The Korean Communications
in Statistics Vol. 6, No. 3, 1999
pp. 761-770

Estimation of Transition Probability on Two Successive Occasions Sampling with Randomized Response Technique

Kay-O Lee[1)]

Abstract

A combination procedure of successive occasions sampling and randomized response method is investigated. Randomized response technique is very simple for use in a telephone survey of a sensitive subject. In the suggested randomized response method, the interviewee replies "yes" or "no" to a randomly selected question and the investigator can estimate the proportion of "yes" or "no" answer. When this procedure is used on successive occasions, not only the proportion supporting a candidate and the time change in this supporting proportion can be derived, but also, the voters' swing in the trend of voters' support can be estimated. A numerical example is given to show how the suggested sampling strategy can be applied to a practical telephone survey.

1. Introduction

Some surveys of public opinion are often performed successively and response rates are very low when the interviewee is asked sensitive issues directly. Especially in telephone survey of the election campaign, a voter does not like to answer when he is asked a direct question and the response rate of such a survey is lower than 20%. Since the survey is executed by telephone, the procedure of data collection should be simple and privacy is to be protected.

In an election campaign, the proportion of people supporting a candidate is frequently changed, so the trend of public opinion is to be closely followed; therefore investigations are repeatedly made on successive occasions with randomized response techniques so as to setup appropriate campaign strategies.

Since the estimation of changes of supporting proportions is very important in an election campaign, it is valuable to study the new sampling scheme, that combines successive occasion sampling theory with a randomized response technique.

Though we often followed randomized response technique, successive occasions sampling

1) Professor, Department of Computer Science and Statistics, Korea Air Force Academy, P.O.Box 335-2, Chongwon, ChungBuk 363-849.

scheme with randomized response technique has not been studied in the literature. In particular, its study would be very useful to analyze the effect of programs and proposals on sensitive characteristics and to estimate changes of supporting proportion.

First of all, a new sampling scheme has been introduced for two successive occasions with randomized response technique for estimating supporting proportion. In section 3, estimation of voter's transition probability on two successive occasions is studied. In section 4, results of studies are evaluated by numerical examples.

2. Sample Selection Scheme

When the subject of a study is more sensitive, people are less cooperative with direct questions that deal with social stigma or privacy. We should be very careful in managing the sample units and in selecting the matched sample units. The whole process of investigation should be simple and protection device of privacy should be transparent to the interviewee. Hence, the data on successive occasion sampling are collected by using Warner's type randomized response technique. Sample selection scheme is carried out as follows.

On the first occasion, the sample of n units is randomly selected and is investigated by randomized response technique. In an example of telephone survey, the frequency table of birth-month can be utilized as a randomizing device. On the second occasion, before selecting sampling units, we divide the first occasion's sample into two groups; one group is composed of sampling units that reply "yes", the other consists of sampling units that reply "no". Sample selection procedures consist of three random processes:

(1) $u((1-\mu)n)$ unmatched sampling units are chosen independently from the N-n remaining sample units, which were not observed on the first occasion.

(2) $m(\mu n)$ matched sampling units are drawn at random and without replacement from the first occasion's sample as follows ; $m_0(\mu n_0)$ sampling units are chosen from the "no" group of the first occasion's sample with size n_0 and $m_1(\mu n_1)$ sampling units are drawn from the "yes" group of the first occasion's sample with size n_1. Assume that sample size n is kept for all occasions to be the same and the proportions of replacement sampling units are also the same for all occasions. Furthermore, the sampling fractions of matched samples are kept for two groups of the first occasion's sample to be the same. The above sample selection procedures can be extended to the third occasion or further occasions as follows. Divide the second (the previous) occasion sample into "yes" and "no" groups according to respondent's answer, and apply the (2) step of sample selection process.

3. Estimation of Transition Probabilities

On a public opinion survey in the election campaign, we have the interest in estimating the proportion of supporting candidate as well as measuring the voter's transition probability on successive occasions. In this section we will study the estimation of transition probability on the basis of the collected data from the given sample selection scheme.

