Journal of the Korean Data & Information Science Society
2015, **26**(6), 1547–1555

http://dx.doi.org/10.7465/jkdi.2015.26.6.1547
한국데이터정보과학회지

A Bayesian uncertainty analysis for nonignorable nonresponse in two-way contingency table

Namkyo Woo[1] · Dal Ho Kim[2]

[12]Department of Statistics, Kyungpook National University

Received 19 August 2015, revised 15 October 2015, accepted 23 October 2015

Abstract

We study the problem of nonignorable nonresponse in a two-way contingency table and there may be one or two missing categories. We describe a nonignorable nonresponse model for the analysis of two-way categorical table. One approach to analyze these data is to construct several tables (one complete and the others incomplete). There are nonidentifiable parameters in incomplete tables. We describe a hierarchical Bayesian model to analyze two-way categorical data. We use a nonignorable nonresponse model with Bayesian uncertainty analysis by placing priors in nonidentifiable parameters instead of a sensitivity analysis for nonidentifiable parameters. To reduce the effects of nonidentifiable parameters, we project the parameters to a lower dimensional space and we allow the reduced set of parameters to share a common distribution. 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 to obtain the finite population proportions.

Keywords: Griddy Gibbs sampler, missingness, nonidentifiable parameter, nonignorable nonresponse, two-way table, uncertainty analysis.

1. Introduction

In survey sampling, data may be summarized in two-way contingency table with missing cells. We consider the problem of nonignorable nonresponse for two-way ($r \times c$) categorical table. There are both item and unit nonresponses. 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. In this situation, we would like to express a degree of uncertainty about ignorability (Nandram and Choi, 2002). The model that includes some difference between them (*i.e.*, nonignorable missing data) may be preferred.

We make a distinction between ignorable and nonignorable nonresponse models. These are associated with the missing data mechanism; see 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 meaningless, 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 nonresponses. However the general difficulty with nonignorable nonresponse models is that there are nonidentifiable parameters (Nandram and Choi, 2008).

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. This is called a Bayesian uncertainty analysis. This passes on the nonidentifiablity effect to a smaller set of hyper parameters. Thus, a Bayesian uncertainty analysis leads to a reduced set of nonidentifiability parameters. The work presented in this paper is closely related to Choi and Kim (2013, 2014).

In this paper, we consider the categorical data in two-way contingency table and we develop a nonignorable nonresponse model with Bayesian uncertainty analysis. The outline of the 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 in the third National Health and Nutrition Examination Survey (NHANES III) conducted by National Center for Health Statistics (1992). 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. Nandram et al. (2005) considered a problem in which an analysis is needed for categorical data from a single two-way table with missing. That is, 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 (Nandram, 2009). In this paper, we index rows by $j = 1, \cdots, r$; columns by $k = 1, \cdots, c$ and the four tables by $t = 1, 2, 3, 4$.

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

Now, let p_{jkt} be the probability that an individual belongs to cell (j, k) of t^{th} sub-table in two-way table, and let π_t be the probability that an individual belongs to the t^{th} sub-table. The parameters π_t are identifiable. However, the parameters p_{jkt} are not identifiable for incomplete tables, $t = 2, 3, 4$. If the p_{jkt} 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.

$$\underset{\sim}{J_s} \mid \underset{\sim}{\pi} \overset{\text{iid}}{\sim} \text{Multinomial}(1,\ \underset{\sim}{\pi});$$

$$\underset{\sim}{I_s} \mid J_{ts} = 1,\ \underset{\sim}{p_t} \overset{\text{iid}}{\sim} \text{Multinomial}(1,\ \underset{\sim}{p_t}).$$

Let $\psi_{jkt} = \pi_t p_{jkt}$. Then because $\sum_t \pi_t = 1$ and $\sum_{jk} p_{jkt} = 1$ for each $t = 1, \cdots, 4$, $\sum_t \sum_{jk} \pi_t p_{jkt} = 1$. Let $w_{jkts} = J_{ts} I_{jks}$. It follows that

$$\underset{\sim}{w_s} \mid \underset{\sim}{p},\ \underset{\sim}{\pi} \overset{\text{iid}}{\sim} \text{Multinomial}(1,\ \underset{\sim}{\psi}).$$

While the parameters π_t are identifiable, the parameters p_{jkt} are not identifiable for $t = 2, 3, 4$. Note that given the constraints and the observed data, inference for the p_{jkt} are independent of the π_t.

