Journal of the Korean Data & Information Science Society
2016, **27**(1), 245–254

http://dx.doi.org/10.7465/jkdi.2016.27.1.245
한국데이터정보과학회지

A Bayesian model for two-way contingency tables with nonignorable nonresponse from small areas

Namkyo Woo[1] · Dal Ho Kim[2]

[12]Department of Statistics, Kyungpook National University

Received 28 September 2015, revised 4 November 2015, accepted 25 November 2015

Abstract

Many surveys provide categorical data and there may be one or more missing categories. We describe a nonignorable nonresponse model for the analysis of two-way contingency tables from small areas. There are both item and unit nonresponse. One approach to analyze these data is to construct several tables corresponding to missing categories. We describe a hierarchical Bayesian model to analyze two-way categorical data from different areas. This allows a "borrowing of strength" of the data from larger areas to improve the reliability in the estimates of the model parameters corresponding to the small areas. Also we use a nonignorable nonresponse model with Bayesian uncertainty analysis by placing priors in nonidentifiable parameters instead of a sensitivity analysis for nonidentifiable parameters. We use the griddy Gibbs sampler to fit our models and compute DIC and BPP for model diagnostics. We illustrate our method using data from NHANES III data on thirteen states to obtain the finite population proportions.

Keywords: Gibbs sampler, nonidentifiable, nonignorable nonresponse, pooling, projection, small areas, two-way table.

1. Introduction

In survey sampling, data may be constructed in two-way categorical tables with missing cells. We consider the problem of nonignorable nonresponse for two-way ($r \times c$) categorical tables, each obtained from a single small area. There are both item and unit nonresponse. Item nonresponse has one of categories missing and unit nonresponse has all categories are missing. We do not know how the missing data are appeared. The model for including some difference between item and unit nonresponse (*i.e.*, nonignorable missing data) may be preferred.

We make a distinction between ignorable and nonignorable nonresponse model. These are associated with the missing data mechanism (Little and Rubin, 2002). According to the probability of response, there are three types of missing data mechanism. Missing completely at random (MCAR) occurs if the missingness is independent of both the observed and the

[1] Ph.D candidate, Department of Statistics, Kyungpook National University, Daegu 702-701, Korea.
[2] Corresponding author: Professor, Department of Statistics, Kyungpook National University, Daegu 702-701, Korea. E-mail: dalkim@knu.ac.kr

unobserved data, and missing at random (MAR) when conditional on the observed data, missingness is independent of the unobserved data. Missing not at random (MNAR) is neither MCAR nor MAR. While under MCAR partially complete data meaningness, under MAR partially complete data are relevant. Models for MCAR and MAR are called ignorable, and models for MNAR missing data mechanism are called nonignorable (Rubin, 1976). Since the missing data are different from the observed data, the issue of MNAR is how to fill in nonresponse. However the general difficulty with nonignorable nonresponse models is that there are nonidentifiable parameters (Nandram and Choi, 2005, 2008, 2010).

In a nonignorable nonresponse model, a sensitivity analysis is necessary to study the effects of nonidentifiable parameters on the parameter of interest. Typically, the sensitivity analysis is performed by setting the nonidentifiable parameters at various plausible values. Rather than performing a sensitivity analysis, we put a prior on the nonidentifiable parameters, that is a Bayesian uncertainty analysis. For Bayesian uncertainty analysis there are two strategies which are projection and pooling. In projection we can project parameters to a lower dimensional space and in pooling we can allow the reduced set of parameters to share a common distribution. This passes on the nonidentifiable effect to a smaller set of hyper parameters. Thus, projection and pooling lead to a reduced set of nonidentifiability parameters.

In this paper, we consider the two-way categorical data obtained independently from several geographical areas and we develop a nonignorable nonresponse model which is used for pooling the data over many small areas with Bayesian uncertainty analysis. The plan of this paper is as follows. In Section 2, we describe the hierarchical Bayesian nonignorable nonresponse model with Bayesian uncertainty analysis. In Section 3, we perform model diagnostic. There are two goodness-of-fit procedures. In Section 4, we illustrate our methodology with public-use data from thirteen states in the third National Health and Nutrition Examination Survey (NHANES III). Section 5 has concluding remarks.

2. Bayesian uncertainty analysis

2.1. The nonignorable nonresponse model

For the problem of nonresponse in a two-way categorical table, we can have both item and unit nonresponse. Also, we may consider the full data array to consist of four tables. One for the complete data and others for incomplete data - one for missing row information, one for missing column, a table for which neither row nor column has been observed. In this paper, we index small area by $i = 1, \cdots, A$, rows by $j = 1, \cdots, r$; columns by $k = 1, \cdots, c$ and the four tables by $t = 1, 2, 3, 4$. We show the four tables in Table 2.1.

