\ \text{cov}(\hat{\pi}_1, \hat{\pi}_2) &= -\left\{ \frac{p_2(1 - p_2)v_1 + p_1(1 - p_1)v_2}{(p_2 - p_1)^2} \right\}, \end{aligned}$$

where

$$v_s = (p_s \pi_1 + (1 - p_s) \pi_2) \{ p_s (1 - \pi_1) + (1 - p_s) (1 - \pi_2) \} / n_s, s = 1, 2.$$

One way to reduce this correlation is to increase the sample size, but this may be costly and prohibitive (or infeasible). Nevertheless, this is a useful result because it warns us to be cautious in running a Gibbs sampler as there will be the weak mixing phenomenon. Similar issues can arise when the EM algorithm is used. Again, the Bayesian method can work with this correlation, provided an efficient blocked Gibbs sampler is constructed.

The design-based estimators obtained by solving the normal equations for MLEs will not necessarily produce MLEs; some truncation will be necessary, and when such truncation is performed, the ad hoc values obtained, 0 or 1, are not really MLEs. [Negative values are set to zero and values larger than one are set to one in an ad hoc manner.] The EM algorithm can actually converge to a value that is not the MLE, but by construction, it will produce a value in $[0, 1]$; e.g., see Wu (1983). At any rate, the Bayesian method does not have these difficulties. There are additional benefits of the Bayesian method in that we can incorporate appropriate prior information about π_1 and π_2 by using a prior distribution that is $\text{Beta}(\alpha, \beta)$, where α and β must be specified (e.g., see Unnikrishnan and Kunte 1999). We cannot be sure about the magnitudes of α and β , unless we do have the appropriate prior information that may be available in some applications. Generally, it is difficult to specify α and β . In the comparisons, to make the EM algorithm and the Bayesian method as similar as possible, we will use independent uniform prior distributions (i.e., $\text{Beta}(1, 1)$) for π_1 and π_2 . One additional advantage of the Bayesian method is that it can provide interval estimators very easily.

As a summary, an objective Bayesian method can solve these problems. It is worth noting that we are not saying that the Bayesian method is the only one that can do so. As we noted, the EM algorithm can fail to converge to the MLEs and it is difficult to get the right standard errors, especially for small sample sizes. We provide evidence to show the better of these two competitors.

4. Bayesian Methodology

In this section, we discuss the Bayesian methodology to analyze data from a randomized response application. We discuss the analysis of sparse counts sampled from a single population using the unrelated question design. We assume that there are g clusters (samples) of individuals, and the respondents within each cluster toss the same random mechanism, but the respondents in different clusters toss possibly different random mechanisms. Because we are studying two items, we need at least two distinct random mechanisms (i.e., at least two clusters, more the better). The unrelated question design has not been run before with more than two samples. In this paper, we are not interested in combining data from a number of small areas, but rather we are interested in a single population. Nandram and Yu (2019) has an example on small area estimation and there are more complicated extensions of the unrelated question design in Yu (2019).

Let p_j denote the probability that the sensitive item is selected for the j^{th} cluster of respondents interviewed, where n_j is the number of respondents in the j^{th} cluster, $j = 1, \dots, g \geq 2$. Note that p_j is known for the j^{th} cluster interviewed. Let π_1 and π_2 denote respectively the probabilities of a ‘yes’ for the sensitive question and the nonsensitive question. Then the probability of getting a ‘yes’ answer from a respondent in the j^{th} cluster is $p_j\pi_1 + (1 - p_j)\pi_2$ and a ‘no’ answer is $p_j(1 - \pi_1) + (1 - p_j)(1 - \pi_2)$. Under random sampling, letting y_j denote the number of ‘yesses’ obtained, we have

$$y_j | \pi_1, \pi_2 \stackrel{iid}{\sim} \text{Binomial}\{n_j, p_j\pi_1 + (1 - p_j)\pi_2\}, j = 1, \dots, g, g \geq 2,$$

where the number of different random mechanisms can be g . Then, the joint probability mass function of $\underline{y} = (y_1, \dots, y_g)'$ is

$$p(\underline{y} | \pi_1, \pi_2) = \prod_{j=1}^g \{p_j\pi_1 + (1 - p_j)\pi_2\}^{y_j} \{p_j(1 - \pi_1) + (1 - p_j)(1 - \pi_2)\}^{n_j - y_j}. \tag{1}$$

Among the g clusters, we assume that there at least two distinct p_j .

It is worth noting that this is a more general model than the one discussed in the literature with just two random mechanisms. By introducing more than two random mechanisms (samples), we can improve the efficiency of the unrelated question design. Clearly, this is better than the standard unrelated question design with two samples and it is more convenient to get responses from smaller groups.

For a full Bayesian analysis, we assume a priori $\pi_s \stackrel{iid}{\sim} \text{Uniform}(0, 1), s = 1, 2$; the posterior density is the same as the likelihood function. However, because of the sparseness of the data, posterior inference may be sensitive to this assumption, and one may need to use a more informative prior (subject to availability). Jayraj, Odumade and Singh (2017) and Unnikrishnan and Kunte (1999) used more general beta priors, but they need to specify the parameters and, of course, there is uncertainty in doing so. Other possibilities are Jeffreys’ prior (not much different from the uniform prior) and Haldane’s prior that may cause posterior impropriety with zero counts of ‘yesses’ or ‘nos’. Yet another useful prior is $\pi_1, \pi_2 \stackrel{iid}{\sim} \text{Uniform}(0, 1), \pi_1 < \pi_2$ (i.e., order statistics from a uniform distribution) that can allow borrowing of strength across π_1 and π_2 . In the college cheating application, this is a realistic piece of prior information. However, for a fair comparison with the EM algorithm, we use the prior $\pi_s \stackrel{iid}{\sim} \text{Uniform}(0, 1), s = 1, 2$.

There are at least two methods to get estimators of π_1 and π_2 within the Bayesian paradigm. First, one can use a grid method to draw π_1 and π_2 separately. This can be done by numerically integrating out one of them, say π_2 , and then draw π_1 using a grid. Then, draw π_2 conditional on π_1 using a grid again. Or, to avoid the numerical integration, one can draw from a bivariate grid on (π_1, π_2) (more computer storage and time are required). Second, one can introduce latent variables as a data augmentation and use the Gibbs sampler; e.g., see Oh (1994). It is convenient that this latter method allows implementation of the more efficient Rao-Blackwellized estimators.

Using the latent variables $(z_j, w_j), j = 1, \dots, g$, the joint posterior density is

$$\begin{aligned} \pi(\pi_1, \pi_2, \underline{z}, \underline{w} | \underline{y}) &\propto \prod_{j=1}^g \binom{y_j}{z_j} (p_j\pi_1)^{z_j} ((1 - p_j)\pi_2)^{y_j - z_j} \\ &\times \prod_{j=1}^g \binom{n_j - y_j}{w_j} (p_j(1 - \pi_1))^{w_j} ((1 - p_j)(1 - \pi_2))^{n_j - y_j - w_j}. \end{aligned}$$

Here, for the sensitive question, respectively z_j and w_j can be interpreted as the number of ‘yesses’ among the y_j ‘yesses’ and the number of ‘nos’ among the $n_j - y_j$ ‘nos’.

