POST-STRATIFICATION FOR PROPORTION ESTIMATION WITH USE OF A SELECTION WEIGHT

Yongxiao Wang, David Takeuchi, University of California, Los Angeles
Yongxiao Wang, UCLA, 405 Hilgard Ave., Los Angeles, CA 90024-1563

Key words: Post-stratification, Collapse, Survey

1.Introduction

The estimation of proportions such as the occurrence or non-occurrence of one or more mental disorders is often the first type of analysis conducted on community psychiatric epidemiologic survey data. Since community surveys are often based on complex sampling designs, statistical weights are used to adjust the prevalence estimators to insure that the results represent the target population. Post-stratification typically compares the known distribution of selected socio-demographic variables found in the target population with the distribution in the final survey sample.

In this paper, we consider a special case of post-stratification weighting where the combination of socio-demographic variables that are included in statistical adjustments create a large number of categories. Cells with small number of cases have the potential to inflate the standard errors for the prevalence estimators. Accordingly, we discuss a technique to reduce the number of post-strata; that is to collapse cells.

To illustrate our method, we analyses data from a large-scale study of Chinese Americans in Los Angeles County, the Chinese American Psychiatry Epidemiological Study (CAPES). An important aim of the investigation is to estimate the prevalence of selected mental disorders among Chinese Americans. In the first wave of the survey, 1747 household interviews with Chinese Americans residing in LA county were completed by September 1994. The research used a multi-stage sampling procedure to select respondents.

Two kinds of weights are considered in this paper. The first is a post-stratification weight that is applied to the post-strata levels and the second is a selection weight, which is the inverse of the number of eligible members in each household. The selection weight is applied to each respondent. We assume, for the purposes of this paper, that the selection weight is not associated with the post-stratification weight. This assumption is for convenience since the problem would be too unwieldy to handle in a paper of this length. We also assume that the target population is infinite. For CAPES, the population for the target population is about 169,000.

In sum, we address the following issues: (a) when and how should we collapse variables and cells to make post-stratification appropriate for the estimation, and (b) how do we evaluate the properties (e.g., bias and standard error of the estimator) of the estimator produced by the weighting methods.

2.When only post-stratification weight is considered

2.1.The proportion estimation

We first discuss the case when only a post-stratification weight is considered. Suppose a post-stratification produces H strata. There are n_i persons in post-stratum i, and total sample size is $n = \Sigma_{i=1}^{H} n_i$. Our goal is to estimate p, the prevalence of a certain disease in the target population. In fact, when only a post-stratification weight is concerned, we can write $p = \Sigma_{i=1}^{H} R_i p_i$, where R_i is the population proportion from a census and p_i is the prevalence of the disorder in the post-stratum i. The estimator of p_i is $\hat{p}_i = m_i / n_i$, the ratio of the number of disorders m_i and the number of the respondents n_i in the post-stratum i. The estimator of p under the post-stratification is $\hat{p}_{ps} = \Sigma_{i=1}^{H} R_i \hat{p}_i$.

In this paper, the variance for $\hat{p}_{ps}$ is a conditional variance given n_i which was suggested by Holt and Smith(1979)

$$Var(\hat{p}_{ps}) = \Sigma_{i=1}^{H} R_i^2 \hat{p}_i (1-\hat{p}_i)/n_i \qquad (2.1)$$

In the literature, an unconditional variance was also used (Cochran 1977, Kish 1965). To use the unconditional variance, one has to assume that the joint distribution of $(n_1, n_2, \ldots, n_H)$ is multinomial with cell probabilities $(R_1, R_2, \ldots, R_H)$. This assumption says that the non-response rates do not vary among the post-strata. In practice, it is common to find higher non-response rates in some post-strata. When the non-response rates are the same across the post-strata, the difference between the two variances is in order $O(n^{-2})$ which is very small when n is large.