Table 2.1 Tables of observed and missing counts for a general table

Table	Observed	Missing
t1	x, y, z, n	None
t2	x	y, z, n
t3	y	x, z, n
t4	None	x, y, z, n

For a two-way categorical table, let $J_{its} = 1$ if the s^{th} individual in i^{th} area belongs to the t^{th} table and $J_{its} = 0$ for the other three tables, and let $I_{ijks} = 1$ if the s^{th} individual in i^{th} area belongs to the (j, k) of the two-way table and $I_{ijks} = 0$ for all other cells. Also let $w_{ijkts} = J_{its} I_{ijks}$.

Now, let p_{ijkt} be the probability that an individual belongs to cell (j,k) of t^{th} sub-table in two-way table for i^{th} area, and let π_{it} be the probability that an individual belongs to the t^{th} sub-table for i^{th} area. The parameters π_{it} are identifiable. However, the parameters p_{ijkt} are not identifiable for $t = 2, 3, 4$. Given the constraints and the observed data, inference for the p_{ijkt} are independent of the π_{it}. If the p_{ijkt} do not depend on t, then these parameters are identifiable. For this case it is the ignorable nonresponse model corresponding the MAR mechanism.

Our basic model is as follows.

$$\underset{\sim}{J}_{is}|\underset{\sim}{\pi}_i \overset{\text{iid}}{\sim} \text{Multinomial}(1, \underset{\sim}{\pi}_i)$$

$$\underset{\sim}{I}_{is}|J_{its} = 1, \underset{\sim}{p}_{it} \overset{\text{iid}}{\sim} \text{Multinomial}(1, \underset{\sim}{p}_{it})$$

where $\underset{\sim}{J}_{is} = (J_{i1s}, \cdots, J_{i4s})'$, $\underset{\sim}{I}_{is} = (I_{i11s}, \cdots, I_{ircs})'$, $\underset{\sim}{\pi}_i = (\pi_{i1}, \cdots, \pi_{i4})'$ and $\underset{\sim}{p}_{it} = (p_{i11t}, \cdots, p_{irct})'$. And let $\psi_{ijkt} = \pi_{it}p_{ijkt}$. Then because $\sum_t \pi_{it} = 1$ and $\sum_{jk} p_{ijkt} = 1$ for each $t = 1, 2, 3, 4$, $\sum_t \sum_{jk} \pi_{it}p_{ijkt} = 1$. It follows that

$$\underset{\sim}{w}_{is}|\underset{\sim}{p}_i, \underset{\sim}{\pi}_i \overset{\text{iid}}{\sim} \text{Multinomial}(1, \underset{\sim}{\psi}_i).$$

where $\underset{\sim}{w}_{is} = (w_{i111s}, \cdots, w_{irc4s})'$, $\underset{\sim}{p}_i = (p_{i111}, \cdot, p_{irc4})'$ and $\underset{\sim}{\psi}_i = (\psi_{i111}, \cdots, \psi_{irc4})'$.

We use a nonignorable nonresponse model with Bayesian uncertainty analysis for reducing the effects of nonidenfitiable parameters. Rather than varying these nonidentifiable parameters at specified plausible values as in a formal non-Bayesian sensitivity analysis, we do so in a coherent Bayesian manner. That is, treat these parameters as hyper parameters by placing a prior on them.

2.2. Projection and pooling

Bayesian uncertainty analysis includes two strategies which are projection and pooling (see Greenland, 2009 and Molenberghs *et al.*, 2001 for details). In first strategy we can project p_{ijkt} to a lower dimensional space. This can be done by expressing the p_{ijkt} as functions of a reduced set of parameters. In current work, we have n_{it} individuals in t^{th} table for i^{th} area and $n_i = \sum_{t=1}^{4} n_{it}$. Also, for the four tables the cell counts are $z_{it} = \sum_{s=1}^{n_{it}} w_{i11ts}$, $x_{it} = \sum_{s=1}^{n_{it}} \sum_{j=1}^{r} w_{ij1ts}$, $y_{it} = \sum_{s=1}^{n_{it}} \sum_{k=1}^{c} w_{i1kts}$ and the corresponding superpopulation proportions are $\theta_{it}, p_{it}, q_{it}$. We show the structure of data in contingency table and cell probabilities in Figure 2.1.

