A Study on Sample Assessment Method in Sampling Survey

Zuyuan Xiong [1], Yi Zhou [2*], Li Zhu [3]

[1] Sichuan Provincial Bureau of Statistics, Sichuan, CHINA

[2]Suining Bureau of Statistics, Southwest University of Finance and Economics, PhD, Sichuan, CHINA

[3]Sichuan Provincial Institute of Statistical Science, Sichuan, CHINA

[*]Corresponding author: Yi Zhou, zhouyi@sc.stats.cn

Abstract

The sampling survey is an essential survey method to estimate certain quantitative characteristics through the samples; however, the estimated value shall not be coincided with real value due to the inevitable sampling error. Therefore, the study on sample assessment method must aim at error of survey in combination with other factors. In order to measure the sample validity, this study drew on data quality assessment approaches and principles, took sample survey approach as a starting point, to design a comprehensive assessment proposal for samples. In addition, the suggestions are also made in this article to optimize samples on three aspects as survey preparation, sample control and error remedy.

Keyword: assessment approach, assessment proposal, sampling method, sample optimization.

1. Definition of related conceptions

1.1 Conception of sample assessment

The sample of sampling survey is the individuals extracted from population under study. Studying on sample assessment shall not only be in favor of error control, but also can direct us to perform reasonable conclusion and estimation, to improve the quality of survey data. The sample assessment is usually composed of before event, during event and after event assessment. The before event assessment is performed prior to details execution of survey, i.e., the assessment to work of survey preparation stage; during event assessment is to assess the execution process of sampling survey; after event assessment is mainly to assess the sampling error and non sampling error, and this is also the core content of this article.

1.2 Brief introduction on sampling method

The sampling method mainly includes probability sampling and non probability sampling (also called sampling with unequal probability or nonrandom sampling). This article is mainly to study the probability sampling, and the sampling error thereof can be calculated and controlled. The table below is the advantage/disadvantage of the general sampling methods.

Table 1: Advantages/disadvantages of various sampling methods

Sampling method	Advantage	Disadvantage
Non probability sampling	Simple and easy operation, low cost, time saving	The subjectivity is strong, the sample representativeness is fairly weak
Simple random sampling	The sampling frame does not need auxiliary information, the development of sampling technology is fairly mature, with existing theory and formula	It is only applicable to limited quantity of population unit, for complicated population the representativeness of the sample cannot be guaranteed; the known information of the population cannot be utilized.
Stratified sampling	The sampling efficiency is high, the sampling method in the stratify may be different, and this is in favor of organization of sampling operation.	There are many individuals in the population, the structure is fairly complicated, and the difference of interior is fairly great
System sampling	It is easy to determine the sample unit; the distribution of sample unit in the population is fairly even in favor of improving assessment precision.	It is not easy of execution for the unit with periodical change, and the variance assessment is fairly complicated.
Cluster sampling	Compilation of sampling frame is fairly simple, survey is convenient, with cost saving	The sample units are fairly concentrated, and the sampling error is fairly great
Multi-phase sampling	The samples are fairly concentrated in favor of survey and saving cost, it does not need to compile the sampling frame for all small sample units.	The sampling is fairly troublesome, and the estimate of population via sample is fairly complicated.

2. Design of sample assessment proposal

During assessment to sample, it is required to comprehensively consider various factors to determine whether the sample can be used, to qualitatively analyze various errors which may exist in the sample, and quantitatively judge the sample by applying the details scoring of the assessment staff to survey and assessment question.

2.1 Determine the comprehensive assessment factor

According to viewpoint of assessment staff, the before event assessment is generalized into 5 factors, during event assessment 4 factors, after event assessment 11 factors, in total 20 assessment factors.

Table 2: Factors for sample assessment in sampling survey

Assessment item	Weight	Assessment sub-item	Weight	Full score	Subitem score	Item score
Before event	W1	1.Reasonableness of the survey time arrangement	ω1	20		S1
		2.Reasonableness of survey outlay	ω2	20		
		3.Examination strength of questionnaire design	ω3	20		
		4.Professional level of survey staff	ω4	20		
		5.The training strength and discussion before survey	ω5	20		
During event	W2	6.The mean survey time for each questionnaire or survey item	ω6	20		S2
		7.The understanding of survey staff to survey questionnaire or survey item	ω7	20		
		8.The applicable extent of sampling survey method when the cost is fixed	ω8	20		
		9.Survey process supervision and strength	ω9	20		
After event	W3	10.Whether sampling frame and auxiliary information are accurate	ω10	20		S3
		11.No reply question quantity	ω11	20		
		12.The mean data type in time of each questionnaire	ω12	20		
		13.Comparison on survey unit quantity and population unit quantity	ω13	20		
		14.Comparison on survey unit quantity and unit quantity of samples acquired	ω14	20		
		15.Whether the sample distribution is approaching to population distribution	ω15	20		
		16.Comparison on sample variance and sample variance of simple random sampling	ω16	20		
		17.Comparison on sample variance and estimated sample variance	ω17	20		
		18.Comparison on sampling error in this survey and sampling error for previous survey	ω18	20		
		19.The proportion of adjusted data quantity of sample in total quantity	ω19	20		
		20.Whether this survey is improved in contrast with previous survey	ω20	20		