The investigation on the first occasion is performed by using a randomizing device which indicates the probability(P) of choosing a sensitive question. The value of P is determined before investigating on the basis of the frequency table of birth-month. For example, when the relative cumulative frequency of voters, whose birth-month is March, May or November, is 0.25, the value of P is 0.75. If an interviewee's birth-month is March, May or November, then the interviewee replies "yes" or "no" to the following statement " I will not vote for the candidate A." and if an interviewee's birth-month is not March, May or November, the interviewee replies "yes" or "no" to the statement " I will vote for candidate A."

On the second occasion, after randomizing device is reconstructed so that the probability of choosing a sensitive question is the same as P, matched and unmatched samples are investigated by the same method as that of the first occasion. Therefore, the observed data are expressed by three parts; one is collected only on the first occasion, another is based on both successive occasions, and the third is observed only on the second occasion.

The observed variable X_{kj} that corresponds to the answer of j th respondent on the kth occasion is defined as follows.

$$X_{kj} = \begin{cases} 1 & \text{if } j\text{th respondent replies "yes"} \\ 0 & \text{if } j\text{th respondent replies "no"} \end{cases} \quad (1)$$

The notations to be used in this section are given as follows.

Π_{kA} = the probability of the event that voters support candidate A on the kth occasion

$\lambda_k = P(X_{kj}=1) = (2P-1)\Pi_{kA} + (1-P)$

$\lambda_{11} = P(X_{2j}=1 | X_{1j}=1)$

$\lambda_{10} = P(X_{2j}=0 | X_{1j}=1)$

$\lambda_{01} = P(X_{2j}=1 | X_{1j}=0)$

$\lambda_{00} = P(X_{2j}=0 | X_{1j}=0)$

The likelihood function is written in terms of three parts of the sample as below.

$$lik(\{obs\};\lambda) = \binom{n}{n_1}\lambda_1^{n_1}(1-\lambda_1)^{n-n_1} \cdot \binom{u}{u_1}\lambda_2^{u_1}(1-\lambda_2)^{u-u_1} \cdot \left\{\left[\binom{m_1}{m_{11}}\lambda_{11}^{m_{11}}\lambda_{10}^{m_1-m_{11}}\right] \cdot \left[\binom{m_0}{m_{01}}\lambda_{01}^{m_{01}}\lambda_{00}^{m_0-m_{01}}\right]\right\} \quad (2)$$

where $\{obs\} = (n, n_1, m_1, m_{11}, m_0, m_{01}, u, u_1)$

n_1 and u_1 denote the number of respondents who reply 'yes' on the kth occasion, k=1,2

m_{11} is the number of respondents that belong to the sample of "yes" group, reply "yes" on the second occasion investigation

m_{01} is the number of respondents that belong to the sample of 'no' group, reply "yes" on the second occasion investigation

$$\lambda_{11} = \frac{1}{\lambda_1}\{\Pi_{1A}P[(2P-1)\Pi_{11} + (1-P)] + (1-\Pi_{1A})(1-P)[(2P-1)\Pi_{01} + (1-P)]\}$$

$$\lambda_{01} = \frac{1}{\lambda_1}\{\Pi_{1A}P[P-(2P-1)\Pi_{11}] + (1-\Pi_{1A})(1-P)[P-(2P-1)\Pi_{01}]\}$$

$$\lambda_{01} = \frac{1}{1-\lambda_1}\{\Pi_{1A}(1-P)[(2P-1)\Pi_{11} + (1-P)] + (1-\Pi_{1A})P[(2P-1)\Pi_{01} + (1-P)]\}$$

$$\lambda_{00} = \frac{1}{1-\lambda_1}\{\Pi_{1A}(1-P)[P-(2P-1)\Pi_{11}] + (1-\Pi_{1A})P[P-(2P-1)\Pi_{01}]\}$$

To simplify the likelihood function, the following assumption is supposed to be true without loss of generality.

P_r (voter will support candidate A on the second occasion, $X_{1j}=1$ | the voter supported candidate A on the first occasion) = P_r (voter will support candidate A on the second occasion | the voter supported candidate A on the first occasion) · P_r ($X_{1j}=1$ | the voter supported candidate A on the first occasion)

Under the above assumption, it can be easily proved that $\lambda_{10} = 1-\lambda_{11}$ and $\lambda_{00} = 1-\lambda_{01}$. Though the parameters of interest, λ_{11} and λ_{01}, are functions of Π_{1A}, Π_{11} and Π_{01}, the parameter λ_1 which is a function of Π_{1A} can be estimated on the basis of only the first occasion sample. The maximum likelihood estimators of Π_{11} and Π_{01} can be obtained as follows. First of all, we compute the estimator of λ_1 on the basis of only the first occasion sample, and an estimator of Π_{1A} is calculated by $\hat{\Pi}_{1A} = \{n_1/n-(1-P)\}/(2P-1)$.