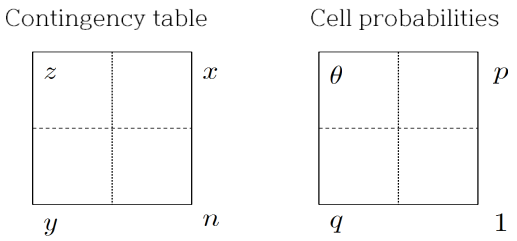

Figure 2.1 Contingency table and cell probabilities

In the next strategy, we can allow the reduced set of parameters to share a common distribution. This passes on the nonidentifiablity effect to a smaller set of hyper parameters. In our model, we use a Dirichlet prior density for categorical cell probabilities as follows

$$(\theta_{it}, p_{it}-\theta_{it}, q_{it}-\theta_{it}, 1-p_{it}-q_{it}+\theta_{it})|\mu_1,\mu_2,\mu_3,\tau \overset{\text{iid}}{\sim} \text{Dirichlet}(\mu_1\tau, (\mu_2-\mu_1)\tau, (\mu_3-\mu_1)\tau, (1-\mu_2-\mu_3+\mu_1)\tau),$$

where $0<\mu_1<\mu_2,\mu_3<1, \tau>0$.

It is worth nothing that if μ_1,μ_2,μ_3 and τ are specified, then the model will be well identified. Usually, sensitivity analysis is performed by taking various plausible values of these parameter. However, it is more sensible to perform a Bayesian uncertainty analysis. That is, treat these parameters as hyper parameters by placing a prior on them. This will permit a study of subjectivity to provide a coherent method to obtain an uncertainty interval for the finite population proportion. But it will not work completely, an adjustment is still needed. One way to do the adjustment is to put a bound on μ_1, say B, and take a proper diffuse prior for τ. That is, these parameters are constrained on the set $S=\{(\mu_1,\mu_2,\mu_3): 0<B\le\mu_1<\mu_2,\mu_3<1\}$. Two strategies lead to a reduced set of nonidentifiability parameters. The specification of priors on these parameters is the Bayesian uncertainty analysis. Henceforth, we will focus on the 2×2 table (*i.e.*, $r=c=2$).

The joint probability mass function of the nonignorable nonresponse model is

$$p(x_{it},y_{it},z_{it}|\theta_{it},p_{it},q_{it}) = \frac{n_{it}!\theta_{it}^{z_{it}}(p_{it}-\theta_{it})^{(x_{it}-z_{it})}(q_{it}-\theta_{it})^{(y_{it}-z_{it})}(1-p_{it}-q_{it}+\theta_{it})^{(n_{it}-x_{it}-y_{it}+z_{it})}}{z_{it}!(x_{it}-z_{it})!(y_{it}-z_{it})!(n_{it}-x_{it}-y_{it}+z_{it})!}$$

Now, hyper prior parameters μ_1,μ_2,μ_3 are constrained on the set $S=\{(\mu_1,\mu_2,\mu_3): 0<B\le\mu_1<\mu_2,\mu_3<1\}$. Also We take a proper diffuse prior for τ,

$$(\mu_1,\mu_2,\mu_3)|B\sim\text{Uniform}(S),\quad B\sim\text{Uniform}(a,b),\quad p(\tau)=\frac{1}{(1+\tau)^2},\quad \tau>0,$$

where a and b should be specified. For example, we can give a nonparametric interval of (a,b). A lower bound a assumes that only observed data including in cell (1, 1) and a upper bound b assumes that all missings are including in cell (1, 1) through all areas.

Let $\underset{\sim}{d}_{mis}$ and $\underset{\sim}{d}_{obs}$ denote all missing data and all observed data respectively. Then, the joint posterior density of all parameters and missing values is

$$\pi(\underset{\sim}{\theta},\underset{\sim}{p},\underset{\sim}{q},\underset{\sim}{d}_{mis},\mu_1,\mu_2,\mu_3,\tau|\underset{\sim}{d}_{obs}) \propto \frac{(\Gamma(\tau))^{4A}}{(1+\tau)^2}$$
$$\times\prod_{i=1}^{A}\prod_{t=1}^{4}\left\{\frac{n!(\theta_{it})^{z_{it}}(p_{it}-\theta_{it})^{(x_{it}-z_{it})}(q_{it}-\theta_{it})^{(y_{it}-z_{it})}(1-p_{it}-q_{it}+\theta_{it})^{(n_{it}-x_{it}-y_{it}+z_{it})}}{z_{it}!(x_{it}-z_{it})!(y_{it}-z_{it})!(n_{it}-x_{it}-y_{it}+z_{it})!}\right.$$
$$\left.\frac{\theta_{it}^{\mu_1\tau-1}(p_{it}-\theta_{it})^{(\mu_2-\mu_1)\tau-1}(q_{it}-\theta_{it})^{(\mu_3-\mu_1)\tau-1}(1-p_{it}-q_{it}+\theta_{it})^{(1-\mu_2-\mu_3+\mu_1)\tau-1}}{\Gamma(\mu_1\tau)\Gamma((\mu_2-\mu_1)\tau)\Gamma((\mu_3-\mu_1)\tau)\Gamma((1-\mu_2-\mu_3+\mu_1)\tau)}\right\}.$$

We use the griddy Gibbs sampler to draw samples from this joint posterior density. The joint conditional posterior distribution of the missing data have standard multinomial forms.

