Journal of Korean
Data & Information Science Society
2007, Vol. 18, No. 4, pp. 1179 ~ 1189

Small Domain Estimation of the Proportion Using Survey Weights

Dal Ho Kim[1)]

Abstract

In this paper, we estimate the proportion of individuals having health insurance in a given year for several small domains cross-classified by age, sex and other demographic characteristics using the data provided by the National Center for Health Statistics(NCHS). We employ Bayesian as well as frequentist methodology to obtain small domain estimates and the associated measures of precision. One of the new features of our study is that we utilize the survey weights along with the model to derive the small domain estimates.

Keywords : Empirical Best Linear Unbiased Predictors, Hierarchical Bayesian, Proportion, Survey Weights, Small Domains

1. Introduction

Our primary interest is to estimate the proportion of individuals having health insurance in a given year for several domains cross-classified by age, sex and other demographic characteristics. In our analysis, we have constructed the domains on the basis of age, sex, race and the region where the person belongs. Using the data provided by the National Center for Health Statistics(NCHS), we can estimate these proportions.

The original survey for any given year contains data on more than 100,000 individuals and on over 800 variables. Of these individuals, we have information on the primary response variable, namely whether a person has health insurance or not. In addition, there is information on demographic characteristics such as age, sex, race, region, education, income status, medical condition, disability conditions (if any) and many other socio-economic factors.

1) Professor, Department of Statistics, Kyungpook National University, Daegu 702-701, Korea
E-mail: dalkim@knu.ac.kr

For the entire US population, the direct estimates for these domains, namely the sample proportions, are fairly reliable, since the sample size for each domain is reasonably large. This need not be the case though when our analysis is targeted towards specific subpopulations, such as hispanics, Asians and similar minority sectors of the community.

For a targeted minority subpopulation, the sample size in a domain is not always very large. Hence, the direct estimates may not be very reliable, being accompanied with large standard errors and coefficients of variations. This calls for the use of small area or small domain estimation techniques, where indirect estimates are obtained for these domains based on implicit or explicit models. These models help building a link between these domains, and thus produce typically estimates of greater precision by borrowing strength. (cf. Rao, 2003).

The ultimate objective of our study is to employ Bayesian as well as frequentist methodology to obtain small domain estimates and find also the associated measures of precision. We will use a hierarchical Bayesian(HB) analogue of the generalized linear mixed model(GLMM) to obtain posterior means and posterior standard errors of the population small domain proportions. The frequentist approach will use the empirical best linear unbiased predictors(EBLUP) of these proportions. One of the new features of our analysis is that we will be utilizing the survey weights along with the model to derive the small domain estimates. Thus, our method, in some sense, can be regarded as design-assisted model-based estimation.

Our eventual analysis will produce small domain estimates of the proportion of uninsured persons for the Hispanic, Asian and other minority groups. But the present talk is focused specifically on the Asian population. We provide the estimates and measures of precision are based on the frequentist and Bayesian approaches, and compare and contrast the two methods. We have the results for 1997-1999, but we report the results only for 1997.

The Asian group is formally composed of the (1) Chinese, (2) Filipino, (3) Asian Indian, and (4) Islanders such as Koreans, Vietnamese, Japanese, Hawaiian, Samoan, Guamanian etc. For the year 1997, there are about 3500 such individuals. These 3500 individuals are assigned to specific domains depending on their age, race, gender and the region they come from. There are 3 age-groups (0-17, 18-64 and 65+), 2 Genders, 4 Races and 4 Regions depending on the size of the Metropolitan Statistical Area (<499,999 ; 500,000-999,999 ; 1,000,000-2,499,999, >2,500,000). Thus, the total number of domains equals $3\times2\times4\times4=96$. When the individuals are distributed to their respective domains, it turns out that many of the domains contain only a few samples. Indeed, there are several domains with a sample of size 1, while domain 58 has sample size zero.

The outline of the remaining sections is as follows. Section 2 addresses the selection of covariates and discusses the general hierarchical Bayesian methodology. Section 3 includes the EBLUP approach. Finally, Section 4 contains the specific application.