Secondly, the estimator of λ_2 is computed on the basis of only the second occasion unmatched sample, so the estimator of Π_{2A} is obtained as $\hat{\Pi}_{2A} = \{u_1/u-(1-P)\}/(2P-1)$.

By inserting $\hat{\Pi}_{1A}$ and $\hat{\Pi}_{2A}$ into the equation (2), we can obtain a conditional likelihood function for given $\hat{\lambda}_1$ and $\hat{\lambda}$ that is based on only two successive occasions matched sample.

$$lik_q(\{m_{11}, m_1, m_0, m_{01}\}, \{\lambda_{11}, \lambda_{01}\}) = (const)\lambda_{11}^{m_{11}}(1-\lambda_{11})^{m_1-m_{11}}\lambda_{01}^{m_{01}}(1-\lambda_{01})^{m_0-m_{01}} \quad (3)$$

In the above equation (3), we can compute constrained maximum likelihood estimators of λ_{11} and λ_{01} by differentiating the equation (3) with respect to λ_{11} and λ_{01}. The computed estimators are $\widehat{\lambda_{11}} = m_{11}/m_1$ and $\widehat{\lambda_{01}} = m_{01}/m_0$. If above estimators are combined with the equation (2), then a system of equations of λ_{11} and λ_{01} can be written as follows.

$$\begin{aligned} \frac{m_{11}}{m_1} &= \frac{2P-1}{(2P-1)\ \widehat{\Pi}_{1A}+1-P}\left[P\ \widehat{\Pi}_{1A}\Pi_{11}+(1-P)(1-\ \widehat{\Pi}_{1A})\Pi_{01}\right]+1-P \\ \frac{m_{01}}{m_0} &= \frac{2P-1}{P-(2P-1)\ \widehat{\Pi}_{1A}}\left[(1-P)\ \widehat{\Pi}_{1A}\Pi_{11}+P(1-\ \widehat{\Pi}_{1A})\Pi_{01}\right]+1-P \end{aligned} \tag{4}$$

In above equations only Π_{11} and Π_{01} are unknown variables and other quantities can be considered as constants. The system of equation (4) is expressed by a simple form as follows.

$$\begin{aligned} a_{11}\Pi_{11}+a_{12}\Pi_{01} &= c_1 \\ a_{21}\Pi_{11}+a_{22}\Pi_{01} &= c_2 \end{aligned} \tag{5}$$

where
$$a_{11} = \frac{\widehat{\Pi}_{1A}P(2P-1)}{(2P-1)\ \widehat{\Pi}_{1A}+(1-P)}$$

$$a_{12} = \frac{(1-\ \widehat{\Pi}_{1A})(1-P)(2P-1)}{(2P-1)\ \widehat{\Pi}_{1A}+(1-P)}$$

$$a_{21} = \frac{\widehat{\Pi}_{1A}(1-P)(2P-1)}{P-(2P-1)\ \widehat{\Pi}_{1A}}$$

$$a_{22} = \frac{(1-\ \widehat{\Pi}_{1A})P(2P-1)}{P-(2P-1)\ \widehat{\Pi}_{1A}}$$

$$c_1 = \frac{m_{11}}{m_1}-(1-P)$$

$$c_2 = \frac{m_{01}}{m_0}-(1-P)$$

If the coefficients of system of equations (5) satisfy that $a_{11}a_{22}-a_{12}a_{21}\neq 0$, then the system of equations is consistent. The condition of consistency of the system of equations is given by the Theorem 1.