We use a parsimonious nonignorable nonresponse model and we use 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. Hyper parameters are constrained on a set. Thus, we take a proper diffuse prior.

2.2. Projection and pooling

Bayesian uncertainty analysis includes two strategies which are projection and pooling. In first strategy we can project p_{jkt} to a lower dimensional space. This can be done by expressing the p_{jkt} as functions of a reduced set of parameters. In current work, we have n_t individuals in t^{th} table. For the four tables the cell counts are $z_t = \sum_{s=1}^{n_t} w_{11ts}$, $x_t = \sum_{s=1}^{n_t} \sum_{j=1}^{r} w_{j1ts}$, $y_t = \sum_{s=1}^{n_t} \sum_{k=1}^{c} w_{1kts}$ and the corresponding superpopulation proportions are θ_t, p_t and q_t. 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

$$(\theta_t,\ p_t - \theta_t,\ q_t - \theta_t,\ 1 - p_t - q_t + \theta_t) \mid \mu_1,\ \mu_2,\ \mu_3,\ \tau \overset{\text{iid}}{\sim}$$
$$\text{Dirichlet}(\mu_1\tau,\ (\mu_2 - \mu_1)\tau,\ (\mu_3 - \mu_1)\tau,\ (1 - \mu_2 - \mu_3 + \mu_1)\tau),$$

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

It is worth noting that if μ_1, μ_2, μ_3 and τ are specified, then the model will be well identified. A sensitivity analysis will proceed by taking various values of these parameters to study effects on the finite population proportion. It is more sensible for a Bayesian to perform a Bayesian uncertainty analysis. 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 $S = \{(\mu_1,\ \mu_2,\ \mu_3) : 0 < B \leq \mu_1 < \mu_2,\ \mu_3 < 1\}$.

Two strategies lead to a reduced set of nonidentifiable 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

$$p(w_t,\ y_t,\ z_t \mid \theta_t,\ p_t,\ q_t) = \frac{n_t!\theta^{z_t}(p_t-\theta_t)^{(x_t-z_t)}(q_t-\theta_t)^{(y_t-z_t)}(1-p_t-q_t+\theta_t)^{(n_t-x_t-y_t+z_t)}}{z_t!(x_t-z_t)!(y_t-z_t)!(n_y-x_t-y_y+z_t)!}.$$

Now, hyper prior parameters $\mu_1,\ \mu_2,\ \mu_3$ 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,

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

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 $(1,1)$.

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

$$\pi(\underset{\sim}{\theta},\ \underset{\sim}{p},\ \underset{\sim}{q},\ \underset{\sim}{d}_{mis},\ \mu_1,\ \mu_2,\ \mu_3,\ \tau \mid \underset{\sim}{d}_{obs}) \propto$$
$$\frac{(\Gamma(\tau))^4}{(1+\tau)^2} \times \prod_{t=1}^{4}\Bigg\{ \frac{n!(\theta_t)^{z_t}(p_t-\theta_t)^{(x_t-z_t)}(q_t-\theta_t)^{(y_t-z_t)}(1-p_t-q_t+\theta_t)^{(n_t-x_t-y_t+z_t)}}{z_t!(x_t-z_t)!(y_t-z_t)!(n_t-x_t-y_t+z_t)!}$$
$$\times \frac{\theta_t^{\mu_1\tau-1}(p_t-\theta_t)^{(\mu_2-\mu_1)\tau-1}(q_t-\theta_t)^{(\mu_3-\mu_1)\tau-1}(1-p_t-q_t+\theta_t)^{(1-\mu_2-\mu_3+\mu_1)\tau-1}}{\Gamma(\mu_1\tau)\Gamma((\mu_2-\mu_1)\tau)\Gamma((\mu_3-\mu_1)\tau)\Gamma((1-\mu_2-\mu_3+\mu_1)\tau)} \Bigg\}.$$

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, $(\theta_t,\ p_t,\ q_t)$ are independent over t, and they have standard Dirichlet distributions given by

$$(\theta_t,\ p_t-\theta_t,\ q_t-\theta_t,\ 1-p_t-q_t+\theta_t \mid n_t,\ z_t,\ x_t,\ y_t,\ \mu_1,\ \mu_2,\ \mu_3,\ \tau) \overset{\text{ind}}{\sim}$$
$$\text{Dirichlet}(z_t+\mu_1\tau,\ x_t-z_t+(\mu_2-\mu_1)\tau,\ y_t-z_t+(\mu_3-\mu_1)\tau,$$
$$n_t-x_t-y_t+z_t+(1-\mu_2-\mu_3+\mu_1)\tau)$$