2. Hierarchical Bayesian Approach

2.1. Selection of Covariates

As mentioned in the introduction, the number of covariates exceeds 800. Inclusion of all of them in the initial model is impractical and unnecessary. We started with a set of 6 covariates and went through a process of forward and backward selection, and finally kept the best model with minimum number of covariates.

Initially we started with the following covariates: (1) legal marital status, (2) family size, (3) education level, (4) total earning from previous year, (5) total family income, and (6) full time working status.

Two-thirds of the data on full time working status were missing for 1997. Hence, this covariate was dropped immediately for model selection. Also, legal marital status and total earning from previous year were both found insignificant. Thus, along with the intercept term the selected covariates are family size, education level, and total family income.

We use SAS Version8 for the initial computations in UNIX workstation. Also, we use the SURVEYREG Procedure for model.

2.2. HB Methodology

We consider random variables y_{ij} such that conditional on p_{ij}, these are independent binary variables with success probabilities $p_{ij}, j=1,\cdots,n_i, i=1,\cdots,k$. In our specific application, p_{ij} denotes the probability of the absence of health insurance for the jth unit in the ith small domain. We model $\theta_{ij}=\text{logit}\,(p_{ij})$ as

$$\theta_{ij}=\boldsymbol{x}_{ij}^T\boldsymbol{b}+u_i$$

where $\boldsymbol{x}_{ij}$ are p-component design vectors, $\boldsymbol{b}$ is the vector of regression parameters, and u_i are the random effects. It is assumed that u_i are iid $N(0,\sigma_u^2)$. Also, let $\text{rank}(X^T)=(\boldsymbol{x}_{11},\cdots,\boldsymbol{x}_{1n_1},\cdots,\boldsymbol{x}_{k1},\cdots,\boldsymbol{x}_{kn_k})=p$.

Finally, it is assumed that $\boldsymbol{b}$ and σ_u^2 are mutually independent with $\boldsymbol{b}\sim\text{Unif}(R^p)$ and $\sigma_u^2\sim\text{IG}(c/2,d/2)$, i.e. $\pi(\sigma_u^2)\propto\exp\left(-\frac{c}{2\sigma_u^2}\right)(\sigma_u^2)^{-d/2-1}, c>0$.

Let $\boldsymbol{y}=(y_{11},\cdots,y_{1n_1},\cdots,y_{k1},\cdots,y_{kn_k})^T$, and $\boldsymbol{\theta}=(\theta_{11},\cdots,\theta_{1n_1},\cdots,\theta_{k1},\cdots,\theta_{kn_k})^T$. Then the joint posterior is given by

$$\pi(\boldsymbol{\theta}, \boldsymbol{b}, \sigma_u^2 \mid \boldsymbol{y}) \propto \prod_{i=1}^{k}\prod_{j=1}^{n_i} \exp[y_{ij}\theta_{ij} - \log(1+\exp(\theta_{ij}))]$$
$$\times (\sigma_u^2)^{-k/2}\exp\left[\frac{1}{2\sigma_u^2}(\theta_{ij} - \boldsymbol{x}_{ij}^T\boldsymbol{b})^2\right]$$
$$\times (\sigma_u^2)^{-d/2-1}\exp\left[-\frac{c}{2\sigma_u^2}\right].$$

This is a nonconjugate Bayesian analysis, and is not implementable analytically. Instead, we use the Markov chain Monte Carlo(MCMC) numerical integration technique. In particular, we employ the Gibbs sampler. To this end, we need to find the full conditionals of θ_{ij}, $\boldsymbol{b}$ and σ_u^2.