For any proper priors on π_1 and π_2 , the joint posterior density of $\pi_1, \pi_2, \underline{z}, \underline{w} | \underline{y}$ is proper because it is proportional to a product of binomial probability mass functions that are all bounded by unity. We have taken $\pi_1, \pi_2 \stackrel{iid}{\sim} \text{Beta}(1, 1)$ (i.e., uniform prior, making the likelihood function and the posterior density the same).

We use a blocked Gibbs sampler that is easy to run because the conditional posterior densities are all in simple forms,

$$z_j | \pi_1, \pi_2, y_j \overset{ind}{\sim} \text{Binomial}\{y_j, \frac{p_j \pi_1}{p_j \pi_1 + (1 - p_j) \pi_2}\}, j = 1, \dots, g,$$

$$w_j | \pi_1, \pi_2, y_j \overset{ind}{\sim} \text{Binomial}\{n_j - y_j, \frac{p_j(1 - \pi_1)}{p_j(1 - \pi_1) + (1 - p_j)(1 - \pi_2)}\}, j = 1, \dots, g.$$

Note that given π_1, π_2, y_j , the z_j and w_j are independent. That is, all the (z_j, w_j) are run in a single block, as part of the blocked Gibbs sampler. More importantly, letting $y. = \sum_{j=1}^g y_j, z. = \sum_{j=1}^g z_j, w. = \sum_{j=1}^g w_j,$

$$\pi_1 | z., w., y. \sim \text{Beta}(z. + 1, w. + 1),$$

$$\pi_2 | z., w., y. \sim \text{Beta}(y. - z. + 1, n. - y. - w. + 1).$$

Again, it is convenient that, given $z., w., y., \pi_1$ and π_2 are independent and they are beta random variables. Essentially, for $j = 1, \dots, g$, given π_1, π_2 and $y_j, (z_j, w_j)$ are all drawn at once (a single block), and given $z., w.$ and $y., (\pi_1, \pi_2)$ are drawn at once (a single block). That is, the blocked Gibbs sampler is performed with two blocks.

This independence is important because it provides a better mixing Gibbs sampler than if they were correlated. One needs to be careful in running any Gibbs sampler. However, it is important to note that the conditional density function of π_1 is a function of the missing data, $z., w.,$ but not a function of the observed data, $y.;$ so that π_1 is weakly identified. This will also impact the estimation of $\pi_2.$ One way to get around this problem is to use a large number of runs with substantial thinning in the blocked Gibbs sampler.

However, we can obtain Rao-Blackwellized estimators (smallest mean squared errors) of π_1 and π_2 because

$$\begin{aligned} \pi(\pi_1, \pi_2 | y) &= \sum_{z_1=0}^{y_1} \sum_{w_1=0}^{n_1-y_1} \dots \sum_{z_g=0}^{y_g} \sum_{w_g=0}^{n_g-y_g} \pi(\pi_1, \pi_2 | z., w., y) \pi(z., w. | y) \\ &= \sum_{z_1=0}^{y_1} \sum_{w_1=0}^{n_1-y_1} \dots \sum_{z_g=0}^{y_g} \sum_{w_g=0}^{n_g-y_g} \pi(\pi_1 | z., w., y) \pi(\pi_2 | z., w., y) \pi(z., w. | y) \end{aligned}$$

as follows. Let $(z_j^{(h)}, w_j^{(h)}), j = 1, \dots, g, h = 1, \dots, M,$ denote a random sample of size M from the posterior density, $\pi(z., w. | y),$ obtained from the Gibbs sampler. Then, the Rao-Blackwellized density estimator of $\pi(\pi_1, \pi_2 | y)$ is

$$\pi(\widehat{\pi_1}, \widehat{\pi_2} | y) = \frac{1}{M} \sum_{h=1}^M \pi(\pi_1 | z^{(h)}, w^{(h)}, y) \pi(\pi_2 | z^{(h)}, w^{(h)}, y).$$

It follows that the Rao-Blackwellized posterior means of π_1 and π_2 are

$$E(\widehat{\pi_1} | y) = \frac{1}{M} \sum_{h=1}^M \frac{z^{(h)} + 1}{z^{(h)} + w^{(h)} + 2} \text{ and } E(\widehat{\pi_2} | y) = \frac{1}{M} \sum_{h=1}^M \frac{y^{(h)} - z^{(h)} + 1}{n^{(h)} - z^{(h)} - w^{(h)} + 2}.$$

Similar formulas, which are a bit cumbersome, can be written down for the posterior variances.

The first method, using grids, does not allow Rao-Blackwellization. This is an important issue because Rao-Blackwellization provides improved estimators beyond the ordinary Monte Carlo sample mean. This was not discussed in the Bayesian approach of the analysis of data from the unrelated question design in Oh (1994), Jayraj, Odumade and Singh (2017) and Unnikrishnan and Kunte (1999).

5. Empirical Studies

In Section 5.1, we describe an application on college cheating and, in Section 5.2, we perform a simulation study to assess the performance of the Bayesian model. Our main objective is to compare the Bayesian method via the Gibbs sampler and the frequentist method via the EM algorithm.

5.1 Illustrative Example

The data on college cheating were collected by a social science professor at a small University. This is an important practical issue because college cheating has become a serious problem in the US (e.g., Shon 2006 and Jones 2011).

The professor asked the students in the class to carry out the survey and find 20-30 different students on campus to ask these questions. He told them to be certain that this is the first time the respondent is doing this questionnaire. The students in the class were divided into teams of 2-3 to make 11 teams and they were asked to visit various locations (campus center, library, fitness center, food courts, etc.) on campus to carry out the survey.

The students were given specific instructions as follows.

“When each of you collect the data, you would need to use at least two different random mechanisms. This has to be done because both the probabilities for a ‘yes’ answer of the sensitive question and the nonsensitive questions are assumed to be unknown. For example, if you use a six-sided die, you may take 1 or 2 for the sensitive question and 3, 4, 5, or 6 for the nonsensitive question for some students and 1,2, 3, or 4 for the sensitive question and 5 or 6 for the nonsensitive question for other students. You must not use the same random mechanism for all the students you interviewed; the example just discussed has two random mechanisms. However, you must record which mechanism you use for a particular student. You would ask each student to toss the die, and answer honestly, ‘yes’ or ‘no’, to the question that turns up, but the student must not tell you which question she/he is answering. When you interview the students, you must ensure that they are giving you independent answers. If you enjoy collecting the data, you may try a third random mechanism.”

Note that the nonsensitive question is a bit sensitive and the optimal design was not used; see Moors (1971). The experiment was performed at 11 locations on campus. There are more descriptions of these data in Nandram and Yu (2017).

However, the data are confidential, so we can only present a limited view. Except for one location, where three samples were taken, two samples were taken from all locations. The smallest number of students sampled is from a location with 18, 9 per sample. The largest number of students sampled is from another location with 30, 20 from one sample and 10 from the other sample. Therefore, the counts are indeed very small (sparse counts).