2.2.Collapsing cells in post-stratification

Post-stratification may create a large number of strata. A significant portion of the post-strata may be unnecessary. Some cells may have a small number of respondents which can contribute to large standard

errors for the estimators. When the variations of the prevalence between cells are small, some post-strata may excessively increase the variances of the estimators. Collapsing the post-strata with similar prevalence is one method to overcome the deficiency. Strategies for collapsing cells have been previously discussed in the literature. Kalton and Maligalig (1991) discuss the criteria for collapsing cells to minimize the mean squared errors. They conclude that the collapsing should be done for those cells with small sample sizes. Little(1993) gives a Bayesian perspective aiming to reduce the posterior variance for continuous outcomes. While most of the literature on this issue have used continuous outcome measures, there have been some investigations that deal with proportion estimation; that is, when the outcome is binary. Trembly (1986), for example, uses a Chi-square statistic test to collapse cells, and the statistic has Pearson Chi-square-like form. His paper considers the proportion estimation but does not discuss the case where some cells have small sample sizes. In this paper, we focus on proportion estimation where the outcome is binary. One special feature for binary outcomes compared to continuous outcomes is that the variance of the proportion estimator is decided by the proportion.

The following claim leads to a criteria for collapsing post-strata. The proof of the Claim is based on simple algebra.

Claim: Under one post-stratification, if cells H-1 and H have the same rates, then to collapse the two cells will neither change the mean of the post-stratification estimator nor increase the variance (2.1).

One way to interpret the claim is, when two cells, say cells i-1 and i, have same proportions of the disease, the post-stratification of i-1 and i increases the variance of the estimator of prevalence when $n_{i-1}/n_i \neq R_{i-1}/R_i$. The amount of the improvement in variance is related to the difference between n_{i-1}/n_i and R_{i-1}/R_i. The greater difference it is, the greater improvement we get over the variance of the estimator.

The above results can be extended to multiple-cell cases. We should collapse a group of cells if these cells have the same proportions of the disorder.

Testing the equality of proportions from two or more binomial distributions has been discussed in literature(see Zar, 1984). Several statistical testing methods may be used here. We describe the methods and circumstances for selecting them.

(1).Pearson Chi-square test

Consider a 2xH crosstabulation, where H is the number of cells. Write the number of respondents with disorders in the first row and number of respondents without disorders in the second row. The classical Pearson Chi-square statistic can be used to test the association between the disease and the post-stratification. We modified the Pearson Chi-square statistic by incorporating the weighting scheme as follows.

Rewrite the classical χ^2 as

$$\chi^2 = \Sigma_{i=1}^{H}[(m_i-\hat{m}_i)^2/\hat{m}_i + (n_i-m_i-(n_i-\hat{m}_i))^2/(n_i-\hat{m}_i)]$$
$$= \Sigma_{i=1}^{H} n_i(m_i-\hat{m}_i)^2/[\hat{m}_i(n_i-\hat{m}_i)]$$

where $\hat{m}_i$ are estimators of m_i. The degrees of freedom for the χ^2 is H-1. The estimator of m_i, $\hat{m}_i$, under the post-stratification is $n_i\hat{p}_{ps}$. We also use $\hat{p}_i$ to substitute m_i/n_i. Thus the χ^2 can be written as

$$\chi^2 = \Sigma_{i=1}^{H} n_i(\hat{p}_i-\hat{p}_{ps})^2/[\hat{p}_{ps}(1-\hat{p}_{ps})]$$

For any post-stratification, a χ^2 statistic can be calculated.

To use the statistic as a criteria for collapsing, suppose we have a post-stratification scheme V_1. We collapse some cells of V_1 and get a post-stratification scheme V_2. The corresponding χ^2 statistics and degrees of freedom of V_k are $\chi^2_{(k)}$ and $df_{(k)}$ for k=1,2. Then $T_1 = \chi^2_{(1)} - \chi^2_{(2)}$ is a χ^2 statistic with degrees of freedom $df_{(1)}-df_{(2)}$. Collapse the cells if T_1 is not significant.

(2).Log-likelihood Ratio test

Log-likelihood ratio test can be applied to this problem. The advantage of the log-likelihood ratio test over the Pearson Chi-square test is that the former does not require m_i's to be large. It requires only that n_i in the strata that will be collapsed are large. We derive the test as follows.

When we treat n_i as given, the likelihood function for the data is

$$\prod_{i=1}^{K}\binom{n_i}{m_i}p_i^{m_i}(1-p_i)^{(n_i-m_i)}$$

Since p_i are independent parameters, the maximum likelihood estimator of p_i is m_i/n_i. If we collapse some of the cells, then we get a new post-stratification with K_1 ($\leq$ K) cells. The maximum likelihood estimators for p_i' are m_i'/n_i', where (n_i', m_i', p_i') is the number of respondents, the number of respondents with a disorder, and the assumed prevalence in ith cell respectively for i=1,..,K_1. The log-likelihood ratio test statistic is

$$T_2 = 2\{\Sigma_{i=1}^{K}[m_i\log(m_i/n_i) + (n_i-m_i)\log(1-m_i/n_i)]$$
$$- \Sigma_{i=1}^{K_1}[m'_i\log(m'_i/n'_i) + (n'_i-m'_i)\log(1-m'_i/n'_i)]\}$$

which is χ^2 distributed with the degrees of freedom K-K_1.

