A Hierarchical Bayesian Multinomial Model with Nonresponse

Geunshik Han, Hanshin University/National Center for Health Statistics
Balgobin Nandram, Worcester Polytechnic Institute
Jai Won Choi, National Center for Health Statistics
Geunshik Han, Dept. of Computer Science and Statistics, 411 Yangsan-dong Osan, Korea

Key Words: Latent Variable; Metropolis-Hastings sampler; Nonignorable nonresponse.

Abstract: We describe a hierarchical Bayesian model to analyze multinomial nonignorable nonresponse data from small areas. We use Dirichlet prior on the multinomial probabilities and beta prior on the response probabilities which permit a pooling of the data from different areas. This pooling is needed because of the weak identifiability of the parameters in the model. Inference is sampling based and Markov chain Monte Carlo methods are used to perform the computations. We apply our method to study body mass index (BMI) data from the third National Health and Nutrition Examination Survey and show that the model works reasonably well.

1. Introduction

The nonresponse rates in many surveys have been increasing steadily (De Heer 1999 and Groves and Couper 1998), making the nonresponse problem more important. For many surveys the responses are polychotomous. For example, the third National Health and Nutrition Examination Survey (NHANES) estimates the proportions of persons belonging to the multilevel of body mass index (BMI).

We consider a hierarchical Bayes model to study nonignorable multinomial nonresponse. Rubin (1987) and Little and Rubin (1987) describe two types of models according to the ignorability of nonresponse. In the ignorable model the distribution of the variable of interest for a respondent is the same as the distribution of that variable for a nonrespondent with the same values of the covariates. In addition, the parameters in the distributions of the variable and response must be distinct (see Rubin 1976). All other models are nonignorable. Our model essentially incorporates both types, and is therefore nonignorable.

Crawford, Johnson and Laird (1993) used nonignorable nonresponse model to analyze data from the Harvard Medical Practice Survey. Stasny, Kadane, and Fritsch (1998) used a Bayesian hierarchical model for the probabilities of voting guilty or not on a particular trial when the views of nonrespondents might differ from those of respondents in various death-penalty beliefs . Park and Brown (1994) used a pseudo-Bayesian method (Baker and Laird 1988), and Park (1998) applied a method in which prior observations are assigned to both observed and unobserved cells to estimate the missing cells of a multi-way categorical table under nonignorable nonresponse.

Stasny (1991) used an empirical Bayes model to study victimization in the National Crime Survey, and used the selection approach . A related method was presented by Albert and Gupta (1985), they made an approximation to obtain a Bayesian approach for a single area, see also (Kaufman and King 1973).

Since Bayesian approach can incorporate prior information about nonrespondents, the Bayesian method is appropriate for the analysis of nonignorable nonresponse problems (Little and Rubin 1987 and Rubin 1987). However the main difficulty is how to describe the relationship between the respondents and nonrespondents. Using the selection approach within the framework of Bayes empirical Bayes (see Deely and Lindley 1981), Stasny (1991) estimated the hyper-parameters by maximum likelihood methods and then assumed them known. We extend this approach in two directions.

First, we consider multinomial data from several areas. It is worthy to note that Basu and Pereira (1982) considered multinomial nonresponse data from a single area using a multinomial Dirichlet model when the hyperparameters are assumed known. Recently, Forster and Smith (1998) used graphical multinomial Dirichlet loglinear models to analyze data from the panel survey in British general election. Again the hyper parameters are assumed known, and one area model is used. Secondly, we obtain a full Bayesian approach for multinomial nonignorable nonresponse data from several areas.

The rest of the paper is organized as follows. In Section 2 we describe the NHANES. In Section 3 we discuss the Bayesian model for nonignorable nonresponse. In particular, a three-stage Bayesian hierarchical multinomial model is applied to the NHANES data to investigate nonresponse problem. In Section 4 we describe an empirical analysis and a simulation study to assess the performance of our model. Finally, Section 5 has the conclusion.

2. NHANES Data and Nonresponse

The NHANES is one of the periodic surveys used to assess the status of health of the U.S. population. Our research is motivated by the presence of nonresponse in the NHANES data of Body Mass Index (BMI). The data for our illustration come from this survey, and were collected October 1988 and September 1994.

