

**EFFICIENCY OF RESPONDENT EFFECT
IN A SUPERVISED INTERVIEW-REINTERVIEW SURVEY**

Roshanak Alimohammadi

Department of Statistics

Faculty of Mathematics

Alzahra University

Vanak, Tehran, IRAN

Abstract: Survey implementation is a common method to data collection. There are several methods for data collecting, such as mailing questionnaire, face to face surveys, telephone and mobile interviewing.

In this article, interview-reinterview surveys are considered, in order to obtain more precise estimator of respondent effect. Therefore, a model of response error for interview-reinterview surveys is proposed. Then, estimators of the model components are obtained by two methods. Furthermore, the efficiency of the model effects is computed by considering respondent error in an interview-reinterview setting.

AMS Subject Classification: 62P99

Key Words: efficiency, interview-reinterview setting, precision, respondent effect

1. Introduction

Precision is a main concept related to data quality. There is an inverse relation between precision and total variance. Therefore, to determine the precision

of data, total variance should be estimated. In the other words nonsampling variance should be estimated, as well as sampling variance. Response error is a main part of nonsampling error. It is defined as the difference between true value and observed value. Modeling of response error has many applications. One of the applications is to estimate survey precision.

In the previous works, response error modeling of face to face surveys is studied (Alimohammadi [1], [2]). In this article, it is assumed that other sources of response error except of the response error model effects may be ignored. In this paper, response error model is considered in the general case and a term of model error is considered to estimate the effect of response error. In this order, a model of response error in interview-reinterview setting is proposed.

2. Survey Errors

By definition, the quality of survey data has some components including accuracy, relevance, timeliness, accessibility and interpretability.

Accuracy is an important and visible aspect of data quality that has been of concern to statisticians and methodologists for many years. It refers to closeness between estimated and true (but unknown) values.

A measure of accuracy should be defined for measuring error of surveys. Variance is the most common measure. Usually the variances are unknown and should be estimated from the collected data. Estimation of sampling variances are computed based on sampling design and form of the estimator. A large part of the classical sampling theory is about simple surveys, but usually other cases can occur in practice. Cochran [5] discussed the basic sampling designs and estimators. Wolter [10] presented diverse methods of variance estimation in complex surveys.

For estimating survey errors, besides sampling error, nonsampling errors should be estimated. Groves [6] and Groves et al [7] have studied nonsampling error in surveys. Nonsampling errors contain coverage error, nonresponse error, response error and data processing error.

2.1. Sampling Error

Sampling error refers to the variability that occurs by chance because a sample is surveyed rather than all units of the population. The value of this error is zero in censuses. The sampling error for produced estimates from simple random samples can be computed easily. In other words, when the aim is

variance estimation of a linear estimator in a simple sampling design, it can be carried out simply, but for some of nonlinear estimator or complex designs, estimation of variance may be more difficult. In such cases, the surveys are called complex surveys. In general, estimation of variance depends on sampling design and form of the estimator. Methods of variance estimating in complex surveys include random groups, balanced half samples, jackknife, generalized variance function and Taylor series method.

2.2. Nonsampling Error

Variance of estimator usually is used for total survey error estimating. Total variance in surveys is the sum of sampling variance and nonsampling variance.

A main part of total survey error is nonsampling error. This error contains various components such as coverage error, response error, nonresponse error and data processing error.

3. A Model of Response Error in an Interview - Reinterview Setting

Survey implementation is a main method of data collection.

Response error is an important and unavoidable part of nonsampling error in surveys.

Mahalanobis [9] was the first statistician who used ANOVA for response error modeling in face to face surveys. Then, several researchers such as Kish [8], Biemer and Stokes [4] have extended their works. Bassie and Fabbris [3] have presented a response error model including supervisor, interviewer and respondent effects in a supervised interview-reinterview setting.

Alimohammadi and Navvabpour [2] proposed a mixed nested ANOVA model for response error by assessing implementation of face to face surveys.

In this section, a model of response error in interview-reinterview setting is proposed. Sometimes, implementation of telephone surveys is more possible in practice. In the model, sources of response error is considered as supervisor error, interviewer error and respondent error as follows:

$$y_{ijk} = \mu_{ijk} + A_i + B_{ij} + C_{ijk} + E_{ijkl}, \quad (3.1)$$

where A_i is supervisor effect, B_{ij} is interviewer error, C_{ijk} is respondent error of ijk th respondent, y_{ijk} is observed value and μ_{ijk} is real value of ijk th unit. E_{ijkl} is residual error (effect of other sources of error).

In model (3.1), effect of respondents nested in interviewers and interviewer effect nested in supervisor effect. Usually in surveys, the effect A is fixed and B and C effects are random variables. Then model (3.1) is considered as a nested mixed effect analysis of variance model.

4. Estimation of the Effects of the Response Error Model

In this section, estimation of model (3.1) components is assessed. Suppose that some assumptions on model (3.1), such as data are balanced in all of the levels. B_{ij} distributed as normal with mean zero and variance σ_B^2 , C_{ijk} distributed as normal with mean zero and variance σ_C^2 and for the real values, $\mu_{ijk} = \mu + R_{ijk}$, where R_{ijk} distributed as normal with mean zero and variance σ_R^2 .