When we fit the Gibbs sampler to the data from each location, we ran 20,000 iterations, used 5000 as a “burn-in” and picked every 15th thereafter. Again, this long run is needed because π_1 is weakly identified. For this sample, all the autocorrelations are nonsignificant, effective sample sizes are 1000 for all locations except two of them, which were nearly 800 (acceptable) for π_1 and π_2 . The Geweke test of stationarity did not reject stationarity. The computations for the 11 locations took just less than one minute. Thus, the Gibbs sampler performed very well for all cases despite the long run that is needed.

We study convergence of the EM algorithm. Let $\hat{\pi}_t$ denote the value of π_1 or π_2 after the t^{th} iteration of the EM algorithm. We consider $\frac{|\hat{\pi}_t - \hat{\pi}_{t-1}|}{\hat{\pi}_{t-1}} \leq .001$ to assess convergence. Typically in this application, if the EM algorithm converges, this criterion is met after 10-20 iterations. If the criterion is not met at 100 iterations of the EM algorithm, we declare that the EM algorithm has not converged, and the current value is selected. We observed four cases in which there are no convergence for locations 3, 4, 6 and 11, and these are for π_1 , the more important parameter. In Table 1 and Table 2, we present some results for inference about π_1 and π_2 .

First, in Table 1, we compare the design-based estimators by solving the normal equations and the estimators obtained from the EM algorithm. Some of the design based estimators are not good; they are mostly out of range, and therefore, the SE, CV and correlation are all incorrect. Of course, one can truncate these estimators to $[0, 1]$ as is usually recommended for design-based estimators (e.g., a negative estimate is set to 0 and an estimator larger than 1 is set to 1). However, the EM algorithm does a relatively much better job, although one cannot be sure that EM algorithm really converges to the correct value, but it is good to see these estimates because they appear reasonable. In the note to Table 1, we have combined all 11 cases into a single data set (i.e., we assume a single (π_1, π_2) for all 11 cases). Here, the EM algorithm converges very quickly in 15 iterations. Using the EM algorithm, the overall estimates (standard errors) for π_1 and π_2 are respectively .279 (.067) and .798 (.037). Therefore, one can believe these estimates, but with respect to π_1 , those for cases 3 and 11 appear to be out of line; the EM algorithm does not converge for case 3. However, note that the EM algorithm has done the right truncation in cases 3 and 11 (5.700 becomes 1.000 and -.556 becomes 0.000, actually MLEs).

Second, in Table 2, we compare the estimates from the EM algorithm and the Bayesian method. Note that the Bayesian method provides much more reasonable estimates, standard errors and coefficient of variations than the EM algorithm. The estimates are similar in some cases. Also, we can see that the absolute correlations between π_1 and π_2 are larger for the Bayesian model; the EM algorithm tends to underestimate variability. We note that the answers from the combined data are very similar except that the posterior standard deviations are larger than the standard errors under the EM algorithm.

Third, in Table 3, we provide a comparison of the Gibbs sampler with a numerical integration method that gives random samples without latent variables rather than a Markov chain sample with latent variables. Let $\pi(\pi_1, \pi_2 | \underline{y})$ denote the joint

posterior density of (π_1, π_2) without the latent variables. Then,

$$\pi(\pi_1, \pi_2 | \underline{y}) = \pi_1(\pi_1 | \pi_2, \underline{y})\pi_2(\pi_2 | \underline{y}), \text{ where } \pi_2(\pi_2 | \underline{y}) = \int_0^1 \pi(\pi_1, \pi_2 | \underline{y})d\pi_1.$$

We use Gauss-Legendre quadrature with 100 roots to do the integration at each choice of a grid of points of π_2 in $[0, 1]$, and the grid method is used to draw π_2 . For each draw of π_2 , π_1 is drawn from $\pi_1(\pi_1 | \pi_2, \underline{y})$ using a grid method. We use 100 grid points in both cases. Table 3 shows that the method based on the Gibbs sampler is very close to the random sampler (numerical method), thereby showing that we have a very efficient Gibbs sampler. Although this problem is somewhat simple, the blocked Gibbs sampler is highly efficient for more complicated problems.

Our next step is to perform a simulation study to see if the Bayesian method is actually better than the EM algorithm. Our strategy is to show that the Bayesian method performs very well in terms of absolute relative bias and coverage. We show that the EM algorithm usually converges to the right value, but sometimes it does not (this is the uncertainty we worry about).

5.2 Simulation Study

We run a simulation study to assess the performance of the Bayesian method. Our objective is modest and we want to study the frequentist properties of the Bayesian method and to compare it with the EM algorithm. For these purposes, we consider two scenarios, one with sparse counts and the other with several samples of these sparse counts, to obtain a more favorable outcome of the EM algorithm.

The design plan is as follows. First, we set $\pi_1 = .35$ and $\pi_2 = .65$, similar to the estimates from the real data on college cheating. We have selected $g \sim \text{Uniform}(2, 5)$, and there are 2 – 5 samples in each of ten clusters (groups) of data. For each sample, we have taken the sample size $n_j \sim \text{Uniform}(10, 20)$, reflecting small sample sizes. We have taken $p_j \sim \text{Uniform}(.10, .30)$ for $j = 1, 3, 5$ and $p_j \sim \text{Uniform}(.70, .90)$ for $j = 2, 4$ consistent with a version of optimality. These parameters are held fixed throughout the experiment.

The simulated counts are obtained as follows. For $j = 1, \dots, g$, we draw $U \sim \text{Uniform}(0, 1)$; if $U \leq p_j$, we draw $I \sim \text{Bernoulli}(\pi_1)$; otherwise we draw $I \sim \text{Bernoulli}(\pi_2)$. For $j = 1, \dots, g$, we perform this procedure n_j times (sample size) to get y_j , the number of ‘yeses’. We use 1,000 runs; each run has ten clusters.

To compare the performance of the EM algorithm and the Bayesian method, we compare two scenarios in which (a) for each individual run (cluster), our model has a single pair of (π_1, π_2) , a total of 10,000 runs and (b) for the combined run (ten clusters) our model has a single pair of (π_1, π_2) , a total of 1,000 runs.

We fit each of the 10,000 simulated runs using the Bayesian model and the EM algorithm in exactly the same manner as described for the data on college cheating. For each of the 10,000 simulated runs, we performed the Geweke test of stationarity, calculated the effective sample sizes and the auto-correlations. We found that the performance of the Gibbs sampler was satisfactory in all cases. We did the same for the 1000 combined runs (ten clusters each) and got very good performance. For the combined data, the EM algorithm converges for all 1,000 runs. Each of the 10,000 runs took 0.059 second and each of the 1,000 runs took .230 second on our Solar Cluster; these include the EM algorithm as well. For the 10,000 runs, there were 230 runs (just about 2% of the runs) in which the EM algorithm did not converge for π_1 but there were no convergence issues for π_2 . For the 230 runs with convergence issues, π_1 converges to 0.000 far from the truth of $\pi_1 = .35$. The Gibbs sampler performed better with estimates of π_1 around .20, not quite .35 though.