Theorem 1. Whenever ① $P\neq 0.5$ and $\widehat{\Pi}_{1A}\neq 0$, or ② $P\neq 0.5$ and $\widehat{\Pi}_{1A}\neq 1$, the system of equations (5) is consistent and its solution is given as follows.

$$\begin{pmatrix} \widehat{\Pi}_{11} \\ \widehat{\Pi}_{01} \end{pmatrix} = \begin{pmatrix} \dfrac{1}{(2P-1)^2\ \widehat{\Pi}_{1A}}\left\{\widehat{\lambda}_1 P\dfrac{m_{11}}{m_1}-(1-P)(1-\widehat{\lambda}_1)\dfrac{m_{01}}{m_0}+(1-P)(1-P-\widehat{\lambda}_1)\right\} \\ \dfrac{1}{(2P-1)^2(1-\ \widehat{\Pi}_{1A})}\left\{-\ \widehat{\lambda}_1(1-P)\dfrac{m_{11}}{m_1}+P(1-\widehat{\lambda}_1)\dfrac{m_{01}}{m_0}+(1-P)(\widehat{\lambda}_1-P)\right\} \end{pmatrix} \tag{6}$$

From equation (6), $\hat{\Pi}_{11}$ and $\hat{\Pi}_{01}$ can be expressed by the terms of $\hat{\lambda}_1$, $\hat{\lambda}_{11}$ and $\hat{\lambda}_{01}$.

$$\hat{\Pi}_{11}=\frac{1}{(2P-1)(\hat{\lambda}_1-1+P)}\{P\hat{\lambda}_1\hat{\lambda}_{11}-(1-P)(1-\hat{\lambda}_1)\hat{\lambda}_{01}\}-\frac{1-P}{2P-1}$$

$$\hat{\Pi}_{01}=\frac{1}{(2P-1)(P-\hat{\lambda}_1)}\{P_1(1-\hat{\lambda}_1)\hat{\lambda}_{01}-(1-P)\hat{\lambda}_1\hat{\lambda}_{11}\}-\frac{1-P}{2P-1}$$

For large samples, variances of $\hat{\Pi}_{11}$ and $\hat{\Pi}_{01}$ can be computed by the Taylor series approximation method.

$$\hat{\Pi}_{11}=\Pi_{11}+\frac{\partial\hat{\Pi}_{11}}{\partial\hat{\lambda}_1}\bigg|_{\hat{\lambda}=\lambda}(\hat{\lambda}_1-\lambda_1)+\frac{\partial\hat{\Pi}_{11}}{\partial\hat{\lambda}_{11}}\bigg|_{\hat{\lambda}=\lambda}(\hat{\lambda}_{11}-\lambda_{11})+\frac{\partial\hat{\Pi}_{11}}{\partial\hat{\lambda}_{01}}\bigg|_{\hat{\lambda}=\lambda}(\hat{\lambda}_{01}-\lambda_{01})+R_n$$

where $\hat{\lambda}=(\hat{\lambda}_1,\hat{\lambda}_{11},\hat{\lambda}_{01})'$ and $\lambda=(\lambda_1,\lambda_{11},\lambda_{01})'$, R_n denotes remainder term.

The property of maximum likelihood estimation implies that $\hat{\lambda}_1$, $\hat{\lambda}_{11}$ and $\hat{\lambda}_{01}$ are asymptotically unbiased for large samples. Hence an asymptotic variance of $\hat{\Pi}_{11}$ can be calculated as follows.

$$\begin{aligned}Var(\hat{\Pi}_{11})=E(\hat{\Pi}_{11}-\Pi_{11})^2=&\left\{\frac{\partial\hat{\Pi}_{11}}{\partial\hat{\lambda}_1}\bigg|_{\hat{\lambda}=\lambda}\right\}^2Var(\hat{\lambda}_1)+\left\{\frac{\partial\hat{\Pi}_{11}}{\partial\hat{\lambda}_{11}}\bigg|_{\hat{\lambda}=\lambda}\right\}^2Var(\hat{\lambda}_{11})\\&\left\{\frac{\partial\hat{\Pi}_{11}}{\partial\hat{\lambda}_{01}}\bigg|_{\hat{\lambda}=\lambda}\right\}^2Var(\hat{\lambda}_{01})+2\left\{\frac{\partial\hat{\Pi}_{11}}{\partial\hat{\lambda}_1}\bigg|_{\hat{\lambda}=\lambda}\right\}\left\{\frac{\partial\hat{\Pi}_{11}}{\partial\hat{\lambda}_{11}}\bigg|_{\hat{\lambda}=\lambda}\right\}Cov(\hat{\lambda}_1,\hat{\lambda}_{11})\\&+2\left\{\frac{\partial\hat{\Pi}_{11}}{\partial\hat{\lambda}_1}\bigg|_{\hat{\lambda}=\lambda}\right\}\left\{\frac{\partial\hat{\Pi}_{11}}{\partial\hat{\lambda}_{01}}\bigg|_{\hat{\lambda}=\lambda}\right\}Cov(\hat{\lambda}_1,\hat{\lambda}_{01})\\&+2\left\{\frac{\partial\hat{\Pi}_{11}}{\partial\hat{\lambda}_{11}}\bigg|_{\hat{\lambda}=\lambda}\right\}\left\{\frac{\partial\hat{\Pi}_{11}}{\partial\hat{\lambda}_{01}}\bigg|_{\hat{\lambda}=\lambda}\right\}Cov(\hat{\lambda}_{11},\hat{\lambda}_{01})\end{aligned}$$