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

$$(\mu_1,\ \mu_2,\ \mu_3,\ \tau \mid \underset{\sim}{r},\ \underset{\sim}{\theta},\ \underset{\sim}{p},\ \underset{\sim}{q}) \propto \frac{(\Gamma(\tau))^4}{(1+\tau)^2}$$
$$\times \prod_{t=1}^{4} \frac{\theta_t^{\mu_1\tau-1}(p_t-\theta_t)^{(\mu_2-\mu_1)\tau-1}(q_t-\theta_t)^{(\mu_3-\mu_1)\tau-1}(1-p_t-q_t+\theta_t)^{(1-\mu_2-\mu_3+\mu_1)\tau-1}}{\Gamma(\mu_1\tau)\Gamma((\mu_2-\mu_1)\tau)\Gamma((\mu_3-\mu_1)\tau)\Gamma((1-\mu_2-\mu_3+\mu_1)\tau)},$$

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

Also, we need the conditional posterior densities for tables t2-t4. For table t2, x_2 is observed. Then,

$$z_2 \mid x_2,\ \theta_2,\ p_2 \sim \text{Binomial}\left(x_2,\ \frac{\theta_2}{p_2}\right),$$
$$y_2 - z_2 \mid n_2,\ x_2,\ z_2, \theta_2,\ p_2,\ q_2 \sim \text{Binomial}\left(n_2 - x_2,\ \frac{q_2-\theta_2}{1-p_2}\right),$$

For table t3, y_3 is observed. Then,

$$z_3 \mid y_3,\ \theta_3,\ q_3 \sim \text{Binomial}\left(y_3,\ \frac{\theta_3}{q_3}\right),$$
$$x_3 - z_3 \mid n_3,\ y_3, z_3,\ \theta_3,\ p_3, q_3 \sim \text{Binomial}\left(n_3 - y_3,\ \frac{p_3 - \theta_3}{1 - q_3}\right).$$

For table t4, all counts are missing. Then,

$$(z_4,\ x_4 - z_4,\ y_4 - z_4,\ n_4 - x_4 - y_4 + z_4)' \mid n_4,\ \theta_4,\ p_4,\ q_4 \sim$$
$$\text{Multinomial}(n_4,\ (\theta_4,\ p_4 - \theta_4,\ q_4 - \theta_4,\ 1 - p_4 - q_4 + \theta_4)').$$

2.3. Inference for the finite population proportion

We assume that a random sample of size n is selected from a finite population of size N and the n 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, $P_{jk}, j = k = 1, 2$. And let N_t denote the total number of responding in the t^{th} table. 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

$$P_{11} = f\bar{z} + (1 - f)\bar{Z},$$
$$P_{12} = f(\bar{x} - \bar{z}) + (1 - f)(\bar{X} - \bar{Z}),$$
$$P_{21} = f(\bar{y} - \bar{z}) + (1 - f)(\bar{Y} - \bar{Z}),$$
$$P_{22} = f(\bar{n} - \bar{x}) - \bar{y} + \bar{z}) + (1 - f)(\bar{N} - \bar{X} - \bar{Y} + \bar{Z}),$$

where $\bar{z}$, $\bar{x} - \bar{z}$, $\bar{y} - \bar{z}$ and $\bar{n} - \bar{x} - \bar{y} + \bar{z}$ are the sample proportions, $\bar{Z}$, $\bar{X} - \bar{Z}$, $\bar{Y} - \bar{Z}$ and $\bar{N} - \bar{X} - \bar{Y} + \bar{Z}$ are the nonsample proportions, and $f = n/N$ 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 $\tilde{N} = N - n$ denote the number of nonsample individuals and $\tilde{N}_t = N_t - n_t$, let $\underset{\sim}{\tilde{N}} = (\tilde{N}_1, \cdots, \tilde{N}_4)'$. Then under the nonignorable nonresponse model,

$$\underset{\sim}{\tilde{N}} \mid \underset{\sim}{\pi} \sim \text{Multinomial}(N - n,\ \underset{\sim}{\pi}),$$
$$Z_t,\ X_t - Z_t,\ Y_t - Z_t,\ N_t - X_t - Y_t + Z_t \mid \tilde{N}_t,\ \theta_t,\ p_t,\ q_t$$
$$\overset{\text{ind}}{\sim} \text{Multinomial}(\tilde{N}_t, (\theta_t,\ p_t - \theta_t,\ q_t - \theta_t,\ 1 - p_t - q_t + \theta_t)).$$