We computed the posterior means (PM), posterior standard deviations (PSD), the numerical standard errors (NSE), the 95% HPD intervals and its width, and the correlation between π_1 and π_2 for all locations. Then, at a macro level, we take the averages and standard errors for all quantities over all simulated runs and all locations. Specifically, we calculated the bias, $B = (PM - T)$, absolute relative bias, $ARB = |PM - T| / T$, and the posterior root mean squared error, $PRMSE = \sqrt{(PM - T)^2 + PSD^2}$, where T denotes the true proportion, π_1 or π_2 (known by simulation). With respect to the intervals, we computed their average width (Wid) and the coverage C , which is the proportion of intervals containing the true value in the 1000 simulated runs. We also computed the average of the correlations between π_1 and π_2 . We computed similar quantities for the EM algorithm, where for convenience, we use the notation where the PM is the MLE, the PSD is the standard error, and the 95% confidence interval is $PM \pm 1.96PSD$ (assuming normality).

In the next two tables, we will provide numerical standard errors as subscripts. This is an idea that was advised by an editor of the Journal of the American Statistical Association in Nandram and Choi (2002 a). This notation was used in many other papers because it is convenient and elegant.

In Table 4, we present summaries for π_1 . First, consider groups 1-10. Bias is smaller for the EM algorithm than the Bayesian method; the Bs are closer to zero, but they are both close to zero. We notice that the PSDs are mostly smaller

than those from the EM algorithm; otherwise they are reasonably close. However, ARB and PRMSE are smaller for the Bayesian method. The Wids are mostly similar; intervals for Bayesian method are generally shorter than those from the EM algorithm. But what matters is coverage and the Cs under the Bayesian method are closer to the nominal value of .95. Second, for the combined runs, the two approaches are close, but again the Bayesian method has better coverage (.938 versus .916). Therefore, for inference about π_1 , particularly for sparse counts, the Bayesian method is preferred.

In Table 5, we present summaries for π_2 , the less important parameter. First, consider groups 1-10. Again bias is smaller for the EM algorithm than the Bayesian method; the Bs are closer to zero, both are close to zero, and they are all negative for the Bayesian method and two of them are positive for the EM algorithm. For π_2 , the PSDs are larger under the Bayesian method; the EM algorithm generally underestimates variability. However, ARBs are smaller, but PRMSEs are larger, for the Bayesian method. The Wids for the EM algorithm are much smaller, thereby making the coverage much too small; the coverage for the Bayesian procedure are much closer to the nominal value of .95. Second, for the combined runs, the two approaches are close, but now the Bayesian method has much better coverage (.945 versus .792). Therefore, again for inference about π_2 , particularly for sparse counts, the Bayesian method is preferred, at least with respect to two measures, absolute relative bias and coverage.

Finally, we have looked at the correlations between π_1 and π_2 averaged over the number of runs. For both the Bayesian method and the EM algorithm, the standard errors are smaller than .002. For the individual locations, the correlations for the Bayesian method (EM algorithm) are $-.485$ ($-.391$), $-.530$ ($-.384$), $-.531$ ($-.384$), $-.482$ ($-.391$), $-.483$ ($-.390$), $-.484$ ($-.391$), $-.441$ ($-.354$), $-.530$ ($-.384$), $-.486$ ($-.391$), $-.532$ ($-.385$) and for the ten groups combined $-.546$ ($-.384$). The correlations for the EM algorithm are smaller in magnitude. This is due to the underestimation of the correlation matrix of π_1 and π_2 , but these correlations are not as large in magnitude as for the real data on college cheating.

6. Concluding Remarks

We have shown how to analyze data from the unrelated question design using a Bayesian method. In particular, we have shown how to provide an efficient blocked Gibbs sampler; Unnikrishnan and Kunte (1999) and Oh (1994) did not emphasize the construction of an efficient Gibbs sampler and no diagnostic tests were demonstrated. We have pointed out several problems associated with design-based method and the EM algorithm for this problem with sparse data, and we have shown how the Bayesian method can overcome these problems. This is evident in our example on college cheating and our simulation study.

In a moderate-size simulation study, we have shown that the Bayesian method via the Gibbs sampler is preferred over the frequentist method via the EM algorithm. The EM algorithm does not always converge, although it does most of the time. But it is important to remove this uncertainty using the Bayesian method so that proper inference can be made. The Gibbs sampler always converges, albeit with very long runs, but the running time is not significant. We have shown that the Bayesian method out performs the EM algorithm in terms of absolute relative bias and coverage.

We consider two extensions, one to finite population sampling and other to optional responses. For randomized response designs, estimation of finite population quantities (means, totals) is not as straight forward in the non-Bayesian method as in the Bayesian method. Incorporating optional responses into the unrelated question design can provide improved inference. We have not considered very complicated extensions; Yu (2019) considered extensions to multinomial models with arbitrary correlation structure for small areas, starting with the approximations in Nandram (1998).

Our first extension is how to do prediction in a finite population under simple random sampling without selection bias. We have actually estimated the parameters π_1 and π_2 , and to make inference about the finite population proportions, we need an additional computational step. Once our individual-area model is fit, we will obtain π_1 and π_2 . Therefore, letting N the total number of individuals in the finite population, the total number of "yeses" are

$$X_s | \pi_1 \stackrel{ind}{\sim} \text{Binomial}(N, \pi_s), s = 1, 2.$$

Then, the finite population proportion $P_s = X_s/N$, $s = 1, 2$, and inference about the P_s can be made in a straight forward manner under the Bayesian model. It is worth noting that we need to sample the entire population of binary variables, not just the non-sample binary variables. Of course, we are assuming that there is no selection bias.

Our second extension is to optional responses. Gupta, Tuck, Gill and Crowe (2013) showed how to accommodate binary data under the optional unrelated question design. In this design, if the sensitive question does not stigmatize a respondent, then she/he is given the option to answer it directly (without using a random mechanism). We assume that some respondents will exercise this option and others will not, and those, who do not choose this option, will follow the unrelated question design. In addition to the parameters in the unrelated question design, we also have the sensitivity-level parameter, which is the probability that a respondent chooses the unrelated question design. This option improves the efficiency of the unrelated question design because not everyone will use the random mechanism (i.e., less noise). There

are actually two cases, one in which we know who exercise the option and the other in which we do not know who exercise this option; Gupta, Tuck, Gill and Crowe (2013) researched this latter case using the frequentist approach, and we have been studying both cases using the Bayesian method.

Finally, we have found that for the unrelated question design, inference about π_1 and π_2 is sensitive to the random mechanism and the unrelated question, especially for small sample sizes. This problem is important and it needs future attention.

Acknowledgement

Dr Nandram's work was supported by a grant from the Simons Foundation (grant 353953, Balgobin Nandram). The authors thank the reviewers for their comments. The Google Drive,

<https://drive.google.com/open?id=1zqZDHY4RIS4bpW4Run4XZy7Nuq7871h->,

has basic codes.