The NHANES consists of two parts, first part is the interview of the sampled person for their personal information and second part is the examination of those sampled. The persons from the sample of households were grouped into a number of subgroups depending on the age, race and sex. Some subgroups were sampled at different rates. Sampled persons were asked to come to station for physical examination. Those who did not come were visited by the examiner for the same purpose. Details of the NHANES sample design are available (Vital and Health statistics, Series 2, Number 113, 1992)

One of the variables of interest in the NHANES is BMI, a convenient index of weight adjusted for height (Kg/m^2) that can be used to broadly categorize bodyweight within age-race-sex groups (Kuczmarski et al. 1997) as low body fat (level 1: BMI < 20), healthy body fat (level 2: $20 \leq$ BMI < 25), hefty or unhealthy (level 3: BMI ≥ 25). We use this classification for the each of 8 age-race-sex groups.

NHANES data are adjusted by multistage of ratio weightings for the data to be consistent with the population (Leyla et al. 1994). The ratio is the proportion of persons in the sample to the number of persons who completed interview and examination. Weighting with nonresponse ratio is one of these stages. In this paper we investigate an alternative method to the nonresponse ratio weighting.

NHANES nonresponse also occurs at several levels in the survey: interview and examination. The interview nonresponse arises from sample persons who did not respond for the interview. Some of those who were already interviewed did not come to station, missing all or part of the examinations.

When the nonresponse of sample persons is adjusted by the ratio estimation within the same adjustment class, the distributions of the respondents and nonrespondents are assumed to be same. Clearly, this ratio estimation can be incorrect when these two groups are different. Therefore there is a need to consider the adjustment by a method other than ratio adjustment. We present Bayesian method as a possible alternative to impute the NHANES nonresponse.

Table 1 shows the number of respondents for each BMI level by age-race-sex group for 34 counties (population at least 500,000). The pattern of respondents differs greatly by age groups (young: age < 45 years ; old: age ≥ 45 years). The nonresponse rate for old age group is negligible. Therefore the main consideration about nonresponse is given to the young age group. There is also higher response rate among females than males.

Table 1: **Number of individuals in each BMI level and number of nonrespondents by age, race and sex over all 34 counties.**

			BMI			
Age	Race	Sex	1	2	3	Nonresponse
Y	W	M	1098	651	597	558
		F	845	434	380	233
	Ot	M	1198	713	665	574
		F	745	463	524	214
O	W	M	46	439	1014	3
		F	51	223	365	4
	Ot	M	79	470	942	8
		F	48	169	552	6

Note Y: Young, O: Old, W: White, Ot:Others

We develop a methodology to analyze the three category BMI data by age, race (white and others) and sex, although our methodology applies generally to any number of cells in several areas (counties in our application).

3. Methodology for Hierarchical Multinomial Model

In this section we describe the Bayesian model for old and young age groups. We use a two part hierarchical model for the BMI data. The first part of the model is for the old individuals. Since their nonresponse rate is very low, it is ignorable. The second part of the model is for the young individuals, since their nonresponse rate is high and differ from respondents, it is nonignorable. Note that we consider inference for each age-race-sex group separately. We will show how to combine them later using logistic regression although this is not the key issue of the paper.

For each age-race-sex group, an individual k in county i belongs to one of J BMI levels, then for k^{th} individual in county i, characteristic variable BMI level j is defined as follows,

$$\mathbf{x}_{ik} = (x_{i1k},\ ..., x_{ijk}, ...,\ x_{iJk})',\quad i = 1, ..., c;$$

$$j = 1, ..., J;\ k = 1, ..., n_i$$

where each $x_{ijk} = 0$ or 1 and $\sum_{j=1}^{J} x_{ijk} = 1$.

The response variable, y_{ijk} is defined for each age-race-sex group

$$y_{ijk} = \begin{cases} 1, & \text{if individual } k \text{ belonging to BMI level } j \\ & \text{in county } i \text{ responded} \\ 0, & \text{if individual } k \text{ belonging to BMI level } j \\ & \text{in county } i \text{ did not respond.} \end{cases}$$

We use a probabilistic structure to model the $\mathbf{x}_{ik}$ and y_{ijk}.