The partial derivatives are given as below.

$$\frac{\partial\hat{\Pi}_{11}}{\partial\lambda_1}\bigg|_{\hat{\lambda}=\lambda}=\frac{1}{(2P-1)}\frac{P(1-P)(\lambda_{01}-\lambda_{11})}{(\lambda_1-1+P)^2}$$

$$\frac{\partial\hat{\Pi}_{11}}{\partial\lambda_{11}}\bigg|_{\hat{\lambda}=\lambda}=\frac{P\lambda_1}{\lambda_1-1+P}$$

$$\frac{\partial\hat{\Pi}_{11}}{\partial\lambda_{01}}\bigg|_{\hat{\lambda}=\lambda}=\frac{-(1-P)(1-\lambda_1)}{\lambda_1-1+P}$$

Variances and covariances are computed on the basis of sample selection process.

$$Var(\hat{\lambda}_1)=(1-f)\frac{\lambda_1(1-\lambda_1)}{n}$$

$$Var(\hat{\lambda}_{11})=(1-\mu)\frac{\lambda_{11}(1-\lambda_{11})}{m_y}$$

$$Var(\widehat{\lambda_{01}}) = (1-\mu)\frac{\lambda_{01}(1-\lambda_{01})}{m_0}$$

$$Cov(\widehat{\lambda_1}, \widehat{\lambda_{11}}) = Cov(\widehat{\lambda_1}, \widehat{\lambda_{01}}) = Cov(\widehat{\lambda_{11}}, \widehat{\lambda_{01}}) = 0$$

where f is sampling fraction and μ is the proportion of matched sample.

The asymptotic variance of $\widehat{\Pi}_{11}$ can be expressed in terms of λ_1, λ_{11} and λ_{01}.

$$Var(\widehat{\Pi}_{11}) = \frac{1}{(2P-1)^2}\left\{\left\{\frac{P(1-P)(\lambda_{01}-\lambda_{11})}{(\lambda_1-1+P)^2}\right\}^2(1-f)\frac{\lambda_1(1-\lambda_1)}{n} + \left\{\frac{P\lambda_1}{\lambda_1-1+P}\right\}^2 \cdot (1-\mu)\frac{\lambda_{11}(1-\lambda_{11})}{m_1} + \left\{\frac{(1-P)(1-\lambda_1)}{\lambda_1-1+P}\right\}^2(1-\mu)\frac{\lambda_{01}(1-\lambda_{01})}{m_0}\right\} \quad (7)$$

To simplify above variance, let's ignore the finite population corrections, then the variance of $\widehat{\Pi}_{11}$ is expressed by the following equation.

$$Var(\widehat{\Pi}_{11}) = \frac{1}{(2P-1)^2(\lambda_1-1+P)^4}\Big[\{P(1-P)(\lambda_{01}-\lambda_{11})\}^2\frac{\lambda_1(1-\lambda_1)}{n} + \frac{(\lambda_1-1+P)^2}{m_1m_0}\{m_0P^2\lambda_1^2\lambda_{11}(1-\lambda_{11}) + m_1(1-P)^2(1-\lambda_1)^2\lambda_{01}(1-\lambda_{01})\}\Big] \quad (8)$$

where $\lambda_1 = (2P-1)\Pi_{1A} + (1-P)$

$$\lambda_{11} = \frac{1}{\lambda_1}\{\Pi_{1A}P[(2P-1)\Pi_{11} + (1-P)] + (1-\Pi_{1A})(1-P)[(2P-1)\Pi_{01} + (1-P)]\}$$

$$\lambda_{01} = \frac{1}{1-\lambda_1}\{\Pi_{1A}(1-P)[(2P-1)\Pi_{11} + (1-P)] + (1-\Pi_{1A})P[(2P-1)\Pi_{01} + (1-P)]\}$$