The full conditionals are given by

$$\sigma_u^2 \mid \boldsymbol{\theta}, \boldsymbol{b}, \boldsymbol{y} \sim \mathrm{IG}\left(\frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(\theta_{ij} - \boldsymbol{x}_{ij}^T\boldsymbol{b})^2 + c}{2}, \frac{k+d}{2}\right);$$
$$\boldsymbol{b} \mid \boldsymbol{\theta}, \sigma_u^2, \boldsymbol{y} \sim N((\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}, \sigma_u^2(\boldsymbol{X}^T\boldsymbol{X})^{-1});$$
$$\theta_{ij} \mid \boldsymbol{b}, \sigma_u^2, \boldsymbol{y} \sim f(y_{ij} \mid \theta_{ij})\exp\left[-\frac{1}{2\sigma_u^2}(\theta_{ij} - \boldsymbol{x}_{ij}^T\boldsymbol{b})^2\right]$$

Our data analysis is based on generating samples from the above conditionals. If $\hat{\theta}_{ij}^{(r)}$ denotes the sampled value of θ_{ij} generated from the rth draw, and the number of draws is R, then the Monte Carlo estimate of $E(\theta_{ij} \mid \boldsymbol{y})$ is $R^{-1}\sum_{r=1}^{R}\hat{\theta}_{ij}^{(r)}$, Similarly, the Monte-Carlo estimate of $V(\theta_{ij} \mid \boldsymbol{y})$ is $R^{-1}\sum_{r=1}^{R}(\hat{\theta}_{ij}^2)^{(r)} - \left(R^{-1}\sum_{r=1}^{R}\hat{\theta}_{ij}^{(r)}\right)^2$.

3. EBLUP Approach

Let y_{ij} denote the response of the jth unit in the ith small area, ω_{ij} the weight attached to y_{ij} (usually inverse of the selection probability), and $\sum_{j=1}^{n_i}\omega_{ij} = 1$. Our basic data consist of $\bar{y}_{iw} = \sum_{j=1}^{n_i} w_{ij}y_{ij}$, $i = 1,\cdots,k$, and not the y_{ij} themselves.

Suppose conditional on p_{ij}, y_{ij} are independent binary variables with success probabilities p_{ij}, $j = 1,\cdots,n_i, i = 1,\cdots,k$. Our interest is $\mu_i = \sum_{j=1}^{n_i} w_{ij}p_{ij}$, $i = 1,\cdots,k$.

Then we compute

$$E(\bar{y}_{iw}) = E(\mu_i) = \sum_{j=1}^{n_i} w_{ij}\, m_{ij} = m_i \text{ (say)};$$

$$\begin{aligned} var(\bar{y}_{iw}) &= var(\mu_i) + E[\sum_{j=1}^{n_i} w_{ij}^2\, p_{ij}\,(1-p_{ij})] \\ &= \sum_{j=1}^{n_i} w_{ij}^2\,[var(p_{ij}) + E\{p_{ij}\,(1-p_{ij})\}] \\ &= \sum_{j=1}^{n_i} w_{ij}^2\,[E(p_{ij}) - \{E(p_{ij})\}^2] \\ &= \sum_{j=1}^{n_i} w_{ij}^2\, m_{ij}\,(1-m_{ij}); \end{aligned}$$

$$\begin{aligned} cov(\bar{y}_{iw}, \mu_i) &= E(\mu_i^2) - (E\mu_i)^2 \\ &= var(\mu_i) = \sum_{j=1}^{n_i} w_{ij}^2\, var(p_{ij}). \end{aligned}$$

Consider now the conjugate prior

$$p_{ij} \overset{ind}{\sim} Beta(\lambda m_{ij}, \lambda(1-m_{ij}))$$

where $logit(m_{ij}) = \boldsymbol{x}_{ij}^T \boldsymbol{b}$. Here $\boldsymbol{x}_{ij}$ is the design vector for the ith small area, $\boldsymbol{b}$ is the regression coefficient. Then $E(p_{ij}) = m_{ij}$ and $var(p_{ij}) = \dfrac{m_{ij}(1-m_{ij})}{\lambda+1}$.

Remark 1. The formulation is non-Bayesian. The word "prior" is used only for convenience. Indeed the marginal distribution of y_{ij} is an overdispersed beta-binomial model.