3.1 Modeling the Old and Young Individuals

For each age-race-sex group we have

$$\mathbf{x}_{ik} \mid \mathbf{p}_i \overset{iid}{\sim} \text{Multinomial } (1, \mathbf{p}_i) \qquad (1)$$

For the old age group we use the ignorable model

$$y_{ijk} \mid \pi_i \overset{iid}{\sim} \text{Bernoulli } (\pi_i), \qquad (2)$$

and at the second stage we take

$$\mathbf{p}_i \mid \boldsymbol{\mu}_1, \tau_1 \overset{iid}{\sim} \text{Dirichlet } (\boldsymbol{\mu}_1 \tau_1), \qquad (3)$$

$$\pi_i \mid \mu_{21}, \tau_{21} \overset{iid}{\sim} \text{Beta } (\mu_{21}\tau_{21}, (1-\mu_{21})\tau_{21}) \qquad (4)$$

where $p(\mathbf{p}_i \mid \boldsymbol{\mu}_1, \tau_1) = \prod_{j=1}^{J} p_{ij}^{\mu_{1j}\tau_1 - 1} / D(\boldsymbol{\mu}_1\tau_1)$, $0 < p_{ij} < 1$, $\sum_{j=1}^{J} p_{ij} = 1$ and $\boldsymbol{\mu}_1 = (\mu_{11}, \mu_{12}, \ldots, \mu_{1J})'$ with $D(\boldsymbol{\mu}_1\tau_1) = \Pi_{j=1}^{J}\Gamma(\mu_{1j}\tau_1)/\Gamma(\tau_1)$, $0 < \mu_{1j} < 1$, $\sum_{j=1}^{J} \mu_{1j} = 1$.

Assumptions (3) and (4) express similarity among the c counties.

For the young age group we have a nonignorable model. Thus, for the response variable we take

$$y_{ijk} \mid \mathbf{x}_{ik}, \pi_{ij} \overset{iid}{\sim} \text{Bernoulli } (\pi_{ij}) \qquad (5)$$

where $x_{ijk} = 1$ or $x_{ijk} = 0$, for $j = 1, 2, ..., J$. Letting $\boldsymbol{\mu}_3 = (\mu_{31}, \mu_{32}, ..., \mu_{3J})'$, at the second stage we also take

$$\mathbf{p}_i \mid \boldsymbol{\mu}_3, \tau_3 \overset{iid}{\sim} \text{Dirichlet } (\boldsymbol{\mu}_3\tau_3) \qquad (6)$$

and

$$\pi_{ij} \mid \mu_{4j}, \tau_{4j} \overset{iid}{\sim} \text{Beta } (\mu_{4j}\tau_{4j}, (1-\mu_{4j})\tau_{4j}) \qquad (7)$$

Like (3) and (4), the assumptions (6) and (7) express similarity among the counties. However the response parameters π_{ij} are weakly identifiable in this case, and (7) helps in the estimation of the π_{ij}.

To ensure a full Bayesian analysis, at the third stage we take the prior for the hyper-parameters as follows. For the ignorable part of the model, the joint prior distribution is

$$\boldsymbol{\mu}_1 \sim \text{Dirichlet } (1, 1, ..., 1),\quad \mu_{21} \sim \text{Beta } (1, 1),$$

$$\tau_1 \sim \Gamma\ (\eta_1^{(0)}, \nu_1^{(0)}),\quad \tau_{21} \sim \Gamma\ (\eta_{21}^{(0)}, \nu_{21}^{(0)}),$$

where the gamma density is given by
$t \sim \Gamma\ (a, b)$ means $f(t) = b^a t^{a-1} e^{-bt}/\Gamma(a)$, $t > 0$.

The corresponding part of the nonignorable model is

$$\boldsymbol{\mu}_3 \sim \text{Dirichlet } (1, 1, ..., 1),\ \mu_{4s} \sim \text{Beta } (1, 1),$$
$$\tau_3 \sim \Gamma\ (\eta_3^{(0)}, \nu_3^{(0)}) \text{ and}$$

$$\tau_{4s} \sim \Gamma\ (\eta_{4s}^{(0)}, \nu_{4s}^{(0)}),\ s = 1, 2.., J.$$

The hyper-parameters $\eta_3^{(0)}$, $\nu_3^{(0)}$, $\eta_{4s}^{(0)}$, $\nu_{4s}^{(0)}$, $s = 1, ..., J$, are to be specified.