In the joint conditional posterior density, $(\theta_{it}, p_{it}, q_{it})$ are independent over t, and they have standard Dirichlet distributions given by

$$(\theta_{it}, p_{it}-\theta_{it}, q_{it}-\theta_{it}, 1-p_{it}-q_{it}+\theta_{it}|n_{it}, z_{it}, x_{it}, y_{it}, \mu_1, \mu_2, \mu_3, \tau) \overset{\text{ind}}{\sim}$$
$$\text{Dirichlet}(z_{it}+\mu_1\tau, x_{it}-z_{it}+(\mu_2-\mu_1)\tau, y_{it}-z_{it}+(\mu_3-\mu_1)\tau,$$
$$n_{it}-x_{it}-y_{it}+z_{it}+(1-\mu_2-\mu_3+\mu_1)\tau).$$

However, the joint posterior density of $(\mu_1, \mu_2, \mu_3, \tau)$ is not have in closed form. Each of them is obtained using a grid method. The joint posterior density is given by

$$(\mu_1, \mu_2, \mu_3, \tau|\underset{\sim}{r}, \underset{\sim}{\theta}, \underset{\sim}{p}, \underset{\sim}{q}) \propto \frac{\Gamma(\tau)^{4A}}{(1+\tau)^2}$$
$$\times \prod_{i=1}^{A}\prod_{t=1}^{4}\left\{\frac{\theta_{it}^{\mu_1\tau-1}(p_{it}-\theta_{it})^{(\mu_2-\mu_1)\tau-1}(q_{it}-\theta_{it})^{(\mu_3-\mu_1)\tau-1}(1-p_{it}-q_{it}+\theta_{it})^{(1-\mu_2-\mu_3+\mu_1)\tau-1}}{\Gamma(\mu_1\tau)\Gamma((\mu_2-\mu_1)\tau)\Gamma((\mu_3-\mu_1)\tau)\Gamma((1-\mu_2-\mu_3+\mu_1)\tau)}\right\},$$

where $0 < B \leq \mu_1 < \mu_2, \mu_3 < 1$ and $\tau > 0$.

Also, we need the conditional posterior densities for incomplete tables t2-t4 in each area. For table t2, x_{i2} is observed. Then,

$$z_{i2}|x_{i2}, \theta_{i2}, p_{i2} \sim \text{Binomial}\left(x_{i2}, \frac{\theta_{i2}}{p_{i2}}\right),$$
$$y_{i2}-z_{i2}|n_{i2}, x_{i2}, z_{i2}, \theta_{i2}, p_{i2}, q_{i2} \sim \text{Binomial}\left(n_{i2}-x_{i2}, \frac{q_{i2}-\theta_{i2}}{1-p_{i2}}\right),$$

For table t3, y_{i3} is observed. Then,

$$z_{i3}|y_{i3}, \theta_{i3}, q_{i3} \sim \text{Binomial}\left(y_{i3}, \frac{\theta_{i3}}{q_{i3}}\right),$$
$$x_{i3}-z_{i3}|n_{i3}, y_{i3}, z_{i3}, \theta_{i3}, p_{i3}, q_{i3} \sim \text{Binomial}\left(n_{i3}-y_{i3}, \frac{p_{i3}-\theta_{i3}}{1-q_{i3}}\right).$$

For table t4, all counts are missing. Then,

$$(z_{i4}, x_{i4}-z_{i4}, y_{i4}-z_{i4}, n_{i4}-x_{i4}-y_{i4}+z_{i4})'|n_{i4}, \theta_{i4}, p_{i4}, q_{i4} \sim$$
$$\text{Multinomial}(n_{i4}, (\theta_{i4}, p_{i4}-\theta_{i4}, q_{i4}-\theta_{i4}, 1-p_{i4}-q_{i4}+\theta_{i4})').$$

2.3. Inference for the finite population proportion

We assume that a random sample of size n_i is selected from each area of size N_i and the n_i selected individuals can be classified into a two-way table of counts.

