RESIDUAL WEIGHTING IN SURVEY SAMPLING

Yiannis C. Bassiakos, Department of Economics, University of Athens, Chrissoula Preza, VPRC Institute
Yiannis C. Bassiakos, Department of Economics, 8 Pesmazoglou St., 105 59,Athens, Greece

Key Words: survey sampling, weights

Introduction

Public Opinion polling has become an important part of the political process in modern democratic states, especially in Europe, where the timing of the elections is usually the prerogative of the siting government. A critical component of any poll is the sample selection process. The most common sample selection method is multi-stage stratified cluster sampling. (msStCS). The method consists of the following steps: (a) the target population is separated in strata according to a number of criteria, e.g. urbanity (urban-suburban-rural), gender (male-female), geographical region of the country etc. Information about the total population in each stratum is obtained from the most recent national census. Then clusters are defined, usually city blocks or other dwelling agglomerations. These clusters are the actual sampling units. Within each stratum, a number of clusters is selected, with selection probability proportional to the size of the cluster. From the sampled clusters households are selected in a systematic fashion and finally individuals, using simple random sampling from the total number of the members of the household. The whole sampling process is designed so that it yields self-weighting estimates of proportions.

The stratified sampling (StRS) method is preferred because the categorical variables, which are used in the stratification, are believed to be related to the dependent variable, vote intention. Indeed, as political scientists affirm, gender, urbanity and geographical region play a role in the voting pattern. The stratification process is either implemented through the definition of the size of the sample within a particular stratum or through quota sampling within a particular cluster (for example, selection of a predetermined number of women with in a particular cluster). The usual sampling method in this case is the selection with simple random sampling (SRS) sub-samples from each stratum with size proportional to that of the stratum in the population. The proportional selection leads to estimators with variance smaller than that obtained with SRS from the population.

What complicates thing more is that sometimes there exist other variables, most notably prior voting history, education and socio-economic status, which also play an important role in the voting decisions. These variables cannot be used for stratification purposes, since information on the size of the relevant strata prior to the actual sample selection is not available. Further more, even for the variables for which information is available, non-response leads to actual samples which differ from their designed counterparts.

A solution to this problem is the introduction of weights. As Danny Pfeffermann (1993) observes, "Sampling weights weigh sample data to correct for the disproportionality of the sample with respect to the target population of interest. The weights reflect unequal sample inclusion probabilities and compensate for differential non-response and frame under-coverage". A common weighting method is defining proportional weights (C.H. Alexander, 1987). More specifically let a population of size $N = A + B$, where A, B the sizes of two strata. Let also a sample of size $n = a + b$ selected from the population with StRS, and a, b the sizes of the sub-samples selected from strata A and B, respectively, so that $a/A = b/B = n/N$, i.e. with proportional representation of each stratum in the sample. This method yields self-weighting estimators, which have smaller variances than the ones produced with SRS, as mentioned before. Let again a different sample selected with StRS, with size $n' = a' + b'$, $n' = n$, where a', b' the sizes of the sub-samples from each stratum. If, in this case, $a'/A \neq b'/B$ i.e. there isn't proportional representation of each stratum within the sample, we can re-establish the proportional representation by multiplying each observation from stratum A with a/a' and each observation from stratum B with b/b'. This is the ratio weighting method.

Usually, the correction obtained by weighting is insignificant, since SRS tends to produce samples, which are already nearly correctly distributed across the strata. When a gross difference from the correct distribution across strata is present, as in the case of informative non-response, or when the samples are selected on purpose divergent from the correct distribution to obtain a minimum necessary sample size within each stratum for estimation purposes, the impact of weighting can be much more pronounced. In this case multiple weighting can lead to improvement or deterioration of the proportion estimates with respect to their bias or/and variance, (J.W. Choi, 1996). The underlying reason for this outcome is the inter-relation of the weighting variables. More specifically the relation between voting history and the rest of the variables is as strong as the relation between the current voting decision and the rest of the variables. If the relation could lead to "total explanation" i.e. if knowledge of gender, urbanity etc could allow exact prediction of the voting history, then weighting by voting history "corrects" the imbalances in the rest of the variables. When multiple weighting is used the result is "over-correction".

Proposed Method

In the present paper a method is proposed to face this problem of "over-correction". The method is inspired from an idea used in Nutritional Epidemiology, where when an independent variable has strong influence to both the dependent and the rest of the independent variables they are regressed on the influential independent variable and their regression residuals are used as new independent variables. The proposed method will be referred to as "residual weighting". The following steps describe the method:

Step 1: weigh all observations by the variable with the highest degree of dependence with the variable to be predicted.

Step 2: weigh by the variable with the second highest degree of dependence with the variable wishing to predict only the observations, which are misclassified by the log-linear model of the variable to be predicted on the first weighting variable.