Let r_i be the number of respondents in county i and y_{ij} the number of respondents for the j^{th} BMI level in county i. Then r_i and y_{ij} are random variables, $n_i - r_i$ is the number of nonrespondents. Since the number of respondents for j^{th} BMI level for the nonrespondents is unknown, we denote them by the latent variables z_{ij}.

The likelihood function for ignorable nonresponse is

$$f(\mathbf{y}, \mathbf{r} \mid \mathbf{p}_i, \pi_i) = \prod_{i=1}^{c} \left\{ \binom{n_i}{r_i} \pi_i^{r_i} (1-\pi_i)^{n_i - r_i} \right\}$$

$$\times \prod_{i=1}^{c} \left\{ \binom{r_i}{y_{i1}, .., y_{iJ}} \prod_{j=1}^{J} \left\{ p_{ij}^{y_{ij} + n_i - r_i} \right\} \right\}.$$

Letting $Z = \{\mathbf{z} : z_{ij} = 0, \ldots, n_i - r_i, \sum_{j=1}^{J} z_{ij} = n_i - r_i,\ i = 1, \ldots, \ell\}$, the likelihood function for the nonignorable model is proportional to $f(\mathbf{y}, \mathbf{r} \mid \mathbf{p}, \boldsymbol{\gamma}, \boldsymbol{\pi})$ where $f(\mathbf{y}, \mathbf{r} \mid \mathbf{p}, \boldsymbol{\gamma}, \boldsymbol{\pi}) = \sum_{\mathbf{z}:\ \mathbf{z} \in Z} f(\mathbf{y}, \mathbf{r}, \mathbf{z} \mid \mathbf{p}, \boldsymbol{\gamma}, \boldsymbol{\pi})$ and

$$f(\mathbf{y}, \mathbf{r}, \mathbf{z} \mid \mathbf{p}, \boldsymbol{\pi}) = \prod_{i=1}^{c} \left\{ \binom{n_i}{r_i} \binom{r_i}{y_{i1}, .., y_{iJ}} \right.$$
$$\left. \times \binom{n_i - r_i}{z_{i1}, .., z_{iJ}} \prod_{j=1}^{J} \{(\pi_{ij} p_{ij})^{y_{ij}} ((1 - \pi_{ij}) p_{ij})^{z_{ij}}\} \right\}$$

Using Bayes' theorem the joint posterior density of all the parameters are constructed for each nonresponse model. We consider inference about $\mathbf{p}_i$.

We use Markov chain Monte Carlo algorithm to obtain the posterior distribution of the $\mathbf{p}_i$ and π_i. Our plan is to obtain Metropolis-Hastings (MH) sampler to get samples from the joint posterior densities and then to use these samples to make posterior inferences about $\mathbf{p}_i$.

3.2 Computations

For the ignorable nonresponse model it is convenient to represent the posterior density function as

$$f(\mathbf{p}, \boldsymbol{\pi}, \boldsymbol{\mu}_1, \tau_1, \mu_{21}, \tau_{21} \mid \mathbf{y}, \mathbf{r})$$
$$= \prod_{i=1}^{c} \{f_1(\mathbf{p}_i \mid \mathbf{y}, \mathbf{r}, \boldsymbol{\mu}_1, \tau_1) f_2(\pi_i \mid \mathbf{y}, \mathbf{r}, \mu_{21}, \tau_{21})\}$$
$$\times\ f_3(\boldsymbol{\mu}_1, \tau_1, \mu_{21}, \tau_{21} \mid \mathbf{y}, \mathbf{r})$$

where $f_1(\cdot)$ is Dirichlet density, $f_2(\cdot)$ is beta density and $f_3(\cdot)$ is joint posterior. Hence f_1 and f_2 are obtained through the Gibbs kernel, while for f_3 we use the MH algorithm of Nandram (1998).