Our target is the finite population proportion for the j^{th} row and the k^{th} column in $i^{th}area$, $P_{ijk}, j = k = 1, 2$. We show the way to inference P_{ijk} under the nonignorable nonresponse model. The idea is same under similar model (Rubin, Stern and Vehovar, 1995). And let N_{it} denote the total number of responding in the t^{th} table of $i^{th}area$. We are assuming that

there is no selection bias, and the sample is representative sample from population. Then using standard notation in survey sampling, we can write our target as

$$\begin{aligned} P_{i11} &= f_i\bar{z}_i + (1-f_i)\bar{Z}_i, \\ P_{i12} &= f_i(\bar{x}_i - \bar{z}_i) + (1-f_i)(\bar{X}_i - \bar{Z}_i), \\ P_{i21} &= f_i(\bar{y}_i - \bar{z}_i) + (1-f_i)(\bar{Y}_i - \bar{Z}_i), \\ P_{i22} &= f_i(\bar{n}_i - \bar{x}_i - \bar{y}_i + \bar{z}_i) + (1-f_i)(\bar{N}_i - \bar{X}_i - \bar{Y}_i + \bar{Z}_i), \end{aligned}$$

where $\bar{z}_i$, $\bar{x}_i - \bar{z}_i$, $\bar{y}_i - \bar{z}_i$ and $\bar{n}_i - \bar{x}_i - \bar{y}_i + \bar{z}_i$ are the sample proportions, $\bar{Z}_i$, $\bar{X}_i - \bar{Z}_i$, $\bar{Y}_i - \bar{Z}_i$ and $\bar{N}_i - \bar{X}_i - \bar{Y}_i + \bar{Z}_i$ are the nonsample proportions, and $f_i = n_i/N_i$ is the sampling fraction.

Note that both the sample proportion and the nonsample proportions are unobserved. Thus, given the sampled data, both of them are random variables. While the sample proportion is obtained directly from the model fitting, the nonsample proportion has to be predicted.

Now we show how to predict the nonsampled proportion. Let $\tilde{N}_i = N_i - n_i$ denote the number of nonsample individuals and $\tilde{N}_{it} = N_{it} - n_{it}$, let $\underset{\sim}{\tilde{N}} = (\tilde{N}_1, \tilde{N}_2, \tilde{N}_3, \tilde{N}_4)'$. Then under the nonignorable nonresponse model,

$$\begin{aligned} \underset{\sim}{\tilde{N}} | \underset{\sim}{\pi} &\sim \text{Multinomial}(N_i - n_i, \underset{\sim}{\pi}), \\ Z_{it}, X_{it} - Z_{it}, Y_{it} - Z_{it}, N_{it} - X_{it} - Y_{it} + Z_{it}, | \tilde{N}_{it_i}, \theta_{it}, p_{it}, q_{it} & \\ \overset{\text{ind}}{\sim} \text{Multinomial}(\tilde{N}_{it_i}, (\theta_{it}, p_{it} - \theta_{it}, q_{it} - \theta_{it}, 1 - p_{it} - q_{it} + \theta_{it})), & \end{aligned}$$

where $i = 1, \cdots, A$ and $t = 2, 3, 4$.

3. Model diagnostic

We perform two goodness-of-fit procedures, the deviance information criterion (DIC) and the Bayesian posterior predictive p-value (BPP). We can assess the overall fit of the models with these procedures. For the nonignorable nonresponse model,

$$p(\underset{\sim}{d} | \underset{\sim}{\theta}, \underset{\sim}{p}, \underset{\sim}{q}) = \prod_{i=1}^{A} \prod_{t=1}^{4} p(x_{it}, y_{it}, z_{it} | \theta_{it}, p_{it}, q_{it}),$$

where given the data $\underset{\sim}{d} = (\underset{\sim}{d}_{obs}, \underset{\sim}{d}_{mis})$.

Let $\theta_{it}^{(g)}, p_{it}^{(g)}, q_{it}^{(g)}$, $i = 1, \cdots, A$, $t = 1, 2, 3, 4$, $g = 1, \cdots, G$, denote the iterates from the griddy Gibbs sampler under the nonignorable nonresponse model and let the posterior mean of them be $\bar{\theta}_{it}, \bar{p}_{it}, \bar{q}_{it}$. For the nonignorable nonresponse model the deviance information criterion is given by

$$DIC = 2\bar{D} - D(\underset{\sim}{\bar{\theta}}, \underset{\sim}{\bar{p}}, \underset{\sim}{\bar{q}}),$$

where $\bar{D} = -2\sum_{g=1}^{G} log\{p(\underset{\sim}{d} | \underset{\sim}{\theta}^{(g)}, \underset{\sim}{p}^{(g)}, \underset{\sim}{q}^{(g)})\}/G$ and $D(\underset{\sim}{\bar{\theta}}, \underset{\sim}{\bar{p}}, \underset{\sim}{\bar{q}}) = -2log\{p(\underset{\sim}{d} | \underset{\sim}{\bar{\theta}}, \underset{\sim}{\bar{p}}, \underset{\sim}{\bar{q}})\}$.