(3).Test for equality of two proportions

If we want to collapse two cells with the number of respondents, the number of disorders and prevalence assumed to be (n_i, m_i, p_i), i=1,2. We do the hypothesis test $p_1 = p_2$.

(3.1).Normal approximation

When the numbers of respondents in both cells, n_1 and n_2, are large, the normal approximation method can be applied. The test statistic is

$$T_3 = |m_1/n_1 - m_2/n_2| / \sqrt{\bar{p}\bar{q}(1/n_1 + 1/n_2)}$$

where $\bar{p} = (m_1 + m_2)/(n_1 + n_2)$ and $\bar{q} = 1 - \bar{p}$. When both n_1 and n_2 are large, the statistic T_3 is approximately normally distributed. It is a two-tail test.

(3.2).Exact test for equality of two proportions

When at least one of n_1 or n_2 is small, we can use a statistic whose p-value for the test can be directly calculated. Assume that n_1 is small (say, less then 10). If the observed difference of two proportions is a, then the two-tail p-value of the test of testing $p_1 = p_2$ is

$$\begin{aligned} &\max_{0\le p\le 1} P\{\left|\frac{m_1}{n_1} - \frac{m_2}{n_2}\right| > |a|\} \\ &= \max_{0\le p\le 1} \sum_{j=0}^{n_1} P\{\left|\frac{m_2}{n_2} - \frac{j}{n_1}\right| > |a|\} P\{m_1 = j\} \end{aligned} \qquad (2.2)$$

where m_2 is a binomial variable with (n_2, p_1). Each of the probabilities in the formula above can be directly calculated if n_2 is also small, or computed by using normal approximation if n_2 is large.

Circumstances to apply the suggested testing methods

Each of the statistics we have suggested above has some advantages and disadvantages under different circumstances. Only when the number of respondents with a disorder m_i and the number of respondents without a disorder n_i-m_i are not too small, the Pearson Chi-square test T_1 can be used. Empirical results suggest that all the numbers in the cells should not be smaller than 5. In some studies such as the CAPES, the number of respondents with a disorder in cells after original post-stratification are typically very small. In such case, the use of the statistic becomes problematic. At the later stage of collapsing, on the other hand, since the numbers of respondents with a disorder usually are increased after combining the primary post-stratification cells, the condition requiring both the numbers of respondents with and respondents without a disorder in each stratum not to be smaller than 5 is more likely to be satisfied. The advantage of the Pearson Chi-square test is that it incorporates information from all the cells and not just the information from the cells that are to be collapsed. Moreover, the Pearson Chi-square test can be extended to the case when some other weights, such as selection weight, are incorporated in the estimation procedure.

To use the log-likelihood ratio test T_2, we need to assume that the numbers of respondents in the cells to be collapsed are large. When only two cells are to be collapsed, the tests suggested in (3) are simple and effective.

Example:

We use the CAPES to show the strategies of collapsing cells. We want to estimate the overall lifetime prevalence of "Any Anxiety Disorder" which is a combination of the following DSM-III-R diagnoses: Generalized Anxiety Disorder, Simple Phobia, Social Phobia, Panic Disorder, Agoraphobia and Post-Traumatic Stress Disorder. The socio-demographic variables we consider here are sex, age, and education level. The detailed definition of the variables and the distributions of these variables in both our sample and the Chinese American population in LA county according to the 1990 U.S. census are shown in Table 1.

Table 1. Comparison of the socio-demographic variables in the sample and Chinese population in LA county from 1990 Census

		Sample	1990 Census
		(in percent)	
Sex	Male	47.8	47.8
	Female	52.2	52.2
Age	18-24	13.8	16.4
	25-34	25.6	29.5
	35-44	31.7	27.6
	45.54	17.4	14.5
	55-65	11.5	12.0
Education level(Year in School)			
	1 (≤11)	18.3	22.7
	2 (=12)	18.9	15.7
	3 (13-15)	31.7	27.6
	4 (≥16)	42.5	36.1

By the nature of this kind of study, we usually like to "force" some variables to be stratified, such as sex. Throughout the example, we will not collapse for sex. Because of the concern of the length of the paper, in Tables 2 and 3, we only show the data from the female group. The actual technique is applied to both gender

groups. The criteria for collapsing cells are in either one of two situations: 1)the cells have a small number of respondents, say smaller than 10; 2)the cells have very similar prevalence rates. The results of collapsing cells are shown in Table 2 and the results of the hypotheses testing Log-likelihood ratio and the Pearson Chi-square tests are shown in Table 3.