The BLUP of μ_i based on $\bar{y}_{iw}$ is

$$\begin{aligned} & m_i + \frac{\sum_{j=1}^{n_i} w_{ij}^2\, m_{ij}\,(1-m_{ij})/(\lambda+1)}{\sum_{j=1}^{n_i} w_{ij}^2\, m_{ij}\,(1-m_{ij})}\,(\bar{y}_{iw} - m_i) \\ &= m_i + \frac{1}{\lambda+1}\,(\bar{y}_{iw} - m_i) = \frac{1}{\lambda+1}\bar{y}_{iw} + \frac{\lambda}{\lambda+1}\, m_i \end{aligned}$$

where $m_i = \sum_{j=1}^{n_i} w_{ij} \frac{\exp(\boldsymbol{x}_{ij}^T \boldsymbol{b})}{1+\exp(\boldsymbol{x}_{ij}^T \boldsymbol{b})} = \sum_{j=1}^{n_i} w_{ij} - \sum_{j=1}^{n_i} \frac{w_{ij}}{1+\exp(\boldsymbol{x}_{ij}^T \boldsymbol{b})}$. Note that this derivation depends on the means, variances and covariances of $\bar{y}_{iw}$ and μ_i.

However, $\boldsymbol{b}$ and λ are unknown, and need to be estimated from the marginal distributions of $\bar{y}_{iw}$, $i=1,\cdots,k$. We use the theory of optimal estimating functions (Godambe and Thompson, 1989) for simultaneous estimation of $\boldsymbol{b}$ and λ. Unfortunately, $var(\bar{y}_{iw})$ is not involving λ. Hence we estimate $\boldsymbol{b}$ only given λ. This requires evaluation of only the first two marginal moments of $\bar{y}_{iw}$, $i=1,\cdots,k$, rather than full knowledge of their distributions.

Following Godambe and Thompson(1989), we begin with the elementary unbiased estimating function $g_i = \bar{y}_{iw} - m_i$. Let

$$D_i = E(-\frac{\partial g_i}{\partial \boldsymbol{b}}) = \sum_{j=1}^{n_i} \frac{w_{ij}\exp(\boldsymbol{x}_{ij}^T \boldsymbol{b})}{[1+\exp(\boldsymbol{x}_{ij}^T \boldsymbol{b})]^2} \boldsymbol{x}_{ij}$$

and

$$\sigma_i^2 = var(g_i) = var(\bar{y}_{iw}) = \sum_{j=1}^{n_i} \frac{w_{ij}^2 \exp(\boldsymbol{x}_{ij}^T \boldsymbol{b})}{[1+\exp(\boldsymbol{x}_{ij}^T \boldsymbol{b})]^2}$$

Then we solve the optimal estimating equation

$$\begin{aligned} 0 = \sum_{i=1}^{k} D_i^T \sigma_i^{-2} g_i &= \sum_{i=1}^{k} \frac{\bar{y}_{iw} - m_i}{\sigma_i^2} D_i \\ &= \sum_{i=1}^{k} \frac{\bar{y}_{iw} - m_i}{\sigma_i^2} \sum_{j=1}^{n_i} \frac{w_{ij}\exp(\boldsymbol{x}_{ij}^T \boldsymbol{b})}{[1+\exp(\boldsymbol{x}_{ij}^T \boldsymbol{b})]^2} \boldsymbol{x}_{ij} \end{aligned}$$

and get the estimate $\hat{\boldsymbol{b}}$. Accordingly, an EBLUP estimator of μ_i is

$$\frac{1}{\lambda+1}\bar{y}_{iw} + \frac{\lambda}{\lambda+1}\sum_{j=1}^{n_i} w_{ij} \frac{\exp(\boldsymbol{x}_{ij}^T \hat{\boldsymbol{b}})}{1+\exp(\boldsymbol{x}_{ij}^T \hat{\boldsymbol{b}})}$$

given λ.

4. Data Analysis

Here the data consist of $y_{ij} = 1$ or 0 if the jth unit in the ith small domain does not (does) have health insurance;

$\tilde{w}_{ij}$ = the sampling weight attached to the jth unit in the ith small domain;

$w_{ij} = \tilde{w}_{ij} / \sum_{j=1}^{n_i} \tilde{w}_{ij}$ so that $\sum_{j=1}^{n_i} w_{ij} = 1$ for each i.

x_{ij1} = the family size of the jth unit in the ith small domain;

x_{ij2} = the education level of the jth unit in the ith small domain;

x_{ij3} = total family income of the jth unit in the ith small domain;

Let $p_{ij} = E(y_{ij})$. We model $\theta_{ij} = \text{logit}(p_{ij}) = b_0 + b_1 x_{ij1} + b_2 x_{ij2} + b_3 x_{ij3} + u_i$, $j = 1, \cdots, n_i$, $i = 1, \cdots, 95$.