For nonignorable nonresponse model it is convenient to represent the posterior density function as

$$f(\mathbf{p}, \boldsymbol{\pi}, \mathbf{z}, \boldsymbol{\mu}_3, \tau_3, \boldsymbol{\mu}_4, \boldsymbol{\tau}_4 \mid \mathbf{y}, \mathbf{r})$$
$$= \prod_{i=1}^{c} \left\{ \left\{ \prod_{j=1}^{J} f_j(\pi_{ij} \mid \mathbf{y}, \mathbf{r}, \mathbf{z}, \mu_{4j}, \tau_{4j}) \right\} \right.$$
$$\times\ f_{J+1}(\mathbf{p}_i \mid \mathbf{y}, \mathbf{r}, \mathbf{z}, \boldsymbol{\mu}_3, \tau_3)\}$$
$$\times\ f_{J+2}(\boldsymbol{\mu}_3, \tau_3, \boldsymbol{\mu}_4, \boldsymbol{\tau}_4, \mathbf{z} \mid \mathbf{y}, \mathbf{r}),$$

where $f_1(\cdot)$, ..., $f_J(\cdot)$ are beta densities , $f_{J+1}(\cdot)$ is Dirichlet density and $f_{J+2}(\cdot)$ is joint posterior. Thus, $f_1, ..., f_{J+1}$ are obtained through the Gibbs kernel, while f_{J+2} is obtained using the MH algorithm of Nandram (1998).

4. An Empirical Analysis

In this section we illustrate our methodology using the NHANES data. Since p_{ij} are similar for each county, we study the weighted posterior mean of p_{ij},

$$q_j = \sum_{i=1}^{c} n_i p_{ij} / \sum_{i=1}^{c} n_i,\ j = 1, 2, 3$$

by age, race and sex for both the young and old age groups.

4.1 Data Analysis

First, we perform a sensitivity analysis to assess the specifications of $\eta^{(0)}$ and $\nu^{(0)}$. We compared 4 choices of hyper-priors $\Omega = (\eta^{(0)}, \nu^{(0)})$ to check its sensitivity to inference. For choice 1, we use 4 times Ω, i.e., $4\Omega = (4\eta^{(0)}, 4\nu^{(0)})$. For choice 2, we use the hyper-prior without any change, i.e., $\Omega = (\eta^{(0)}, \nu^{(0)})$. For choice 3, we use the one fourth of Ω, i.e., $\Omega/4 = (\eta^{(0)}/4, \nu^{(0)}/4)$ and fourth choice is $\Omega = (0, 0)$. The simulation results for the sensitivity to the inference of q_j shows that the point estimates and standard deviations of the proportion are very similar over the four choices of hyper-priors for young age group. For old age group there are some changes in point estimates and standard deviations. Generally the nonignorable model performs better than the ignorable model, as the nonignorable model is not sensitive to choices of hyper-priors.

In Table 2 we present 95% credible intervals for the weighted posterior means, q_j. For the young age group, the weighted posterior mean is the highest for q_1 of BMI level 1, and q_2 of BMI level 2 is the lowest. The lower bounds for q_1 and q_3 are similar for the young age group except white male, and those for q_2 are similar except non-white male group. For old age group, the weighted posterior mean is highest for q_3 of BMI level 3, and q_1 of BMI level 1 is lowest. Specifically q_1, q_2 are high and q_3 is low for white-female group.

4.2 Linear and Nonlinear Logistic Regression Model

It is possible to relate the p_{ijl} to age, race and sex using a linear or nonlinear logistic regression model.

Table 2: **95% credible intervals for the weighted posterior means, $q_j = \sum_{i=1}^{c} n_i p_{ij} / \sum_{i=1}^{c} n_i$ by age, race and sex for each age group**

Age	Race	Sex	95% credible interval q_1	q_2	q_3
Y	W	M	(.382 .470)	(.174 .252)	(.314 .412)
		F	(.425 .525)	(.171 .269)	(.243 .371)
	Ot	M	(.381 .455)	(.176 .241)	(.333 .419)
		F	(.385 .482)	(.130 .230)	(.329 .442)
O	W	M	(.022 .041)	(.255 .326)	(.643 .710)
		F	(.059 .068)	(.431 .451)	(.486 .505)
	Ot	M	(.035 .076)	(.282 .352)	(.592 .670)
		F	(.040 .093)	(.206 .265)	(.661 .731)

Note Y: Young, O: Old, W: White, Ot:Others

Table 3: **Comparison of 95 % credible intervals for θ_1, θ_2 and α_1,α_2,...,α_8 by regression type.**