Similarly, the variance of $\widehat{\Pi}_{01}$ can be obtained as below.

$$Var(\widehat{\Pi}_{01}) = \frac{1}{(2P-1)^2(P-\lambda_1)^4}\Big[\{P(1-P)(\lambda_{01}-\lambda_{11})\}^2\frac{\lambda_1(1-\lambda_1)}{n} + \frac{(P-\lambda_1)^2}{m_1m_0} \cdot \{m_0(1-P)^2\lambda_1^2\lambda_{11}(1-\lambda_{11}) + m_1P^2(1-\lambda_1)^2\lambda_{01}(1-\lambda_{01})\}\Big] \quad (9)$$

The estimators of $Var(\widehat{\Pi}_{11})$ and $Var(\widehat{\Pi}_{01})$ can be obtained by replacing λ_1, λ_{11} and λ_{01} with their estimators in the above equations.

4. Numerical Example

In the previous sections we have studied estimation of transition probability and its variance on two successive occasions sampling estimation of transition probabilities. The practicability of the results of studies will be approved by a numerical example. In this section, we mainly deal with the estimation of transition probabilities.

4.1 Description of Parameters

For illustrating numerical example, the values of parameters are large enough to obtain reasonable estimators from observed data. That is $\Pi_{1A}=0.55$, $\Pi_{11}=0.35$, $\Pi_{01}=0.25$.

Unless the values of Π_{11} and Π_{01} are large, estimates of Π_{11} and Π_{01} are either negative or greater than Π_{1A}. We don't deal with the feasible region of estimators, Raghavarao(1978) and Singh(1978) have given the feasible region of estimator.

The values of parameters and sizes of samples in Table 1 are calculated theoretically as follows; λ_1 is computed by $(2P-1)\Pi_{1A}+(1-P)$, λ_{11} and λ_{01} are calculated by formulas in equation(2), the number of persons who reply "yes", n_1, is obtained by $m\lambda_1$, $m_1=0.5n_1$, $m_0=0.5n_0$, variances of $\hat{\Pi}_{11}$ and $\hat{\Pi}_{01}$are computed by the equations(8) and (9).

The four values of P are considered in this numerical example. We can observe that as the value of P goes to large, the variance becomes small.

Table 1. Values of quantities for some P

P	0.75	0.8	0.85	0.9
λ_1	0.525	0.530	0.535	0.54
λ_{11}	0.4143	0.3998	0.3862	0.3733
λ_{01}	0.3895	0.3640	0.3374	0.3096
n_1	210	212	214	216
m_1	105	106	107	108
m_0	95	94	93	92
$Var(\hat{\Pi}_{11})$	0.02082	0.01094	0.00647	0.00911
$Var(\hat{\Pi}_{01})$	0.02827	0.01425	0.00802	0.00488

4.2 Estimates of Transition Probabilities and Variances

The properties of estimators of Π_{11} and Π_{01} are investigated on 14 sets of data. Estimates of transition probabilities and their variances are computed on the basis of artificial data instead of the actual data that would be collected from practical field survey. Results of computation are illustrated in Table 2.

As the number of persons who reply "yes" goes large, $\hat{\Pi}_{11}$ and $\hat{\Pi}_{01}$ become large on the all 14 sets of data. We can observe that estimators of Π_{11} and Π_{01} are changed sharply on the little change of m_{11} and m_{01}, but estimates of $Var(\hat{\Pi}_{11})$ and $Var(\hat{\Pi}_{01})$ are very stable. For example, when $P=0.75$, m_{11} varies 41 to 45 and m_{01} changes 35 to 39, estimate of Π_{11} has the range of (0.300, 0.373) and estimate of Π_{01} changes 0.211 to 0.300, but their variances change relatively small.