There are two sets of direct domain estimates.

$$\text{(i)}\ \hat{p}_{i0} = n_i^{-1} \sum_{j=1}^{n_i} y_{ij},\quad \text{(ii)}\ \hat{p}_{iw} = \sum_{j=1}^{n_i} w_{ij} y_{ij}$$

The corresponding sets of hierarchical Bayes estimates are given by

$$\text{(iii)}\ \hat{p}_{i0}^{HB} = n_i^{-1} \sum_{j=1}^{n_i} E(\theta_{ij} \mid \boldsymbol{y}),\quad \text{(iv)}\ \hat{p}_{iw}^{HB} = \sum_{j=1}^{n_i} w_{ij} E(\theta_{ij} \mid \boldsymbol{y}).$$

We find Monte Carlo estimates of (iii) and (iv) as described in Section 2. Our hyperprior considers: $c = 0.2, 0.02, 0.002$; $d = 0.2, 0.02, 0.002$. The results are very insensitive to the choice of the hyperpriors, and are reported only for $c = d = 0.02$. Also we find EBLUP estimates for $\lambda = 0.1, 0.5, 1.0$. Table 1 provides the different estimates of the proportions of uninsured people for the different domains (with the exception of domain 58 which has sample size zero), and the posterior standard deviation associated with $\hat{p}_{iw}^{HB}$, i.e. the Monte Carlo estimate of $V(\sum_{j=1}^{n_i} w_{ij} p_{ij} \mid \boldsymbol{y})$.

It follows from the Table 1 that the HB estimates are hardly different from the direct estimates. The intuitive reason for this seems to be the fact that possibly in this situation model variability far too outweighs the sampling variability. Since the HB estimates are, in some sense, weighted averages of direct estimates and regression estimates with the respective weights close to the inverses of the sampling variance and the model variance, this phenomenon is not totally impossible. However, EBLUP estimates with $\lambda = 0.5$ and $\lambda = 1.0$ are quite comparable with HB estimates. We may also point out once again that the HB method is capable of providing not only the estimates, but also the associated measures of precision.

Table 2 provides the summary table for the proprtion of uninsured for the three age groups 0-17, 18-64 and 65+ individually for Chinese (Asian 1), Filipino (Asian

2), Asian Indian (Asian 3) and other Asians (Asian 4). It turns out that at this higher level of aggregation, both the EB and HB small domain estimates are fairly close to the corresponding direct estimates except possibly for the age-group 65+.

Table 1. Small area estimates of the proportions of uninsured people :year 1997

Domain	n_i	Raw1	Raw2 (wt)	HB1	HB2	se(HB2)	EBLUP $\lambda=.1$	EBLUP $\lambda=.5$	EBLUP $\lambda=1$
1	11	.182	.103	.177	.121	.045	.108	.120	.128
2	8	.125	.088	.115	.092	.044	.090	.094	.096
3	36	.000	.000	.048	.047	.033	.012	.044	.066
4	23	.174	.194	.166	.178	.037	.188	.172	.161
5	29	.379	.382	.317	.319	.051	.370	.339	.318
6	18	.111	.086	.125	.109	.037	.096	.121	.139
7	96	.167	.156	.165	.159	.017	.158	.164	.167
8	58	.328	.323	.276	.272	.037	.311	.280	.258
9	4	.000	.000	.049	.049	.067	.015	.056	.084
10	1	.000	.000	.054	.054	.133	.016	.057	.086
11	14	.000	.000	.058	.059	.050	.020	.073	.109
12	6	.000	.000	.066	.063	.066	.024	.089	.134
13	10	.100	.038	.124	.081	.051	.049	.080	.101
14	5	.200	.163	.171	.148	.071	.161	.156	.152
15	38	.184	.179	.171	.168	.028	.178	.173	.169
16	21	.333	.295	.269	.242	.048	.280	.242	.215
17	31	.258	.263	.235	.239	.036	.261	.255	.252
18	20	.200	.164	.187	.163	.036	.166	.171	.175
19	103	.136	.131	.145	.142	.017	.135	.147	.156
20	66	.273	.303	.240	.261	.034	.293	.264	.245
21	1	.000	.000	.039	.039	.111	.015	.053	.080
22	2	.500	.553	.405	.442	.167	.530	.468	.425
23	11	.000	.000	.062	.064	.057	.022	.080	.119
24	7	.000	.000	.075	.075	.072	.028	.101	.151
25	31	.065	.082	.094	.104	.030	.085	.092	.098
26	12	.000	.000	.047	.047	.045	.012	.044	.066
27	34	.029	.022	.064	.058	.028	.030	.051	.066
28	36	.056	.050	.080	.075	.026	.054	.066	.075
29	55	.255	.284	.225	.244	.034	.272	.241	.219
30	25	.000	.000	.051	.051	.040	.013	.046	.069
31	67	.090	.087	.104	.103	.020	.091	.101	.108
32	50	.120	.132	.130	.138	.022	.133	.135	.137