Models with smaller DIC are more preferred over those with larger DIC. However, since DIC tends to select over-fitted models, thus, Yan and Sedransk (2007) described the Bayesian predictive p-values as a backup.

Let y_{ijkt} be cell counts in (j,k) cell of t^{th} table in i^{th} area and $\underset{\sim}{y}_{it}$ be the multinomial distribution with probabilities p_{ijkt}. Clearly, $E(y_{ijkt}|p_{ijkt}) = n_{it}p_{ijkt}$ and $Var(y_{ijkt}|p_{ijk}) = n_{it}p_{ijkt}(1-p_{ijkt})$. For the nonignorable nonresponse model, the discrepancy function is

$$T(\underset{\sim}{y};\underset{\sim}{p}) = \sum_{i=1}^{A}\sum_{t=1}^{4}\sum_{jk}\frac{(y_{ijkt} - E(y_{ijkt}|p_{ijkt}))^2}{Var(y_{ijkt}|p_{ijkt})}.$$

Then, we can obtain the respective Bayesian predictive p-values corresponding to the models, $p\{T(\underset{\sim}{y}^{(rep)};\underset{\sim}{p}) \geq T(\underset{\sim}{y}^{(obs)};\underset{\sim}{p})\}$. Here, these probabilities are calculated over their corresponding iterates $\underset{\sim}{p}^{(g)}$, $g = 1, \cdots, G$. If the value of this probability is close to 0.5, it indicates good fit of the model.

4. Data Analysis

We have data from the NHANES III. Variables which we use are family income divided by family size (FI) and the Bone mineral density (BMD) from thirteen states. FI is 0 if family income is less than \$20,000 which means the low level of income, 1 if it is greater than \$20,000 and 2 if it is missing. BMD is used to diagnose osteoporosis, a disease of elderly. Osteoporosis is BMD less than $0.64mg/cm^2$. Therefore BMD category is 0 if BMD is greater than $0.64mg/cm^2$, 1 if osteoporosis and 2 if missing. We present the categorical counts of full data for the area in Table 4.1.

Table 4.1 Classification of bone mineral density (BMD) and family income (FI) by states

		State												
FI	BMD	4	6	12	17	25	26	29	36	39	42	44	48	53
0	0	29	349	55	33	15	24	12	74	10	26	20	210	24
0	1	2	32	15	0	2	2	3	9	3	8	5	9	3
0	2	11	178	18	18	9	10	9	43	9	35	19	88	9
1	0	33	316	47	36	33	43	20	75	29	31	20	90	49
1	1	2	12	5	1	1	2	3	7	1	3	1	8	2
1	2	11	136	27	26	18	16	11	63	12	39	16	35	12
2	0	4	93	11	19	5	7	2	16	5	4	8	24	6
2	1	1	4	2	2	5	0	4	6	0	0	0	2	1
2	2	1	20	1	4	1	1	1	1	0	0	1	14	0

We use the griddy Gibbs sampler to fit models to the NHANES III data. We fit the Nonignorable nonresponse model with Bayesian uncertainty analysis. We iterated 11,000 times and first 1,000 used as a 'burn-in' and we took each iterate thereafter. We used the trace plots and the autocorrelations for checking the quality of the sample. Therefore we found negligible autocorrelations among the iterates, and so it is good that 'thinning' is not needed.

We compare the ignorable nonresponse model and the nonignorable nonresponse model. The results for posterior means of the finite population proportion in four cells of thirteen areas show in Table 4.2-4.5. For numerical summaries we use the posterior mean (PM), posterior standard deviation (PSD) and 95% credible interval (CI). PMs are similar between models. However, almost PSDs with the nonignorable nonresponse model is smaller than

with the ignorable nonresponse model. Also, we perform two model diagnostic procedures, DIC and BPP. The result of diagnostic statistics are shown in Table 4.6. DIC of the non-ignorable nonresponse model is lower than the ignorable nonresponse model. Also the BPP of model is not close to 0 or 1. That is, the nonignorable nonresponse model is significantly better than the ignorable nonresponse model.