3. Model diagnostic

We perform two goodness-of-fit procedures, the deviance information criterion (DIC) together with the complexity (PD) or effective number of parameters 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,

$$p(\underset{\sim}{d} \mid \underset{\sim}{\theta},\ \underset{\sim}{p},\ \underset{\sim}{q}) = \prod_{t=1}^{4} p(x_t,\ y_t,\ z_t|\theta_t,\ p_t,\ q_t),$$

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

Let $\theta_t^{(g)},\ p_t^{(g)},\ q_t^{(g)},\ t = 1, \cdots, 4,\ g = 1, \cdots, G$, denote the iterates from the griddy Gibbs sampler under the nonignorable nonresponse model and let the posterior mean of them be $\bar{\theta}_t, \bar{p}_t, \bar{q}_t$. 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 $\bar{D} = -2\sum_{g=1}^{G} log\{p(\underset{\sim}{d} \mid \underset{\sim}{\theta}^{(g)},\ \underset{\sim}{p}^{(g)},\ \underset{\sim}{q}^{(g)})\}/G$ and $D(\underset{\sim}{\bar{\theta}},\ \underset{\sim}{\bar{p}},\ \underset{\sim}{\bar{q}}) = -2log\{p(\underset{\sim}{d} \mid \underset{\sim}{\bar{\theta}},\ \underset{\sim}{\bar{p}},\ \underset{\sim}{\bar{q}})\}$.

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

Let y_{jkt} be cell counts in (j,k) cell of t^{th} table and $\underset{\sim}{y}_t$ be the multinomial distribution with probabilities p_{jkt}. Clearly, $E(y_{jkt} \mid p_{jkt}) = n_t p_{jkt}$ and $Var(y_{jkt} \mid p_{jk}) = n_t p_{jkt}(1 - p_{jkt})$. For the nonignorable nonresponse model, the discrepancy function is

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

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. Numerical results

4.1. Data analysis

We have data from NHANES III. We use two variables which are family income divided by family size (FI) and the Bone mineral density (BMD). FI is 1 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 $0.64 mg/cm^2$. Therefore BMD category is 0 if BMD is greater than $0.64 mg/cm^2$, 1 if osteoporosis and 2 if missing. We present the categorical counts of full data for the area in Table 4.1. The data from NHANES III is 5% sampling data, so $N = 20 \times n$.

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

BMD	FI			Total
	0	1	Missing	
0	881	822	204	1907
1	93	48	27	168
Missing	456	422	45	923
Total	1430	1292	276	2,998

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 drew 11,000 iterates 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 show in Table 4.2. For numerical summaries we use the posterior mean (PM), posterior standard deviation (PSD) and 95% credible interval (CI).

Table 4.2 Comparison of the posterior means (PM), posterior standard deviation (PSD) and 95% credible interval (CI) for P_{jk} from the ignorable and nonignorable nonresponse model

P_{jk}	Ignorable nonresponse model			Nonignorable nonresponse model		
	PM	PSD	CI	PM	PSD	CI
P_{11}	0.4725	0.0096	(0.4536, 0.4909)	0.4717	0.0232	(0.4241, 0.5165)
P_{12}	0.4463	0.0094	(0.4280, 0.4651)	0.4314	0.0263	(0.3566, 0.4686)
P_{21}	0.0533	0.0049	(0.0439, 0.0634)	0.0568	0.0182	(0.0315, 0.1015)
P_{22}	0.0279	0.0037	(0.0210, 0.0358)	0.0402	0.0221	(0.0161, 0.1065)

Also, we perform two model diagnostic procedures, the deviance information criterion (DIC) and Bayesian predictive p-values (BPP). The result of diagnostic statistics are shown in Table 4.3. Models with smaller DIC are more preferred. However DIC tends to select over-fitted models, we compute BPP as a backup. Then, if the BPP value of this probability is close to 0.5, it indicates good fit of the model.