	Linear	Nonlinear
θ_1	(-1.743 -1.469)	(-1.731 -1.466)
θ_2	(0.028 0.196)	(0.025 0.193)
α_1	(-1.167 -0.751)	(-1.159 -0.751)
α_2	(-1.395 -0.939)	(-1.385 -0.937)
α_3	(-1.127 -0.723)	(-1.119 -0.728)
α_4	(-1.112 -0.659)	(-1.103 -0.658)
α_5	(1.198 1.514)	(1.188 1.498)
α_6	(0.513 0.689)	(0.506 0.685)
α_7	(0.715 1.210)	(0.725 1.225)
α_8	(0.809 1.310)	(0.803 1.300)

Let p_{ijl} denote the probability that a respondent in l^{th} (l=1,...,8) age-race-sex group in county i responds in the j^{th} BMI level and let $\sum_{s=1}^{j} p_{isl}$ be the cumulative probability. Letting $y_{ijl} = \log\left\{\sum_{s=1}^{j} p_{isl} / \left(1 - \sum_{s=1}^{j} p_{isl}\right)\right\}$, we take

$$y_{ijl} = (\theta_j - (\mu_i + \alpha_l))/\psi_i \qquad (8)$$

subject to the constraints $\sum_{i=1}^{c} \mu_i = 0$, $\sum_{l=1}^{8} \alpha_l = 0$, and $\sum_{i=1}^{c} \ln \psi_i = 0$. In (8) θ_j, μ_i, α_l and ψ_i have posterior distributions whose properties are inherited from the posterior distributions of p_{ijl}. Each iterate of the MH algorithm provides a value for p_{ijl} which is used in (8), and a nonlinear least squares problem is solved by using an iterative method to get values of θ_j, μ_i, α_l and ψ_i. Alternatively, we can also use the much simpler linear logistic model in which the ψ_i in (8) are taken equal to unity. In this case at the h^{th} iteration of MH algorithm the least squares estimators of θ_j, ϕ_i, μ_i and α_l exist in closed form. Specifically, for $\phi_i = 0$, we have least squares estimates $\hat{\mu}_i = \bar{y}_{...} - \bar{y}_{i..}$, $\hat{\theta}_j = \bar{y}_{.j.}$, $\hat{\alpha}_l = \bar{y}_{...} - \bar{y}_{..l}$, where $\bar{y}_{...} = \sum_{i=1}^{c}\sum_{j=1}^{J}\sum_{l}^{8} y_{ijl}/8cJ$, $\bar{y}_{i..} = \sum_{j=1}^{J}\sum_{l=1}^{8} y_{ijl}/8J$, $\bar{y}_{.j.} = \sum_{i=1}^{c}\sum_{l=1}^{8} y_{ijl}/8c$ and $\bar{y}_{..l} = \sum_{i=1}^{c}\sum_{j=1}^{J} y_{ijl}/cJ$. The nonlinear least squares problem is solved by using an iterative method to get values of $\hat{\theta}_j$, $\hat{\phi}_i$, $\hat{\mu}_i$ and $\hat{\alpha}_l$.

We present 95% credible intervals for θ_1, θ_2 and $\alpha_1, \ldots, \alpha_8$ for the young and old age group by regression type in Table 3. For cut-points θ_j ($\theta_1 < \theta_2$), θ_1 gives large negative effect compare to θ_2. The relative measure α_l ($l = 1,..,4$) of young age group gives negative effect, while the relative measure α_l ($l = 5,..,8$) of old age group give positive effects. The 95% credible intervals for linear and nonlinear estimates are essentially the same.

5. Conclusion

We have discussed the problem of nonignorable nonresponse for the estimation of the BMI proportions. Our main issue was to get an efficient alternative method to estimate BMI in the NHANES in such a way as to incorporate nonignorability. We introduced a three stage hierarchical model to solve the nonresponse problem. We have extended the model of Stasny (1991) in two directions. First we consider multinomial data other than binomial data and second we study a full Bayesian analysis. We applied our model to the NHANES data, and have shown that the nonresponse was reasonably addressed by our model. The MCMC method allowed us to assess the complex structure of the multinomial nonresponse estimation. Our empirical analysis indicate good performance of the model for this data. Thus, the method of ratio estimation currently used in NHANES may be replaced by our Bayesian method as the nonrespondents' characteristics might differ from those of the respondents.

For the NHANES there are substantial differences in the proportion of individuals in the 3 BMI levels for males versus females and young versus old. While we have shown that inference about BMI is not sensitive to prior specification, we might want to use other prior densities for the sum of the Dirichlet or beta parameters (i.e., a uniform shrinkage prior) or an improper prior for vector of means for the Dirichlet or the one mean for the beta distribution.