4.1. Estimation of Response Error Model Effects

In this section, estimation of respondent error is computed in interview-reinterview setting.

It can be shown that the mean of squares of model (3.1) effects are as follows:

$$MSA = \frac{SSA}{a-1}, MSB = \frac{SSB}{a(b-1)} \text{ and } MSC = \frac{SSC}{ab(c-1)}.$$

Then ANOVA estimators of variance components of the model effects are obtained as follows:

$$\begin{aligned}\hat{\sigma}_B^2 &= \frac{MSB - MSC}{cn}, \\ \hat{\sigma}_C^2 &= \frac{MSC - MSE}{n}, \\ \hat{\sigma}_E^2 &= MSE.\end{aligned}$$

In some situations, estimates of the variance components by ANOVA method may be negative. Therefore maximum likelihood estimation method is applied to determine the response error model effects. In this order, the related likelihood function based on the model is computed and partial derivative roots of the likelihood function are calculated. Then conditions of second derivative test is applied to determine coordinates of maximum point and based on the assessments maximum likelihood estimators of variance components are presented.

It can be shown that, maximum likelihood estimators of the response error model are obtained as follows:

For $(1 - 1/b)MSB \geq MSC$ and $MSC \geq MSE$, ML estimator of the model effects are computed as:

$$\hat{\sigma}_B^2 = \frac{(1 - 1/b)MSB - MSC}{cn}, \quad \hat{\sigma}_C^2 = \frac{MSC - MSE}{n} \quad \text{and} \quad \hat{\sigma}_E^2 = MSE.$$

If $(1 - 1/b)MSB \geq MSC$ and $MSC < MSE$, ML estimator of the model effects are computed as:

$$\hat{\sigma}_B^2 = \frac{(1 - 1/b)MSB - MSC - MSE}{cn}, \quad \hat{\sigma}_C^2 = 0 \quad \text{and} \quad \hat{\sigma}_E^2 = \frac{MSC + MSE}{b(cn - 1)}.$$

If $(1 - 1/b)MSB < MSC$ and $MSC \geq MSE$, ML estimator of the model effects are computed as:

$$\hat{\sigma}_B^2 = 0, \quad \hat{\sigma}_C^2 = \frac{1}{n} \frac{SSA + SSB}{bc} - MSE \quad \text{and} \quad \hat{\sigma}_E^2 = MSE.$$

If $(1 - 1/b)MSB < MSC$ and $MSC < MSE$, ML estimator of the model effects are computed as:

$$\hat{\sigma}_B^2 = 0, \quad \hat{\sigma}_C^2 = 0 \quad \text{and} \quad \hat{\sigma}_E^2 = \frac{SST}{bcn}.$$

An advantage of maximum likelihood estimators is to set them in the parameters space. Then these estimators of variance components are nonnegative.

4.2. Application of the Obtained Estimators

To compute the survey precision, the variance of the considered survey statistic should be estimated. Alimohammadi [1] proposed a relation to compute survey precision based on another response error model under some circumstances. Respondent variance is one of the components to compute survey precision. Then precise estimation of the components resulted in more precise estimators of survey precision.

5. Conclusion

The estimation of data quality in surveys is important because of various usages of the survey results.

The response error is an unavoidable error in survey. The telephone interview is a proper mode of data collection.

In this article, a model for interview-reinterview surveys is considered. To compute the survey precision, variance components of response error model should be determined. Therefore, estimators of the model effects are computed. The results of this article may be applied to compute the effect of response error in total survey error.

Acknowledgements

This research is supported financially by Alzahra University.

References

- [1] R. Alimohammadi, Comparison of variance components estimation methods of response error model in surveys, *Applied Mathematical Sciences*, **5**, No 45-48 (2011), 2405-2410.
- [2] R. Alimohammadi and H. Navvabpour, Response error modeling in face to face surveys, *Research Journal of Science* (Isfahan University), **33** (2008), 1-14.
- [3] F. Bassie and L. Fabbris, Estimation of nonsampling error in interview reinterview supervised in interpenetrated assignments, In: *Survey Measurement and Process Quality*, Wiley, New York (1997).
- [4] P.P. Biemer, and S.L. Stokes, Approaches of modeling of measurement error. In: P.P. Biemer and al. (Eds), *Measurement Error in Surveys*, Wiley, New York (1991).
- [5] W.G. Cochran, *Sampling Technics*, 3th Ed., Wiley, New York (1977).
- [6] R.M. Groves, *Survey Errors and Survey Costs*, Wiley, New York (2004).
- [7] R.M. Groves, F.J. Fowler, M.P. Couper, J.M. Lepkowski, E. Singer and R. Tourangeau, *Survey Methodology*, Wiley, New York (2004).
- [8] L. Kish, Studies of interviewer variance for attitudinal variables, *Journal of the American Statistical Association*, **57** (1962), 92-115.
- [9] P.C. Mahalanobis, Recent experiments in statistical sampling in the Indian statistical institute, *Royal Statistical Society*, **109** (1946), 325-378.

- [10] K. Wolter, *Introduction to Variance Estimation*, Springer-Verlag, New York (2007).