DIC of the nonignorable 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 4.3 Model diagnostic ststistics

Model	DIC	BPP
Ignorable nonresponse model	84.5832	0.3679
Nonignorable nonresponse model	78.0339	0.4987

4.2. Simulation study

We perform a simulation study to further assess the performance between the ignorable nonresponse model and the nonignorable nonresponse model. We keep $r = c = 2$ and the sample size n in the original data. After fitting the nonignorable nonresponse model to NHANES III data, we obtained the posterior means $\hat{\mu}_1 = 0.4777$, $\hat{\mu}_2 = 0.8607$, $\hat{\mu}_3 = 0.5540$ and $\hat{\tau} = 32.8260$ for hyper parameters μ_1, μ_2, μ_3 and τ. Thus, we generate cell proportions for the nonignorable nonresponse model

$$(\theta_t,\ p_t - \theta_t,\ q_t - \theta_t,\ 1 - p_t - q_t + \theta_t) \mid \hat{\mu}_1,\ \hat{\mu}_2,\ \hat{\mu}_3,\ \hat{\tau} \overset{\text{iid}}{\sim}$$
$$\text{Dirichlet}(\hat{\mu}_1\hat{\tau},\ (\hat{\mu}_2 - \hat{\mu}_1)\hat{\tau},\ (\hat{\mu}_3 - \hat{\mu}_1)\hat{\tau},\ (1 - \hat{\mu}_2 - \hat{\mu}_3 + \hat{\mu}_1)\hat{\tau}).$$

With these values we generate the cell counts for the nonignorable nonresponse model by

drawing from

$$z_t,\ x_t - z_t,\ y_t - z_t,\ n_t - x_t - y_t + z_t \mid \theta_t,\ p_t,\ q_t \overset{\text{ind}}{\sim} \text{Multinomial}(n,\ (\theta_t,\ p_t - \theta_t,\ q_t - \theta_t,\ 1 - p_t - q_t + \theta_t)).$$

We repeat this procedure to get 1,000 data sets. Then we fit each of these data sets using both the ignorable nonresponse model and the nonignorable nonresponse model. To confirm differences of the goodness-of-fit between models, we calculate $\text{DIC}^{(h)}$ and $\text{BPP}^{(h)}$ corresponding to the h^{th} dataset, $h = 1, \cdots, 1,000$. Table 4.4 shows comparison of the ignorable nonresponse model and the nonignorable nonresponse model. DIC under the nonignorable nonresponse model are smaller than those under the ignorable nonresponse model. And BPP under the nonignorable nonresponse model is close to 0.5. Based on these measures, the nonignorable nonresponse model is preferred over the ignorable nonresponse model.

Table 4.4 Comparison of models via simulated data

Model	DIC		BPP	
	AVG	SE	AVG	SE
Ignorable nonresponse model	133.5551	48.5675	0.1016	0.1750
Nonignorable nonresponse model	86.3267	4.3396	0.4982	0.0153

5. Concluding remarks

The purpose of this paper has been to develop a methodology to analyze data from incomplete two-way categorical table. 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 the ignorable nonresponse model and the nonignorable nonresponse model. Using the data on FI and BMD from NHANES III, we have used the griddy Gibbs sampler to fit the model. We have compared our model with the ignorable nonresponse model and the nonignorable nonresponse models. Also we perform the model diagnostic with two procedures which are DIC and BPP. Therefore, the nonignorable nonresponse model is more preferred. Finally, our simulation study supports that the data from the nonignorable nonresponse model can be suitable for the nonignorable nonresponse model.

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.

Little, R. J. A. and Rubin, D. B. (2002). *Statistical analysis with missing data*, 2nd Ed., John Wiley & Sons, New York.

Nandram, B. (2009). Bayesian inference of the cell probabilities of a two-way categorical table under nonignorability. *Communications in Statistics - Theory and Methods*, **38**, 3015-3030.

Nandram, B. and Choi, J. W. (2002). A Bayesian analysis of a proportion under nonignorable nonresponse. *Statistics in Medicine*, **21**, 1189-1212.

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

Nandram, B., Cox, L. and Choi, J. W. (2005). Bayesian analysis of nonignorable missing categorical data: An application to bone mineral density and family income. *Survey Methodology*, **31**, 213-225.

National Center for Health Statistics. (1992). Third National Health and Nutrition Examination Survey. *Vital and Health Statistics Series 2*, **113**.

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.