The noteworthy fact is that $Var(\hat{\Pi}_{11})$ and $Var(\hat{\Pi}_{01})$ given in Table 2 are almost same as $Var(\hat{\Pi}_{11})$ and $Var(\hat{\Pi}_{01})$ given in Table 1.

The sizes of matched samples m_1 and m_0 are as large as around 100, which makes it possible to apply the large sample theory to compute asymptotic estimator and its variance. That non-response or false answer gives rise to significant bias and error in estimation can be inferred from that the estimates of and change sharply in accordance with the variations of m_{11} and m_{01}.

Table 2. Estimates of Π_{11}, Π_{01}, $Var(\hat{\Pi}_{11})$ and $Var(\hat{\Pi}_{01})$

P	n_1	m_1	m_0	m_{11}	m_{01}	$\hat{\Pi}_{11}$	$\hat{\Pi}_{01}$	$Var(\hat{\Pi}_{11})$	$Var(\hat{\Pi}_{01})$
0.75	210	105	95	41	35	0.3000	0.2111	0.02042	0.02766
0.75	210	105	95	43	37	0.3364	0.2556	0.02758	0.02825
0.75	210	105	95	45	39	0.3727	0.3000	0.02103	0.02873
0.8	212	106	94	40	32	0.3131	0.2099	0.01071	0.01383
0.8	212	106	94	42	34	0.3434	0.2469	0.01091	0.01421
0.8	212	106	94	44	36	0.3737	0.2839	0.01107	0.01453
0.85	214	107	93	39	29	0.3200	0.2120	0.00632	0.00771
0.85	214	107	93	41	31	0.3460	0.2438	0.00645	0.00797
0.85	214	107	93	42	32	0.3589	0.2596	0.00651	0.00809
0.85	214	107	93	43	33	0.3720	0.2755	0.00656	0.00821
0.9	216	108	92	38	26	0.3239	0.2153	0.00406	0.00463
0.9	216	108	92	40	28	0.3466	0.2431	0.00416	0.00483
0.9	216	108	92	42	29	0.3579	0.2569	0.00425	0.00492
0.9	216	108	92	42	30	0.3693	0.2708	0.00424	0.00821

5. Conclusion and Remarks

We devise a sampling scheme which combines successive occasions sampling theory with a randomized response technique. This can reduce bias and raise the rate of response in sensitive subject surveys. We suggest a new sampling scheme for estimating the change of supporting proportions on two successive time periods. Especially, the first occasion's sample is divided into two parts on the basis of respondent's reply, and the matched sample is drawn from each group of "yes" and "no". Estimators of the proportion of the supporting candidate as well as the change of proportions are derived. The transition probabilities are valuable information on election survey are studied on the basis of two successive occasion sample.

Finally, numerical examples show that the suggested sampling strategy can be applied to the practical field survey, and the derived estimators of Π_{11} and Π_{01} are sensitive to change of data.

The study is dealt with only binomial cases. In further research this work can be extended to multi-proportion case and the case of existing respondent's false reply. Also, it will be of interest to develop the randomizing devise that can be applied to telephone survey. The distribution of birth-month of voters can be considered as a randomized device for an election telephone survey.

References

[1] Abul-Ela,A.A., Greenberg,B.G., and Horvitz,D.G.(1967) "A multi-proportions randomized response model," *Journal of the American Statistical Association* 62, 990-1008.

[2] Chaudhuri,A., and Mukerjee,R.(1988) Randomized Response theory and techniques, Mercell Dekker, Inc. New York.

[3] Cochran,W.G.(1977) Sampling Techniques, John Wiley and Sons, New York.

[4] Kendall,M.G.,and Stuart,A.(1973) The Advanced Theory of Statistics, Charles Greffin Company, London.

[5] Lee,Kay-O(1985) Comparison of Successive occasions sampling schemes from spatially correlated processes, Ph. D. Thesis, North Carolina State University.

[6] Raghavarao, D.(1978) "On an Estimation Problem in Warner's RR Technique," *Biometrics* 34, 87-90

[7] Singh, J. (1978) "A note on Maximum Likelihood Estimation from Randomized Response Models," American Statistical Association 1978 Proceedings of Social Statistics Section, 282-283