Table 2(a) is the result of original post-stratification. We collapse (A) education levels 3 and 4 to form a new education group 3-4; (B) age groups 45-54 and 55-65 to form a new age group 45-65.

The p-values for the two tests are above 0.3, suggesting that the collapsing steps are justified. The result after collapsing steps (A) and (B) is shown in Table 2(b).

Having collapsed the categories for variables, we consider collapsing cells. The result of using statistical tests is shown in the third to sixth rows in Table 3. The normal approximate test T_3 is only for testing the equality of two proportions. We see that the p-values from the log-likelihood ratio and the direct approximation tests are close while the Pearson Chi-square test gives different p-values.

The final result after the post-stratification and the collapsing steps is shown in Table 2(c). In the female group, there are five cells which form the final post-stratification. We see that it is much easier and clearer to interpret the results of the study after collapsing than under the original post-stratification. We can draw three important conclusions from the tables. First, respondents with low education have higher rates of the disorder. Among the respondents with low education, older people may have a slightly higher rate of the disorder. Second, respondents who are older than 25 and have at least 12 years of education have a low prevalence. Third, there are some interactions among the socio-demographic variables in the prevalence estimates. Similar conclusions are also found for the male group.

The estimators of the overall prevalence and their standard errors are listed in Table 4. We do not see much difference among the estimators. As suggested earlier, since the distribution of the socio-demographic variables is quite similar to the target population, we would not expect much change in the prevalence estimators after post-stratification adjustments. We do expect, however, that collapsing cells will result in smaller estimated standard errors.

We see from the example that selecting cells for collapsing is somewhat subjective. The preference is given to the post-stratification scheme that is easily interpretable. The criteria for accepting the statistical testing result may vary, say one may like to choose the criterion as 0.10 or 0.15. We think that we should consider both the p-values and the nature of a study to decide what to be collapsed.

Table 2. Steps of collapsing post-strata(Female only)

(a)

Educ|age	18-24	25-34	35-44	45-54	55-65
1	10.0% 10	9.5% 42	13.1% 61	15.4% 39	10.2% 49
2	0.5% 21	0% 37	4.9% 61	5.1% 39	8.8% 34
3	10.5% 57	2.5% 40	9.7% 62	0% 25	0% 6
4	12.5% 32	5.7% 105	2.7% 111	8.1% 62	5.9% 17

(b)

Educ|age	18-24	25-34	35-44	45-65
1	10.0% 10	9.5% 42	13.1% 61	12.5% 88
2	9.5% 21	0% 37	4.9% 61	6.8% 73
3-4	11.2% 89	4.8% 145	5.2% 173	5.5% 110

(c)

Educ|age	18-24	25-34	35-44	45-65
1	9.6% 52			12.8% 149
2	10.9% 110	0% 37		
3-4				5.3% 562

In each cell, the percentages are the prevalence, the second row is the numbe of respondents

3.When a selection weight is used

3.1.The estimation

We now consider the case with the use of a selection weight. In the CAPES, for example, only one person was interviewed from each household with eligible members, so the respondents in the sample do not have the same chance of selection. One usual way to adjust for this selection bias is to use a weight u_{ij} which is proportional to the inverse of the number of the eligible member in the household. Under the assumption of independence between the socio-

Table 3. Results About the Testing

Collapsing Steps	T_1 (df) p-values	T_2 (df) p-values	T_3 p-values
(A)	10.15(10) 0.428	11.52(10) 0.318	
(B)	3.713(6) 0.715	4.006(6) 0.676	
(1,1) and (1,2)	0.000(1) 1.000	0.002(1) 0.964	0.046 0.963
(2,1) and (3-4,1)	0.133(1) 0.715	0.053(1) 0.818	0.226 0.821
(2,3),(2,4),(3-4,2), (3-4,3) and (3-4,4)	0.501(2) 0.778	0.412(2) 0.814	
(1,3) and (1,4)	0.033(1) 0.856	0.012(1) 0.912	0.111 0.912

The indicators in the first collum are: (Education level, Age). Education level=1, 2, and 3-4; age=18-24(1), 25-34(2), 35-44(3) and 45-65(4).