Application

The following example is based on data from an exit-poll conducted in Greece for the 1996 national elections, on behalf of the National Television, by VPRC Institute, a public opinion survey research company. The data set contained 10486 voters of which 10482 had full data and were used for the analysis, with information on current vote, vote in the prior election (1993), gender, age, education level, location of residence (urban-suburban-rural) and part of the country the voter voted at. For the present problem we discuss only the prior vote and gender. Two different applications of the method are presented: 1. Five different cases of 100 replications of samples of size 500, where the post-stratification variables (prior vote and gender) are independent. 2. Four different cases of 100 replications of samples of size 500, where the post-stratification variables are dependent. The full data set is treated as the population and the samples as results of polling. In the dependent case only 8919 cases are used, since the people who did not vote in the 1993 election are excluded from the analysis. The results are classified according to a χ^2 goodness-of-fit criterion.

While running a regression (log-linear model) it was observed that the misclassified observations corresponded to the "vote transfers" i.e. the subjects, which in the preceding election voted for party 1 and in the upcoming election voted for party 2. This was a natural result since voting history is the best predictor of vote intentions, as political scientists affirm. Accordingly, the vote transfers were considered "residuals" for the purpose of this analysis.

The argument of the applicability of the method is as follows. If the proportion of transfer votes is the same for each party within each category of the rest of the explanatory variables, then weighting based on the vote in the preceding election could be enough to correct the imbalance of the sample. As a matter of fact, any additional weighting might cause the deterioration of the estimates. But if there are different proportions of transfer votes within some categories of the explanatory variables, then additional weighting is necessary to correct the imbalance of vote transfers within each category. The problem is that the proportion of vote transfers within each category is not known in the population. As a result the weights must be calculated using the marginals of the category. One could use prior knowledge from other polls to estimate the missing proportions but such information could be biased and its validity cannot be checked.

The main comparison is between the residual method and the product method. The main reason for this comparison is that pollsters usually try to balance all explanatory variables, which show discrepancies with their known population values. Consequently the most common mistake is over-weighting.

Table 1 below contains information about the distribution of vote93 and gender for the first example. Note that party6 did not participate in the 1993 election

TABLE 1. Vote93 by Gender (population)

party	gender		total
	male	female	
party1	54,7%	45,3%	39,2%
party2	51,2%	48,8%	29,9%
party3	54,4%	45,6%	3,3%
party4	59,5%	40,5%	5,1%
party5	47,5%	52,5%	4,2%
party6	n/a	n/a	n/a
party7	62,2%	37,8%	2,6%
party8	61,3%	38,7%	1,4%
no answer	48,2%	51,8%	7,1%
didn't vote	48,4%	51,6%	7,2%
TOTAL	53,0%	47,0%	100,0%

From this population, samples of size 500 were selected using StRS, with the following distribution for the variable vote93 and gender, within each sample, as listed in Table 2.

TABLE 2. Vote93 by Gender (samples)

party	gender		total
	male	female	
party1	60,0%	40,0%	30,0%
party2	60,0%	40,0%	40,0%
party3	60,0%	40,0%	4,0%
party4	60,0%	40,0%	6,0%
party5	60,0%	40,0%	3,0%
party6	n/a	n/a	n/a
party7	60,0%	40,0%	4,0%
party8	60,0%	40,0%	2,0%
no answer	60,0%	40,0%	6,0%
didn't vote	60,0%	40,0%	5,0%
TOTAL	60,0%	40,0%	100,0%

As it can be seen from the two tables the samples had forced upon them a deviation from the population, so that weighting was necessary. The sampling process was repeated using 5 different vote transfer proportions from Party 1 to Party 5. The five different vote transfer proportion cases in the samples were ordered from the bigger to the smaller deviation from the true vote transfer proportion. For each vote transfer proportion case, 100 samples were drawn and four different weighting methods were used. One using the proposed residual weighting method, one using the product of weights method, one using gender alone and finally one using vote in the prior election (vote93) alone. The difference of the estimated proportion for each party from the true proportion was computed using a χ^2 statistic for goodness of fit for each weighting method. The values of the χ^2 statistic were used to select the best method in each replication. Table 3 presents the vote transfers between the two elections in the population, by gender.

TABLE 3. Vote93 moves by Gender (population)

party	gender		total
	male	female	
party1	17,5%	15,1%	16,4%
party2	10,4%	9,9%	10,2%
party3	55,1%	63,7%	59,0%
party4	11,4%	12,5%	11,8%
party5	21,4%	26,3%	24,0%
party6	n/a	n/a	n/a
party7	33,9%	51,0%	40,4%
party8	42,4%	53,5%	46,7%
no answer	n/a	n/a	n/a
didn't vote	n/a	n/a	n/a

The next table, Table 4, presents the imbalances of the five vote transfer proportion cases (from party 1) in the samples:

TABLE 4. Moves from party1 by Gender (samples)

case	gender		total
	male	female	
case1	21,1%	18,3%	20,0%
case2	20,0%	20,0%	20,0%
case3	18,9%	21,7%	20,0%
case4	17,8%	23,3%	20,0%
case5	16,7%	25,0%	20,0%
(true value)	*17,5%*	*15,1%*	*16,4%*

As it can be seen from Tables 5 and (especially) 6 the residual weighting outperforms product weighting in the first four cases (biggest deviations), the gender weighting in all cases and the vote93 weighting in the last two.