Table 1. Small area estimates of the proportions of uninsured people: year 1997 (continued)

Domain	n_i	Raw1	Raw2 (wt)	HB1	HB2	se(HB2)	EBLUP $\lambda=.1$	EBLUP $\lambda=.5$	EBLUP $\lambda=1$
33	4	.000	.000	.046	.044	.062	.011	.040	.060
34	5	.000	.000	.057	.058	.070	.014	.052	.078
35	4	.000	.000	.045	.045	.064	.008	.030	.045
36	6	.000	.000	.047	.047	.057	.016	.058	.087
37	33	.152	.164	.157	.165	.030	.163	.163	.163
38	13	.077	.057	.109	.094	.045	.065	.085	.099
39	27	.000	.000	.043	.041	.033	.010	.038	.057
40	24	.000	.000	.044	.044	.034	.008	.030	.045
41	64	.141	.186	.154	.184	.023	.185	.183	.182
42	34	.118	.108	.135	.127	.028	.112	.122	.128
43	83	.096	.083	.113	.103	.019	.087	.099	.107
44	70	.157	.185	.154	.172	.023	.183	.178	.175
45	6	.000	.000	.058	.057	.063	.018	.066	.099
46	6	.000	.000	.059	.059	.064	.016	.058	.086
47	6	.167	.125	.182	.155	.071	.140	.179	.207
48	5	.000	.000	.056	.056	.068	.018	.067	.100
49	10	.400	.392	.316	.308	.078	.370	.312	.273
50	7	.000	.000	.045	.044	.050	.012	.043	.065
51	17	.118	.093	.132	.110	.036	.096	.104	.110
52	23	.174	.154	.174	.160	.034	.154	.153	.153
53	23	.261	.265	.235	.237	.041	.260	.247	.238
54	38	.158	.160	.160	.162	.028	.165	.179	.189
55	37	.081	.080	.108	.104	.028	.088	.112	.127
56	66	.364	.399	.302	.326	.048	.381	.332	.298
57	1	.000	.000	.050	.050	.127	.012	.044	.066
58	0	-	-	-	-	-	-	-	-
59	2	.500	.587	.370	.424	.194	.543	.424	.343
60	1	.000	.000	.093	.093	.180	.017	.062	.092
61	10	.200	.185	.177	.166	.051	.179	.164	.154
62	10	.400	.343	.324	.281	.067	.326	.282	.252
63	11	.091	.091	.102	.100	.042	.092	.097	.101
64	24	.375	.359	.306	.292	.054	.340	.290	.255

Table 1. Small area estimates of the proportions of uninsured people: year 1997 (continued)