Table 4.2 Comparison of the PM, PSD and 95% CI for for P_{i11}

	Ignorable			Nonignorable		
State	PM	PSD	95%CI	PM	PSD	95%CI
4	0.4285	0.0442	(0.3436, 0.5154)	0.4239	0.0406	(0.3452, 0.5043)
6	0.4982	0.0156	(0.4672, 0.5284)	0.4788	0.0218	(0.4344, 0.5186)
12	0.4146	0.0343	(0.3478, 0.4823)	0.4205	0.0321	(0.3566, 0.4834)
17	0.4283	0.0404	(0.3496, 0.5076)	0.4194	0.0363	(0.3475, 0.4899)
25	0.3137	0.0446	(0.2281, 0.4017)	0.3425	0.0403	(0.2646, 0.4213)
26	0.3606	0.0424	(0.2786, 0.4448)	0.3666	0.0386	(0.2910, 0.4424)
29	0.3451	0.0511	(0.2477, 0.4454)	0.3694	0.0445	(0.2823, 0.4569)
36	0.4049	0.0287	(0.3485, 0.4609)	0.4076	0.0287	(0.3497, 0.4624)
39	0.3246	0.0499	(0.2283, 0.4217)	0.3390	0.0435	(0.2536, 0.4246)
42	0.3883	0.0397	(0.3103, 0.4675)	0.3990	0.0386	(0.3226, 0.4740)
44	0.4341	0.0471	(0.3428, 0.5272)	0.4296	0.0423	(0.3467, 0.5128)
48	0.6442	0.0233	(0.5972, 0.6889)	0.6036	0.0268	(0.5486, 0.6535)
53	0.3430	0.0415	(0.2632, 0.4255)	0.3511	0.0374	(0.2778, 0.4245)

Table 4.3 Comparison of the PM, PSD and 95% CI for P_{i12}

	Ignorable			Nonignorable		
State	PM	PSD	95%CI	PM	PSD	95%CI
4	0.4868	0.0441	(0.4016, 0.5734)	0.4854	0.0405	(0.4037, 0.5638)
6	0.4395	0.0151	(0.4100, 0.4690)	0.4404	0.0182	(0.4036, 0.4746)
12	0.4380	0.0352	(0.3699, 0.5061)	0.4373	0.0328	(0.3732, 0.5008)
17	0.5155	0.0411	(0.4353, 0.5964)	0.5024	0.0373	(0.4281, 0.5748)
25	0.5649	0.0467	(0.4736, 0.6556)	0.5499	0.0420	(0.4663, 0.6320)
26	0.5689	0.0439	(0.4824, 0.6543)	0.5481	0.0397	(0.4714, 0.6262)
29	0.4815	0.0514	(0.3785, 0.5800)	0.4898	0.0445	(0.4023, 0.5777)
36	0.4801	0.0292	(0.4235, 0.5383)	0.4816	0.0296	(0.4214, 0.5381)
39	0.5785	0.0511	(0.4790, 0.6783)	0.5548	0.0444	(0.4681, 0.6428)
42	0.4750	0.0382	(0.3993, 0.5500)	0.4708	0.0384	(0.3928, 0.5432)
44	0.4541	0.0458	(0.3644, 0.5444)	0.4514	0.0416	(0.3694, 0.5322)
48	0.2950	0.0217	(0.2540, 0.3391)	0.3170	0.0219	(0.2745, 0.3603)
53	0.5733	0.0430	(0.4892, 0.6580)	0.5598	0.0392	(0.4830, 0.6373)

Table 4.4 Comparison of the PM, PSD and 95% CI for P_{i21}

	Ignorable			Nonignorable		
State	PM	PSD	95%CI	PM	PSD	95%CI
4	0.0472	0.0207	(0.0149, 0.0952)	0.0518	0.0198	(0.0202, 0.0973)
6	0.0451	0.0072	(0.0318, 0.0605)	0.0560	0.0159	(0.0319, 0.0924)
12	0.1026	0.0225	(0.0630, 0.1511)	0.0970	0.0208	(0.0610, 0.1412)
17	0.0260	0.0155	(0.0032, 0.0612)	0.0424	0.0179	(0.0137, 0.0831)
25	0.0730	0.0286	(0.0258, 0.1354)	0.0634	0.0217	(0.0275, 0.1118)
26	0.0384	0.0180	(0.0110, 0.0810)	0.0473	0.0180	(0.0181, 0.0886)
29	0.0948	0.0352	(0.0377, 0.1769)	0.0790	0.0261	(0.0354, 0.1369)
36	0.0616	0.0162	(0.0340, 0.0973)	0.0622	0.0186	(0.0325, 0.1053)
39	0.0659	0.0270	(0.0239, 0.1283)	0.0682	0.0244	(0.0290, 0.1232)
42	0.0938	0.0265	(0.0500, 0.1521)	0.0840	0.0277	(0.0411, 0.1483)
44	0.0833	0.0289	(0.0367, 0.1483)	0.0823	0.0272	(0.0378, 0.1428)
48	0.0339	0.0095	(0.0175, 0.0547)	0.0505	0.0175	(0.0232, 0.0915)
53	0.0506	0.0199	(0.0184, 0.0953)	0.0529	0.0178	(0.0236, 0.0934)