References

- Blair, G., Imai, K., & Zhou, Y-Y. (2015). Design and Analysis of the Randomized Response Technique. *Journal of the American Statistical Association, Review, 110*, 1304-1319.
- Bourke, P. D., & Moran, M. A. (1988). Estimating Proportions From Randomized Response Data Using the EM Algorithm. *Journal of the American Statistical Association, 83*, 964-968.
- Boruch, R. F. (1971). Assuring Confidentiality of Responses in Social Research: A Note on Strategies. *American Sociologist, 6*, 308-311.
- Chaudhuri, A., & Mukerjee, R. (1988). *Randomized Response: Theory and Techniques*. Dekker: New York.
- Chaudhuri, A. (2016). *Randomized Response and Indirect Questioning Techniques in Surveys*. CRC Press, Chapman and Hall, New York, USA.
- Chaudhuri, A., & Christofides, T. C. (2013). *Indirect Questioning in Sample Surveys*. Springer-Verlag, Heidelberg, Germany.
- Christofides, T. C. (2005). Randomized Response Techniques for Two Sensitive Characteristics at the Same Time. *Metrika, 62*, 53-63.
- Dass, B. K., & Chhabra, A. (2017). Generalized Multi-Stage Optional Unrelated Question RRT Models. *Statistics and Applications, 15*(1-2), 7-18.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society, Series B, 39*(1), 1-38.
- Dihidar, K., & Basu, L. (2017). Privacy Protection in Estimating Sensitive Population Proportion by a Modified Unrelated Question Model. *Statistics and Applications, 15*(1-2), 19-25.
- Erhardt, E. B., Nandram, B., & Choi, J. W. (2012). Bayesian Simultaneous Intervals for Small Areas: An Application to Variation in Maps. *International Journal of Statistics and Probability, 1*(2), 229-243.
- Eriksson, S. A. (1973). A New Model for Randomized Response. *International Statistical Review, 41*, 101-113.
- Folsom, R. E., Greenberg, B. G., & Horvitz, D. G. (1973). The Two Alternate Questions RR Model for Human Surveys. *Journal of the American Statistical Association, 68*, 525-530.
- Fox, J. A., & Tracy, P. E. (1986). *Randomized Response: A Method for Sensitive Surveys*, Sage: London.
- Greenberg, B. G., Abul-Ela, A. L. A., Simmons, W. R., & Horvitz, D. G. (1969). The Unrelated Question Randomized Response Model: Theoretical Framework. *Journal of the American Statistical Association, 64*, 520-539.
- Greenberg, B. G., Kuebler, R. R., Abernathy, J. R., & Horvitz, D. G. (1971) Application of the Randomized Response Technique in Obtaining Quantitative Data. *Journal of the American Statistical Association, 66*, 243-250.
- Groenitz, H. (2017). Improving Estimation Accuracy in Nonrandomized Response Questioning Methods by Multiple Answers. *International Journal of Statistics and Probability, 6*(5), 101-109.
- Groenitz, H. (2015). Using Prior Information in Privacy-Protecting Survey Designs for Categorical Sensitive Variables. *Statistical Papers, 56*, 167-189.
- Gupta, S., Tuck, A., Gill, T. S., & Crowe, M. (2013). Optional Unrelated Question Randomized Response Models. *Involve: A Journal of Mathematics, 6*(4), 483-492.

- Jayraj, A., Odumade, O., & Singh, S. (2017). A New Quasi Empirical Bayes Estimate in Randomized Response Sampling. *Communications in Statistics - Simulation and Computation*. <https://doi.org/10.1080/03610918.2017.1332208>
- Jones, D. L. R. (2011). Academic Dishonesty: Are More Students Cheating? *Business Communication Quarterly*, 74(2), 141-150.
- Johnson, M. L., Sedory, S. A., & Singh, S. (2016). Alternative Methods to Make Efficient Use of Two Decks of Cards in Randomized Response Sampling. *Sociological Methods and Research*, 1-30.
- Kim, J. M., & Warde, W. D. (2004). A Stratified Warner Randomized Response Model. *Journal of Statistical Planning and Inference*, 120, 155-165.
- Kuk, A. Y. (1990). Asking Sensitive Questions Indirectly. *Biometrika*, 77, 436-438.
- Lanke, J. (1975). On the Choice of the Unrelated Question in Simmons' Version of Randomized Response. *Journal of the American Statistical Association*, 70, 80-83.
- Lee, C-S., Sedory, S. A., & Singh, S. (2013). Simulated Minimum Sample Size Requirements in Various Randomized Response Models. *Communications in Statistics - Simulation and Computation*, 42(4), 771-789.
- Mangat, N. S. (1994). An Improved Randomized Response Strategy. *Journal of the Royal Statistical Society, Series B*, 56(1), 93-95.
- Mangat, N. S. (1992). Two Stage Randomized Response Sampling Procedure Using Unrelated Question. *Journal of the Indian Society of Agricultural Statistics*, 44(1), 82-87.
- Mangat, N. S., & Singh, R. (1990). An alternative randomized response procedure. *Biometrika*, 77(2), 439-442.
- Mangat, N. S., Singh, R., & Singh, S. (1992). An Alternative Randomized Response Procedure. *Calcutta Statistical Association Bulletin*, 42, 277-281.
- Mahmood, M., Singh, S., & Horn, S. (1998). On the Confidentiality Guaranteed Under Randomized Response Sampling: A Comparison with Several New Techniques. *Biometrical Journal*, 40, 237-242.
- McLachlan, G. J., & Krishnan, T. (2008). *The EM Algorithm and Extensions*, 2nd edition, Wiley: Hoboken, New Jersey.
- Moors, J. (1971). Optimization of the Unrelated Question Randomized Response Model. *Journal of the American Statistical Association*, 66, 627-629.
- Nandram, B. (1998). A Bayesian Analysis of the Three-stage Hierarchical Multinomial Model. *Journal of Statistical Computation and Simulation*, 61(1-2), 97-126.
- Nandram, B., & Yu, Y. (2017). Bayesian Analysis of Sparse Counts Under the Unrelated Question Design. *JSM Proceedings*. Alexandria, VA: American Statistical Association, 1162-1175.
- Nandram, B., & Yu, Y. (2019). Bayesian Analysis of a Sensitive Proportion for a Small Area. *International Statistical Review*, 87(1), 104-120. <https://doi.org/10.1111/insr.12286>
- Nandram, B., & Choi, J. W. (2010). A Bayesian Analysis of Body Mass Index Data From Small Domains Under Nonignorable Nonresponse and Selection. *Journal of the American Statistical Association*, 105, 120-135.
- Nandram, B., & Choi, J. W. (2002a). Hierarchical Bayesian Nonresponse Models for Binary Data from Small Areas with Uncertainty about Ignorability. *Journal of the American Statistical Association*, 97, 381-388.
- Nandram, B., & Choi, J. W. (2002b). A Bayesian Analysis of a Proportion under Nonignorable Nonresponse. *Statistics in Medicine*, 21, 1189-1212.
- Nandram, B., & Choi, J. W. (2005). Hierarchical Bayesian Nonignorable Nonresponse Regression Models for Small Areas: An Application to the NHANES Data. *Survey Methodology*, 31(1), 73-84.
- Nandram, B., & Woo, N. (2015). A Bayesian uncertainty analysis for nonignorable nonresponse. *In Current Trends in Bayesian Methodology with Applications*, 415C433. (Edited by S. K. Upadhyay, U. Singh, D. K. Dey and A. Longanathan). Chapman and Hall/CRC, Boca Raton.
- Perri, P. F. (2008). Modified Randomized Devices for Simmons' Model. *Model Assisted Statistics and Applications*, 3, 233-239.
- Oh, M. (1994). Bayesian Analysis of Randomized Response Models: A Gibbs Sampling Approach. *Journal of the Korean Statistical Society*, 23, 463-482.
- Odumade, O., & Singh, S. (2009). Efficient Use of Two Deck of Cards in Randomized Response Sampling. *Communica-*