Table 4. The Estimators of Overall Prevalence rates with Their Standard Errors(in percent)

Scheme	Est. of p	SE
Unweighted	5.953	0.566
At original poststratification	6.138	0.615
At final poststratification	6.265	0.595
At final poststratification with a selection weight	6.059	0.652

demographic variables and the selection weight, we assume that $\Sigma_{j=1}^{n_i} u_{ij} = 1$. The estimator of the proportion in the post-stratum i is $\hat{p}_i^{(s)} = \Sigma_{j=1}^{n_i} u_{ij} X_{ij}$, where $X_{ij}=1$ if the respondent j in the post-stratum i is a disorder and $X_{ij}=0$ if not. The estimator of the prevalence for the overall target population is $\hat{p}_{ps}^{(s)} = \Sigma_{i=1}^{H} R_i \hat{p}_i^{(s)}$. The difference between $\hat{p}_{ps}^{(s)}$ and $\hat{p}_{ps}$, the estimator for which only the post-stratification weight is applied is that $\hat{p}_i^{(s)}$ replaces $\hat{p}_i$. The difference between $\hat{p}_i^{(s)}$ and $\hat{p}_i$ is that a selection weight is applied to the former while the latter is simply the proportion of disorders sampled in post-stratum i.

The mean and variance of $\hat{p}_{ps}^{(s)}$

To derive the expression and the estimators for the mean and the variance of $\hat{p}_{ps}^{(s)}$, we make two assumptions. As mentioned before, we first assume that n_i, the numbers of respondents for the post-stratum i, are given. Thus, we only need to focus on the mean and variance for $\hat{p}_i^{(s)}$ which involves the selection weight only. The second assumption is that the selection weight does not relate to disease status. This assumption may be violated in some cases, but without the assumption the derivation can be quite intricate in terms of expression and notation. On the other hand, in practice, it is rare to assume the selection weight is associated with disorder status when the expressions of the mean and the variance for the estimator are derived.

When deriving the mean and variance for the estimator, we consider two cases by treating the selection weight u_{ij} either as given or random.

(i).When the selection weight is given

When u_{ij} are given, the derivation is straight forward.

The mean of $\hat{p}_i^{(s)}$ is

$$E(\hat{p}_i^{(s)}) = \Sigma_{j=1}^{n_i} u_{ij} E[X_{ij}].$$

Here, we have $E[X_{ij}]=p_i$ since we assume that the disease status does not relate to the selection weight. Since we have also assumed $\Sigma_{j=1}^{n_i} u_{ij} = 1$, then $\hat{p}_i^{(s)}$ is an unbiased estimator of p_i which implies that $\hat{p}_{ps}^{(s)}$ is an unbiased estimator of p.

Under the same assumption, the variance of $\hat{p}_{ps}^{(s)}$ is

$$Var(\hat{p}_{ps}^{(s)}) = \Sigma_{i=1}^{H} R_i^2 p_i(1-p_i) \Sigma_{j=1}^{n_i} u_{ij}^2.$$

Thus, the standard error of $\hat{p}_{ps}^{(s)}$ is the square root of

$$\Sigma_{i=1}^{H} R_i^2 \hat{p}_i^{(s)}(1-\hat{p}_i^{(s)}) \Sigma_{j=1}^{n_i} u_{ij}^2. \qquad (3.1)$$

(ii).When the selection weight is random

When u_{ij} are random, we can assume that for each i, u_{ij} are identically distributed. Since we have already assumed that $\Sigma_{j=1}^{n_i} u_{ij} = 1$, then u_{ij} are not independent variables. The mean of u_{ij} is $1/n_i$.