Table 4.5 Comparison of the PM, PSD and 95% CI for P_{i22}

	Ignorable			Nonignorable		
State	PM	PSD	95%CI	PM	PSD	95%CI
4	0.0376	0.0187	(0.0101, 0.0824)	0.0389	0.0168	(0.0133, 0.0782)
6	0.0173	0.0045	(0.0095, 0.0269)	0.0249	0.0096	(0.0114, 0.0480)
12	0.0448	0.0165	(0.0185, 0.0823)	0.0452	0.0162	(0.0199, 0.0829)
17	0.0302	0.0164	(0.0061, 0.0687)	0.0358	0.0171	(0.0104, 0.0766)
25	0.0483	0.0256	(0.0096, 0.1067)	0.0442	0.0205	(0.0135, 0.0927)
26	0.0321	0.0161	(0.0081, 0.0695)	0.0380	0.0170	(0.0124, 0.0776)
29	0.0786	0.0329	(0.0277, 0.1546)	0.0618	0.0237	(0.0254, 0.1177)
36	0.0534	0.0164	(0.0260, 0.0898)	0.0485	0.0183	(0.0213, 0.0923)
39	0.0311	0.0187	(0.0051, 0.0768)	0.0381	0.0192	(0.0109, 0.0841)
42	0.0429	0.0188	(0.0137, 0.0860)	0.0462	0.0230	(0.0144, 0.1017)
44	0.0285	0.0170	(0.0050, 0.0700)	0.0367	0.0190	(0.0094, 0.0811)
48	0.0270	0.0082	(0.0132, 0.0453)	0.0289	0.0090	(0.0144, 0.0493)
53	0.0331	0.0162	(0.0085, 0.0708)	0.0363	0.0155	(0.0118, 0.0722)

Table 4.6 Model diagnostic statistics

Model	DIC	BPP
Ignorable Nonresponse Model	557.0223	0.2064
Nonignorable Nonresponse Model	550.5476	0.5257

5. Concluding remarks

The purpose of this paper has been to develop a methodology to analyze data from incomplete two-way categorical tables from small areas. We have constructed the nonignorable nonresponse model with a reduced set of nonidentifiable parameters, each of the three incomplete tables has a set of parameters. We allowed these parameters to share a common effect, thereby passing on the nonidentifiability effects to a manageable set of parameters. For a Bayesian uncertainty analysis we set artificial priors on these hyper parameters. This allows a study of subjectivity in uncertainty for the finite population proportion is obtained.

We have shown that there are differences between ignorable nonresponse model and nonignorable nonresponse model. Using the data on FI and BMD from NHANES III, we have used the griddy Gibbs sampler to fit the model. Also we perform the model diagnostic with two procedures which are DIC and BPP. Finally, we observe that posterior inference about an illustrative example on estimating the finite population proportion for the 2×2 table over thirteen subnational areas.

References

Choi, S. M. and Kim, D. H. (2013). Bayesian estimation for finite population proportion under selection bias via surrogate samples. *Journal of the Korean Data & Information Science Society*, **24**, 1543-1550.

Choi, S. M. and Kim, D. H. (2014). Sensitivity analysis in Bayesian nonignorable selection model for binary responses. *Journal of the Korean Data & Information Science Society*, **25**, 187-194.

Greenland, S. (2009). Relaxation penalties and priors for plausible modeling of nonidentified bias sources, *Statistical Sciences*, **24**, 195-210.

Little, R. J. A. and Rubin D. B. (2002). *Statistical analysis with missing data.* Wiley, New York.

Molenberghs, G., Kenward, M. G. and Goetghebeur, E. (2001). Sensitivity analysis for incomplete contingency table: The Slovenian plebiscite case. *Journal of the American Statistical Association*, **97**, 381-388.

Nandram, B. and Choi, J. W. (2005). Hierarchical Bayesian nonignorable nonresponse regression models for small areas: An application to the NHANES data. *Survey Methodology*, **31**, 73-84.

Nandram, B. and Choi, J. W. (2008). A Bayesian allocation of undecided voters. *Survey Methodology*, **34**, 37-49.

Nandram, B. and 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.

National Center for Health Statistics. (1992). Third national health and nutrition examination survey. *Vital and Health Statistics Series 2*, **113**.

Rubin, D. B., Stern, H. S. and Behovar, V. (1995). Handling "Don't Know" survey responses: The case of the Slovenian plebiscite. *Journal of the American Statistical Association*, **72**, 822-828.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, **63**, 581-592.

Yan, G. and Sedransk, J. (2007). Bayesian diagnostic techniques for detecting hierarchical structure. *Bayesian Analysis*, **2**, 735-760.