- tions in *Statistics - Theory and Methods*, 38, 439-446.
- O'Hagan, A. (1987). Bayes Linear Estimators for Randomized Response Models. *Journal of the American Statistical Association*, 82, 580-585.
- Shaw, P., & Chaudhuri, A. (2017). Empirical Bayes Estimation Method in Some Randomized Response Techniques. *Statistics and Applications*, 15(1-2), 101-116.
- Shon, P. C. H. (2006). How College Students Cheat on In-Class Examinations: Creativity, Strain, and Techniques of Innovation. *Plagiarism: Cross Disciplinary Studies in Plagiarism, Fabrication, and Falsification*, 130-148.
- Song, J. J., & Kim, J-M. (2017). Bayesian Estimation of Rare Sensitive Attribute. *Communications in Statistics - Simulation and Computation*. <https://doi.org/10.1080/03610918.2015.1109655>
- Tan, M. T., Tian, G.-L., & Tang, M-L. (2009). Sample Surveys with Sensitive Questions: A Randomized Response Approach. *The American Statistician*, 63, 9-16.
- Tian, G.-L., Yuen, K. C., Tang, M.-L., & Tan, M. T. (2009). Bayesian Non-randomized Response Models for Surveys with Sensitive Questions, *Statistics and Its Interface*, 2, 13-25.
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The Psychology of Survey Response*. Cambridge University Press: Cambridge, England.
- Tourangeau, R., & Yan, T. (2007). Sensitive Questions in Surveys. *Psychological Bulletin*, 133, 859-883.
- Tracy, D. S., & Mangat N. S. (1996). On Respondent's Jeopardy in Two Alternate Questions Randomized Response Model. *Journal of Statistical Planning and Inference*, 55(1), 107-114.
- Unnikrishnan, N. K., & Kunte, S. (1999). Bayesian Analysis of randomized Response Models. *Sankhya, Series B*, 61, 422-432.
- Warner, S. L. (1965). Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias. *Journal of the American Statistical Association*, 60, 63-69.
- Winkler, R. L., & Franklin, L. A. (1979). Warner's Randomized Response Model: A Bayesian Approach. *Journal of the American Statistical Association*, 74, 207-214.
- Woo, N., Nandram, B., & Kim, D. (2018). Bayesian Small Area Models For Three-way Contingency Tables with Nonignorability. *Statistica Sinica*, 28, 1839-1866.
- Wu, C. F. J. (1983). On the Convergence properties of the EM Algorithm. *Annals of Statistics*, 11(1). 95-103.
- Yu, Y. (2019). Bayesian Small-Area Analyzes of the Unrelated Question Design with Multiple Sensitive Questions. *PhD Dissertation*, Mathematical Sciences, Worcester Polytechnic Institute.

Appendix A

EM Algorithm

We obtain the EM algorithm for a single area, where

$$y_j | \pi_1, \pi_2 \stackrel{ind}{\sim} \text{Binomial}(n_j, p_j\pi_1 + (1 - p_j)\pi_2), j = 1, \dots, g.$$

We introduce latent variables $(z_j, w_j), j = 1, \dots, g$, as in the Gibbs sampler.

Then, for the expectation step,

$$z_j | \pi_1, \pi_2, y_j \stackrel{ind}{\sim} \text{Binomial}\left\{y_j, \frac{p_j\pi_1}{p_j\pi_1 + (1 - p_j)\pi_2}\right\}, j = 1, \dots, g,$$

$$w_j | \pi_1, \pi_2, y_j \stackrel{ind}{\sim} \text{Binomial}\left\{n_j - y_j, \frac{p_j(1 - \pi_1)}{p_j(1 - \pi_1) + (1 - p_j)(1 - \pi_2)}\right\}, j = 1, \dots, g.$$

Therefore,

$$E(z_j | \pi_1, \pi_2, y_j) = \frac{y_j p_j \pi_1}{p_j \pi_1 + (1 - p_j) \pi_2}, \quad E(w_j | \pi_1, \pi_2, y_j) = \frac{(n_j - y_j) p_j (1 - \pi_1)}{p_j (1 - \pi_1) + (1 - p_j) (1 - \pi_2)},$$

$j = 1, \dots, g$.

For the maximization step, it is convenient that, given π_1, π_2, y_j, z_j and w_j are independent. It is worth noting that the z_j and w_j are replaced by their expectations. More importantly, letting $y. = \sum_{j=1}^g y_j, z. = \sum_{j=1}^g z_j, w. = \sum_{j=1}^g w_j,$

$$\hat{\pi}_1 = \frac{z.}{z. + w.},$$

$$\hat{\pi}_2 = \frac{y. - z.}{n. - z. - w.}.$$

It is convenient that again, given $z., w., y., \pi_1$ and π_2 are independent and they are beta random variables.

The estimated covariance matrix of $\hat{\pi}_1$ and $\hat{\pi}_2$, based on the negative inverse Hessian matrix, is

$$C = \begin{pmatrix} 1/a & -c/ab \\ -c/ab & 1/b \end{pmatrix} / (1 - c^2/ab).$$

where

$$a = \sum_{j=1}^g p_j^2 \left\{ \frac{y_j}{(p_j \hat{\pi}_1 + (1 - p_j) \hat{\pi}_2)^2} + \frac{(n_j - y_j)}{(p_j (1 - \hat{\pi}_1) + (1 - p_j) (1 - \hat{\pi}_2))^2} \right\},$$

$$b = \sum_{j=1}^g (1 - p_j)^2 \left\{ \frac{y_j}{(p_j \hat{\pi}_1 + (1 - p_j) \hat{\pi}_2)^2} + \frac{(n_j - y_j)}{(p_j (1 - \hat{\pi}_1) + (1 - p_j) (1 - \hat{\pi}_2))^2} \right\},$$

$$c = \sum_{j=1}^g p_j (1 - p_j) \left\{ \frac{y_j}{(p_j \hat{\pi}_1 + (1 - p_j) \hat{\pi}_2)^2} + \frac{(n_j - y_j)}{(p_j (1 - \hat{\pi}_1) + (1 - p_j) (1 - \hat{\pi}_2))^2} \right\}.$$

This covariance matrix, C , gives an underestimate of the true variability of $\hat{\pi}_1$ and $\hat{\pi}_2$ because $\hat{\pi}_1$ and $\hat{\pi}_2$ themselves are substituted into the population covariance matrix. But this can be improved a little bit if the log-likelihood, without the latent variables, is used to form the Hessian matrix.