It is not difficult to conclude that $\hat{p}_{ps}^{(s)}$ is an unbiased estimator of p, since we only need to show that $\hat{p}_i^{(s)}$ is unbiased for p_i

$$E(\hat{p}_i^{(s)}) = E\left[\Sigma_{j=1}^{n_i} E(u_{ij}X_{ij}|u_{ij})\right] = p_i E\left[\Sigma_{j=1}^{n_i} u_{ij}\right] = p_i.$$

Notice that

$$Var\left[E\left(\Sigma_{j=1}^{n_i} u_{ij}X_{ij}|u_{ij}\right)\right] = p_i^2\, Var\left[\Sigma_{j=1}^{n_i} u_{ij}\right] = 0.$$

We have

$$Var(\hat{p}_i^{(s)}) = E\left[Var\left(\Sigma_{j=1}^{n_i} u_{ij}X_{ij}|u_{ij}\right)\right]$$
$$= p_i(1-p_i)E\left[\Sigma_{j=1}^{n_i} u_{ij}^2\right] = n_i p_i(1-p_i)E\left[u_{il}^2\right].$$

which implies that the standard error of $\hat{p}_{ps}^{(s)}$ is also the square root of (3.1). Thus whenever u_{ij} are treated as given or random, the two estimators of the variance of $\hat{p}_{ps}^{(s)}$ are the same.

Example(cont'd.):

Following the example in Section 2.2, we calculate the estimator of the mean and standard error with the use of the selection weight. The post-stratification scheme in Table 2(a) is used here. The result is in the last row of Table 4. As expected, the standard error of the estimator is increased by using the selection weight.

3.2. Collapsing cells when the selection weight is used

Since collapsing cells is the technique to reduce unnecessary post-strata, one can imagine that the selection weight may have little to do with it. Especially when the numbers of respondents in the involved cells are small, the prevalence rates in those cells are typically very sensitive to the selection weight. When the numbers of respondents in post-strata are large, it is possible to extend the Pearson Chi-square test to the case though we believe that for most cases, it will not make much of a difference.

In the case where a selection weight is used, the estimator of the prevalence in each of the stratum is changed from $\hat{p}_i$ to $\hat{p}_i^{(s)}$. The argument for collapsing cells is parallel and the change to the χ^2 in Section 2 is to use $\hat{p}_i^{(s)}$ and $\hat{p}_{ps}^{(s)}$ replacing $\hat{p}_i$ and $\hat{p}_{ps}^{(s)}$ respectively. The modified χ^2 is now

$$\chi^2 = \sum_i n_i(\hat{p}_i^{(s)} - \hat{p}_{ps}^{(s)})^2 / \left[\hat{p}_{ps}^{(s)}(1-\hat{p}_{ps}^{(s)})\right]$$

4. Some Remarks

We highlight two points about post-stratification here. First, we should only do post-stratification for socio-demographic variables that have some association with an outcome. This recommendation is made to avoid (a) over-adjustments that will increase the standard error even though it will not generate a biased estimator, and (b) having cells with small numbers of respondents. Second, collapsing cells with the same proportion will improve the precision of the estimator. The improvement in precision will be most evident when the difference between the proportions of respondents in the sample from each post-stratum and the target population is large.

The claim in Section 2 can be extended to the case of estimating the mean of a continuous outcome. The condition for collapsing two or more cells then will be that the cells have same means and variances. In this condition, the condition of having same means is to make sure that the estimator after the collapsing step is still unbiased while the condition of having same variances is to ensure a smaller standard deviation of the estimator. Note when the outcome variable is normally distributed in those cells, the condition of having the same means and variances is equivalent to the condition of having the same distribution.

The collapsing step for the post-stratification is not very meaningful for estimating the prevalence of some rare disease, because the number of persons with a disorder is too small. For example, in the CAPES with 1747 respondents, some disorders may have a prevalence of about 1-2%. The estimators can be very sensitive to general weighting methods.

Acknowledgements

Support for this research was provided by NIMH Grants #47460 and 44331.

Reference

1. Cochran W.G. (1977) "Sampling Technique(3rd ed.)", New York. John Wiley.
2. Holt, D. and Smith, T.M.F., (1979) "Post Stratification". Journal of the Royal Statistical Society, Ser. A., 142, 33-46.
3. Kalton, G. and Magligalig, D.S. (1991) "A comparison of methods if weighting adjustment for nonresponse", Proceedings of the 1991 Annual Research Conference, U.S. Bureau of the Census, 409-428.
4. Kish L., (1965) "Survey Sampling". New York. Weily.
5. Little, R.J.A. (1993) "Post-stratification: A modeler's Perspective". Journal of the American Statistical Association, 88, 1001-1012.
6. Tremblay, V. (1986), "Practical Criteria for Definition of Weighting Classes". Survey Methodology, 12, 85-97.
7. Zar, J.H. (1984) Biostatistics Analysis, 2nd edition. Prentice-Hall, Inc., N.J.