TABLE 5. Times each weighting method was better overall

case	weighting method			
	residual	product	gender	vote
case1	3	2	3	92
case2	8	4	9	74
case3	13	18	3	59
case4	30	30	5	51
case5	26	57	5	35

TABLE 6. Pairwise comparisons

case	weighting method		
	residual vs product	residual vs gender	residual vs vote
case1	86	87	4
case2	74	86	12
case3	59	96	28
case4	51	94	60
case5	35	94	84

The forced independence of prior vote and gender has lead to independence of current vote and gender in the samples. As a result in the first three cases weighting should be based on prior vote alone. Product weighting is actually leading to over-weighting. As the dependence between prior vote and gender strengthens the product based weight improves, as does the residual weight, which outperforms the product weight in the first four cases.

In the second example Table 7 contains the distribution of vote93 and gender.

TABLE 7. Vote93 by Gender (population)

party	gender		total
	male	female	
party1	54,7%	45,3%	45,7%
party2	51,2%	48,8%	35,0%
party3	54,6%	45,4%	3,8%
party4	59,5%	40,5%	6,0%
party5	47,5%	52,5%	5,0%
party6	n/a	n/a	n/a
party7	61,7%	38,3%	4,5%
TOTAL	53,7%	46,3%	100,0%

From this population, 500 observations were selected using StRS, with the following distribution for the variable vote93 and gender, within each sample, as listed in Table 8.

TABLE 8. Vote93 by Gender (samples)

party	gender		total
	male	female	
party1	60,0%	40,0%	36,0%
party2	40,0%	60,0%	40,0%
party3	60,0%	40,0%	6,0%
party4	60,0%	40,0%	6,0%
party5	60,0%	40,0%	4,0%
party6	n/a	n/a	n/a
party7	60,0%	40,0%	8,0%
TOTAL	52,0%	48,0%	100,0%

The sampling process was repeated using 4 different vote transfer proportions from Party 1 to 5. The 4 vote transfer proportion cases in the samples were ordered from the bigger to the smaller deviation from the true proportion. Tables 9 and 10 present the vote transfers between the two parties in the population and the samples, respectively, by gender.

TABLE 9. Vote93 moves by Gender (population)

party	gender		total
	male	female	
party1	17,0%	14,4%	15,8%
party2	9,8%	9,5%	9,6%
party3	54,8%	63,2%	58,6%
party4	11,1%	12,1%	11,4%
party5	21,4%	26,3%	24,0%
party6	n/a	n/a	n/a
party7	35,5%	50,0%	41,0%

TABLE 10. Moves from party1 by Gender (samples)

case	gender		total
	male	female	
case1	40,74%	9,7%	28,3%
case2	22,22%	9,7%	17,2%
case3	17,59%	9,7%	14,4%
case4	12,96%	9,7%	11,7%
(true value)	*17,0%*	*14,4%*	*15,8%*

As it can be seen from Tables 11 and (especially) 12 the residual weighting outperforms product and gender weighting in all cases and the vote93 weighting in the last two.

TABLE 11. Times each weighting method was better overall

case	weighting method			
	residual	product	gender	vote
case1	40	1	0	59
case2	25	0	0	75
case3	64	0	0	36
case4	76	0	7	17

TABLE 12. Pairwise comparisons

case	weighting method		
	residual vs product	residual vs gender	residual vs vote
case1	98	100	40
case2	100	100	25
case3	100	100	64
case4	100	93	81

Conclusion

The residual method outperforms the product weighting method under severe imbalance conditions. It will be the subject of further research to describe the typology of the conditions under which each weighting method performs better. Such a typology can be a useful tool in the hands of professional pollsters to help them produce results with the maximum possible accuracy.

It remains to be seen if the various weighting methods in general can be improved to overcome such disadvantages as the poor quality of data on the necessary categorical variables.

References

Alexander C.H. (1987), *A Model Based Justification for Survey Weights*, Proceedings of the Section on Survey Research Methods, ASA, pp. 183-188.

Choi J.W. (1996), *The Weighting and Variances of Survey Data*, Proceedings of the Section on Survey Research Methods, ASA, pp. 864-869.

Pfeffermann D. (1993), *The Role of Sampling Weights when Modeling Survey Data*, International Statistical Review, 61, pp. 317-337.

Zeischang D.K. (1990), *Sample Weighting Methods and Estimation of Totals in the Consumer Expenditure Survey*, Journal of the American Statistical Association, No 412, pp. 986-1001.