Appendix B

Table 1. Comparison of the estimates (E), standard errors (SE), coefficients of variation (CV) and correlations (Cor) obtained by solving the normal equations without restrictions and the EM algorithm

	π_1			π_2			Cor
	E	SE	CV	E	SE	CV	
a. Normal equations							
1	.050	.325	6.492	1.100	.248	.225	-.829
2	.450	.335	.745	.600	.273	.455	-.817
3	5.700	4.210	.739	-.300	.990	-3.301	-.995
4	.731	.510	.698	.038	.692	17.991	-.944
5	.535	.195	.364	.823	.243	.295	-.591
6	-.150	.548	-3.654	1.050	.390	.371	-.943
7	1.291	.821	.636	.582	.312	.537	-.919
8	.071	.286	4.009	.786	.202	.258	-.707
9	.115	.188	1.633	.577	.215	.373	-.605
10	.889	.355	.400	.222	.367	1.650	-.800
11	-.556	.251	-.452	1.444	.296	.205	-.811
b. EM algorithm							
1	0.157	0.225	1.437	1.000	0.117	0.117	-0.544
2	0.450	0.229	0.509	0.600	0.122	0.203	-0.535
3	1.000	0.912	0.912	0.755	0.165	0.219	-0.831
4	0.692	0.195	0.283	0.093	0.141	1.510	-0.500
5	0.535	0.170	0.317	0.823	0.134	0.163	-0.376
6	0.026	0.235	8.963	0.928	0.110	0.118	-0.615
7	0.609	0.209	0.343	0.790	0.100	0.127	-0.485
8	0.072	0.234	3.266	0.786	0.117	0.149	-0.500
9	0.115	0.163	1.413	0.577	0.121	0.209	-0.392
10	0.889	0.245	0.275	0.222	0.155	0.697	-0.491
11	0.000	0.292	9.000	0.995	0.173	0.174	-0.507

NOTE: Using the EM algorithm, the overall estimates for π_1 and π_2 are respectively .279 (.067) and .798 (.037); see Appendix A for the EM algorithm.

Table 2. Comparison of the Bayesian method and Expectation Maximization algorithm (EM) algorithm using posterior means (PM), posterior standard deviations (PM), posterior coefficient of variations (PCV) and correlations (Cor)

	π_1			π_2			Cor
	PM	PSD	PCV	PM	PSD	PCV	
a. Bayesian method							
1	0.365	0.195	0.536	0.810	0.140	0.172	-0.582
2	0.483	0.236	0.488	0.561	0.205	0.365	-0.714
3	0.554	0.282	0.510	0.810	0.113	0.140	-0.418
4	0.475	0.219	0.462	0.417	0.264	0.633	-0.737
5	0.563	0.158	0.280	0.708	0.187	0.264	-0.450
6	0.383	0.255	0.666	0.669	0.197	0.294	-0.799
7	0.596	0.199	0.334	0.752	0.135	0.179	-0.585
8	0.270	0.186	0.691	0.671	0.161	0.239	-0.538
9	0.217	0.145	0.670	0.522	0.174	0.333	-0.452
10	0.671	0.219	0.327	0.402	0.229	0.569	-0.617
11	0.182	0.155	0.848	0.747	0.180	0.241	-0.416
12	0.279	0.082	0.295	0.793	0.070	0.089	-0.715
b. EM algorithm							
1	0.157	0.225	1.437	1.000	0.117	0.117	-0.544
2	0.450	0.229	0.509	0.600	0.122	0.203	-0.535
3	1.000	0.912	0.912	0.755	0.165	0.219	-0.831
4	0.692	0.195	0.283	0.093	0.141	1.510	-0.500
5	0.535	0.170	0.317	0.823	0.134	0.163	-0.376
6	0.026	0.235	8.963	0.928	0.110	0.118	-0.615
7	0.609	0.209	0.343	0.790	0.100	0.127	-0.485
8	0.072	0.234	3.266	0.786	0.117	0.149	-0.500
9	0.115	0.163	1.413	0.577	0.121	0.209	-0.392
10	0.889	0.245	0.275	0.222	0.155	0.697	-0.491
11	0.000	0.292	9.000	0.995	0.173	0.174	-0.507
12	0.274	0.067	0.244	0.798	0.037	0.047	-0.485

NOTE: Under the EM algorithm, the PM, PSD, PCV refer to estimate, standard error and coefficient of variation; see Appendix A for the EM algorithm. Group 12 means all data combined.

Table 3. Comparison of the two Bayesian methods, one based on the Gibbs sampler and the other based on numerical integration, using posterior means (PM), posterior standard deviations (PM), posterior coefficient of variations (PCV) and correlations (Cor)

	π_1			π_2			Cor
	PM	PSD	PCV	PM	PSD	PCV	
a. Gibbs sampler							
1	0.353	0.198	0.562	0.810	0.138	0.171	-0.604
2	0.482	0.239	0.495	0.567	0.209	0.369	-0.730
3	0.555	0.279	0.502	0.808	0.110	0.136	-0.395
4	0.479	0.225	0.469	0.411	0.260	0.633	-0.736
5	0.568	0.163	0.286	0.710	0.187	0.264	-0.469
6	0.387	0.249	0.643	0.661	0.195	0.295	-0.785
7	0.601	0.201	0.335	0.753	0.134	0.178	-0.532
8	0.265	0.186	0.700	0.679	0.157	0.232	-0.554
9	0.216	0.146	0.677	0.530	0.173	0.327	-0.447
10	0.667	0.222	0.333	0.427	0.233	0.545	-0.612
11	0.191	0.152	0.796	0.736	0.180	0.245	-0.414
12	0.279	0.082	0.295	0.793	0.070	0.089	-0.715
b. Numerical integration							
1	0.354	0.197	0.556	0.811	0.143	0.177	-0.613
2	0.477	0.237	0.497	0.573	0.201	0.351	-0.729
3	0.559	0.276	0.493	0.808	0.111	0.137	-0.384
4	0.485	0.220	0.454	0.416	0.266	0.640	-0.722
5	0.563	0.164	0.292	0.716	0.191	0.267	-0.463
6	0.370	0.251	0.679	0.685	0.189	0.276	-0.807
7	0.602	0.194	0.322	0.754	0.132	0.175	-0.577
8	0.270	0.191	0.708	0.674	0.157	0.233	-0.531
9	0.222	0.150	0.675	0.522	0.179	0.343	-0.472
10	0.659	0.221	0.335	0.427	0.228	0.534	-0.587
11	0.184	0.151	0.819	0.749	0.169	0.226	-0.383
12	0.277	0.080	0.287	0.792	0.069	0.087	-0.717

NOTE: Group 12 means all data combined.