It is feasible to use a nonignorable model that incorporates the extent of nonignorability from the ignorable model.

References

Albert, J. H. and Gupta, A. K. (1985), "Bayesian Methods for Binomial Data with Applications to a Nonresponse Problem," *Journal of the American statistical Association*, 80, 167-174.

Baker, S. G. and Laird, N. M. (1988), "Regression analysis for categorical variables with Outcome Subject to Nonignorable nonresponse," *Journal of the American statistical Association*, 83, 62-69.

Basu, D. and Pereira, C. A. (1982), "On the Bayesian Analysis of Categorical Data: The Problem of Nonresponse," *Journal of Statistical Planning and Inference*, 6, 345-362.

Crawford, S. L., Johnson, W. G. and Laird, N. M. (1993), "Bayes Analysis of Model-Based Methods for Nonignorable Nonresponse in the Harvard Medical Practice Survey (with discussions)," *In Case Studies in Bayesian Statistics.* Gatsonis C., Hodges, J.S., Kass, R.E. and Sinpurwalla, N.D. (eds.) Springer-Verlag: New York, pp. 78-117.

De Heer, W. (1999), "International Response Trends: Results of an International Survey," *Journal of Official Statistics*, 15, 129-142.

Deely, J. J. and Lindley, D. V. (1981), "Bayes Empirical Bayes," *Journal of the American Statistical Association*, 76, 833-841.

Forster, J. J. and Smith, W. F. (1998), "Model-based inference for categorical survey data subject to nonignorable nonresponse," *Journal of the Royal Statistical society*, Series B, 60, 57-70.

Groves, R. M. and Couper, M. P. (1998), *Nonresponse in Household Interview Surveys*, New York: Wiley.

Heckman, J. (1976), "The Common Structure of Statistical Models of truncation, sample selection and limited dependent variables and a simple estimator for such models," *Annals of Economic and Social Measurement*, 5, 475-492.

Kuczmarski, R. J., Carrol, M. D., Flegal, K. M. and Troiano, R. P. (1997), "Varying Body Mass Index cutoff points to describe overweight prevalence among U.S. adults: NHANES III (1988 to 1994)," *Obesity Research*, 5, 542-548.

Kaufman, G. M. and King, B. (1973), "A Bayesian Analysis of Nonresponse in Dichotomous Processes," *Journal of the American Statistical Association*, 68, 670-678.

Little R. J. A. and Rubin D. B. (1987), *Statistical Analysis with Missing Data*, New York: Wiley.

Mohadjar, L., Bell, B. and Waksberg, J. (1994), "National health and Nutrition Examination Survey III-Accounting for item nonresponse bias," National Center for Health Statistics.

Nandram, B. (1998), "A Bayesian Analysis of the Three-Stage Hierarchical Multinomial Model," *Journal of Statistical Computation and Simulation*, 61, 97-126.

National Center for Health Statistics (1994), "Plan and Operation of the Third National Health and Nutrition Examination Survey," *Vital and Health Statistics Series 1*, 32.

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

Olson, R. L. (1980), "A Least Squares Correction for Selectivity Bias," *Econometrica*, 48, 1815-1820.

Park, T. (1998), "An Approach to Categorical Data Nonignorable Nonresponse," *Biometrics*, 54, 1579-1590.

Park, T. and Brown, M. B. (1994), "Models for Categorical Data with Nonignorable Nonresponse," *Journal of the American Statistical Association*, 89, 44-52.

Rubin D. B. (1976), "Inference and Missing Data," *Biometrika*, 63, 581-590.

Rubin D. B. (1987), *Multiple Imputation for Nonresponse in Surveys*, New York: Wiley.

Stasny, E. A. (1991), "Hierarchical Models for the Probabilities of a Survey Classification and Nonresponse: An Example from the National Crime Survey," *Journal of the American Statistical Association*, 86, 296-303.

Stasny, E. A., Kadane, J. B., and Fritsch, K, S. (1998), "On the Fairness of death Penalty Jurors: A Comparison of Bayesian Models with Different levels of Hierarchy and Various Missing-Data Mechanisms," *Journal of the American Statistical Association*, 93, 464-477.