Domain	n_i	Raw1	Raw2 (wt)	HB1	HB2	se(HB2)	EBLUP $\lambda=.1$	EBLUP $\lambda=.5$	EBLUP $\lambda=1$
65	19	.105	.091	.127	.117	.038	.101	.126	.144
66	22	.318	.328	.259	.263	.055	.310	.263	.230
67	35	.171	.174	.169	.166	.030	.174	.172	.172
68	44	.300	.278	.253	.240	.035	.268	.239	.220
69	1	.000	.000	.043	.043	.116	.016	.059	.089
70	1	.000	.000	.038	.038	.108	.006	.023	.035
71	1	.000	.000	.079	.079	.163	.032	.115	.173
72	1	.000	.000	.087	.087	.171	.033	.123	.184
73	66	.091	.089	.121	.119	.025	.095	.110	.120
74	51	.118	.132	.133	.143	.023	.132	.133	.134
75	97	.083	.075	.113	.108	.024	.081	.098	.110
76	53	.245	.232	.222	.210	.027	.226	.208	.196
77	76	.197	.208	.198	.206	.020	.209	.212	.214
78	79	.139	.171	.146	.169	.019	.172	.173	.174
79	168	.173	.179	.175	.179	.014	.180	.183	.185
80	91	.363	.357	.303	.299	.038	.343	.305	.279
81	3	.000	.000	.080	.080	.103	.031	.113	.169
82	11	.000	.000	.051	.051	.047	.017	.064	.096
83	7	.000	.000	.073	.074	.073	.029	.105	.157
84	9	.111	.183	.129	.175	.063	.183	.184	.185
85	55	.055	.055	.097	.097	.031	.062	.081	.094
86	32	.063	.063	.089	.089	.028	.068	.082	.092
87	94	.128	.126	.141	.140	.018	.128	.131	.133
88	44	.159	.146	.159	.149	.025	.147	.148	.149
89	102	.177	.173	.182	.180	.018	.177	.189	.198
90	78	.180	.188	.177	.183	.019	.187	.185	.184
91	167	.228	.234	.208	.213	.018	.228	.213	.203
92	121	.298	.297	.256	.255	.029	.286	.257	.238
93	12	.000	.000	.077	.075	.062	.030	.110	.165
94	14	.000	.000	.058	.059	.050	.022	.081	.121
95	19	.000	.000	.073	.072	.054	.025	.091	.136
96	11	.091	.156	.123	.166	.058	.164	.183	.197

Table 2. Proportions without health insurance coverage by age group and asian group: year 1997

	Raw1	Raw2 (wt)	HB1	HB2	EBLUP $\lambda=.1$	EBLUP $\lambda=.5$	EBLUP $\lambda=1$
0-17 years							
Total	.121	.115	.134	.129	.117	.122	.126
Asian 1	.151	.135	.149	.137	.135	.137	.138
Asian 2	.052	.054	.083	.083	.059	.075	.086
Asian 3	.232	.214	.207	.192	.207	.189	.177
Asian 4	.116	.113	.134	.131	.116	.123	.128
18-64 years							
Total	.202	.208	.191	.195	.205	.199	.194
Asian 1	.219	.217	.202	.201	.215	.209	.205
Asian 2	.130	.142	.138	.146	.142	.144	.145
Asian 3	.236	.242	.213	.216	.237	.223	.213
Asian 4	.220	.226	.206	.210	.223	.214	.208
65+ years							
Total	.028	.035	.080	.085	.053	.101	.134
Asian 1	.022	.024	.077	.078	.044	.096	.132
Asian 2	.024	.018	.072	.068	.033	.072	.100
Asian 3	.125	.147	.141	.155	.150	.159	.166
Asian 4	.023	.039	.081	.091	.059	.112	.148

References

1. Ghosh, M. and Maiti, T. (2003). Small area estimation based on NEF-QVF models and survey weights. Accepted for publication in *Biometrika.*
2. Godambe, V.P. and Thompson, M.E. (1989). An extension of quasi-likelihood estimation (with discussion). *Journal of Statistical Planning and Inference,* 22, 137-152.
3. Prasad, N.G.N. and Rao, J.N.K. (1990). The estimation of the mean squared error of small area estimators. *Journal of the American Statistical Association,* 85, 163-171.
4. Rao, J.N.K. (2003). *Small Area Estimation.* Wiley, New York.

[received date : Sep. 2007, accepted date : Oct. 2007]