Table 4. Frequentist properties of the Bayesian method versus the EM algorithm using bias, absolute relative bias, posterior standard deviation, root mean squared error, interval widths and coverage for π_1

Group	<i>B</i>	ARB	PSD	PRMSE	Wid	<i>C</i>
<u>a. Bayesian method</u>						
1	.037 _{.004}	.334 _{.008}	.158 _{.000}	.207 _{.002}	.583 _{.003}	.947 _{.003}
2	.017 _{.004}	.298 _{.007}	.130 _{.000}	.177 _{.002}	.493 _{.002}	.927 _{.002}
3	.021 _{.004}	.285 _{.007}	.131 _{.000}	.175 _{.002}	.496 _{.002}	.929 _{.002}
4	.034 _{.004}	.332 _{.008}	.157 _{.001}	.207 _{.002}	.581 _{.003}	.934 _{.003}
5	.031 _{.005}	.333 _{.008}	.157 _{.001}	.208 _{.002}	.578 _{.003}	.934 _{.003}
6	.036 _{.005}	.335 _{.008}	.157 _{.001}	.208 _{.002}	.581 _{.003}	.939 _{.003}
7	.045 _{.004}	.337 _{.008}	.159 _{.001}	.209 _{.002}	.588 _{.003}	.950 _{.003}
8	.026 _{.004}	.303 _{.007}	.131 _{.000}	.179 _{.002}	.494 _{.002}	.918 _{.002}
9	.038 _{.004}	.338 _{.008}	.157 _{.001}	.209 _{.002}	.583 _{.003}	.945 _{.003}
10	.021 _{.004}	.282 _{.007}	.131 _{.000}	.175 _{.002}	.497 _{.002}	.944 _{.002}
11	.004 _{.002}	.118 _{.003}	.052 _{.000}	.070 _{.001}	.201 _{.000}	.938 _{.008}
<u>b. EM algorithm</u>						
1	.006 _{.006}	.416 _{.009}	.166 _{.000}	.234 _{.002}	.600 _{.003}	.915 _{.009}
2	-.002 _{.005}	.342 _{.008}	.127 _{.000}	.186 _{.002}	.480 _{.002}	.898 _{.010}
3	.002 _{.005}	.324 _{.008}	.127 _{.000}	.182 _{.002}	.484 _{.002}	.902 _{.009}
4	.002 _{.006}	.414 _{.010}	.167 _{.000}	.235 _{.002}	.598 _{.004}	.904 _{.009}
5	-.001 _{.006}	.421 _{.010}	.166 _{.000}	.237 _{.002}	.594 _{.004}	.894 _{.010}
6	.005 _{.006}	.418 _{.010}	.166 _{.000}	.236 _{.002}	.599 _{.004}	.905 _{.009}
7	.008 _{.006}	.423 _{.010}	.169 _{.000}	.239 _{.002}	.608 _{.004}	.918 _{.009}
8	.008 _{.005}	.342 _{.008}	.128 _{.000}	.187 _{.002}	.483 _{.002}	.885 _{.010}
9	.008 _{.006}	.417 _{.010}	.166 _{.000}	.236 _{.002}	.600 _{.004}	.907 _{.009}
10	.002 _{.005}	.323 _{.008}	.128 _{.000}	.182 _{.002}	.484 _{.002}	.900 _{.009}
11	.002 _{.002}	.119 _{.003}	.048 _{.000}	.067 _{.001}	.187 _{.000}	.916 _{.009}

NOTE: All entries are in (0,1) in magnitude and they are averages of the 1000 runs. The notation a_b means that a is the average and b is the standard error. Group 11 has ten runs combined. Under the EM algorithm, PSD refers to standard error.

Table 5. Frequentist properties of the Bayesian method versus the EM algorithm using bias, absolute relative bias, posterior standard deviation, root mean squared error, interval widths and coverage for π_2

Group	<i>B</i>	ARB	PSD	PRMSE	Wid	<i>C</i>
<u>a. Bayesian method</u>						
1	-.025 _{.004}	.160 _{.004}	.135 _{.000}	.181 _{.002}	.508 _{.002}	.932 _{.002}
2	-.020 _{.004}	.160 _{.004}	.134 _{.000}	.180 _{.002}	.507 _{.002}	.936 _{.002}
3	-.021 _{.004}	.156 _{.004}	.135 _{.000}	.179 _{.002}	.508 _{.002}	.941 _{.002}
4	-.026 _{.004}	.162 _{.004}	.135 _{.001}	.182 _{.002}	.508 _{.002}	.932 _{.002}
5	-.030 _{.004}	.161 _{.004}	.135 _{.000}	.182 _{.002}	.509 _{.002}	.928 _{.002}
6	-.026 _{.004}	.161 _{.004}	.135 _{.000}	.181 _{.001}	.510 _{.002}	.938 _{.002}
7	-.052 _{.005}	.200 _{.005}	.174 _{.001}	.231 _{.002}	.630 _{.003}	.926 _{.003}
8	-.022 _{.004}	.157 _{.004}	.135 _{.000}	.180 _{.002}	.508 _{.002}	.948 _{.002}
9	-.028 _{.004}	.156 _{.004}	.135 _{.000}	.180 _{.002}	.511 _{.002}	.937 _{.002}
10	-.016 _{.004}	.163 _{.004}	.134 _{.000}	.181 _{.002}	.505 _{.002}	.928 _{.002}
11	-.002 _{.002}	.061 _{.001}	.049 _{.000}	.067 _{.001}	.189 _{.000}	.945 _{.007}
<u>b. EM algorithm</u>						
1	-.002 _{.005}	.186 _{.004}	.100 _{.000}	.167 _{.002}	.391 _{.000}	.802 _{.013}
2	.001 _{.005}	.185 _{.004}	.090 _{.000}	.160 _{.002}	.351 _{.000}	.743 _{.014}
3	-.001 _{.005}	.180 _{.004}	.090 _{.000}	.157 _{.002}	.351 _{.000}	.769 _{.013}
4	-.002 _{.005}	.188 _{.004}	.100 _{.000}	.169 _{.002}	.392 _{.001}	.783 _{.013}
5	-.006 _{.005}	.186 _{.005}	.100 _{.000}	.168 _{.002}	.392 _{.000}	.784 _{.013}
6	-.002 _{.005}	.188 _{.004}	.100 _{.000}	.168 _{.002}	.392 _{.000}	.795 _{.013}
7	-.008 _{.007}	.274 _{.006}	.131 _{.000}	.235 _{.003}	.508 _{.001}	.720 _{.014}
8	-.002 _{.005}	.182 _{.004}	.090 _{.000}	.158 _{.002}	.351 _{.000}	.756 _{.014}
9	-.005 _{.005}	.180 _{.004}	.100 _{.000}	.165 _{.002}	.392 _{.001}	.800 _{.013}
10	.005 _{.005}	.190 _{.004}	.089 _{.000}	.162 _{.002}	.351 _{.000}	.726 _{.014}
11	-.001 _{.002}	.061 _{.001}	.031 _{.000}	.054 _{.001}	.123 _{.000}	.792 _{.013}

NOTE: All entries are in (0,1) in magnitude and they are averages of the 1000 runs. The notation a_b means that a is the average and b is the standard error. Group 11 has ten runs combined. Under the EM algorithm, PSD refers to standard error.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).