Total	100%	Total score(TS)= W1*S1+W2*S2+W3*S3

2.2 Weighting processing

The sample assessment= (before event assessment, during event assessment, after event assessment). In each sampling survey, the Delphi method or AHP analytic hierarchy process shall be adopted subject to goal of survey and sampling method used, and the comment of each expert shall be generalized to perform subjective weighting to before event, during event and after event assessment, and the weighting matrix derived is :W=(W1, W2, W3)； ω=(ω1,ω2,…,ω20).

Wherein, W1=ω1+ω2+…+ω5； W2=ω6+ω7+ω8+ω9； W3=ω10+ω11+…+ω20.

If it is deemed that the assessment items (in total 20 indicators) are all of the same importance, i.e., weight of each sub-item is equivalent, the weighting matrix of assessment item is W=(25%, 20%, 55%), and weighting matrix of each sub-item is ω=(ω1,ω2,…,ω20)=(5%,5%,…,5%).

If it is deemed that the before event, during event and after event assessment are of the same importance, the weighting of each subitem factor shall then be determined. The weighting matrix of assessment item W= (33.3%, 33.3%, 33.3%), the weighting of each sub-item is determined according to key point of assessment.

It is to be mentioned that non probability sampling has no sampling error, and its after event assessment is unconcerned with item 15, 16, 17, 18, therefore the scoring of these four items are determined as full score during assessment to non-probability sampling.

2.3 Establish comprehensive assessment model

At first, the assessment staff shall score the factor subject to size or serious degree of the issue reflected, the standard of scoring refers to five types of satisfaction (very satisfied, fairly satisfied, generally satisfied, not satisfied, extremely not satisfied) assessment guideline. E.g., the assessment staff may divide the satisfaction of survey time arrangement (of before event assessment) into: very reasonable, fairly reasonable, generally reasonable, not reasonable, extremely not reasonable. The score value is 16~20, 12~16, 8~12, 4~8, 0~4. The other items score can also be divided into 16~20, 12~16, 8~12, 4~8, 0~4 based on the satisfaction degree of assessment staff.

Secondly, according to weighting and factor scoring standard of before event, during event and after event assessment, the assessment result of the sample can be divided into five types, as shown in table below.

Table 3: Type and assessment standard of assessment sample

Sample Type	Excellent	Good	General	Bad	Very bad
Assessment standard	[16~20)	[12~16)	[8~12)	[4~8)	[0~4)

Finally, the assessment staff shall calculate the score of main item subject to factors provided by the assessment, and derive the comprehensive score of the sample subject to weighting of before event, during event and after event, and then it shall be

compared with assessment standard to drive the type of the assessment sample.

3. Related suggestions for optimizing sample

3.1 The work before survey shall be prepared well

Reasonably arrange survey time and survey cost. on one hand, the details plan and budget report shall be made subject to survey program, on the other hand, it is required to refer to previous survey experience.

Determine appropriate survey method. The difference of survey time shall directly impact the sampling error and non sampling error size, and the good survey method can also provide work convenience to survey staff. Prior to item survey, it shall be carefully selected and repeatedly considered, and the feasibility and availability of the survey method shall be discussed in details.

The survey staff with fairly high professional level or with experience shall be selected. There are three advantages for doing so: first, they have fairly accurate understanding to survey questionnaire and item; second, it can improve the efficiency and precision of survey; third, it can effectively supervise the survey process. Therefore, during survey the fairly professional survey staff or survey unit shall be selected to complete the detailed survey work.

Improve test strength of questionnaire and trainings strength of survey. The examination of questionnaire and the survey training discussion can effectively improve the work efficiency of survey staff and the cognition extent to survey item, furthermore, it is also easy of deriving fairly accurate data during survey process.

3.2 Control and remedy measures of sample

Control and remedy of error of sampling frame. For the sample with existence of sampling frame error, the requirement for re-setting sampling frame to extract the sample may not be feasible during practical work. Generally the fairly effective method is to extract the sample from the incomplete sampling frame, partial samples shall be added or deleted on the basis of previous samples, to drop the error of sampling frame.

Control and remedy to no replay error. First, it is improved by substitution method, but this remedy measure shall incur other error after the no-reply is changed into with-reply; second, the survey value is derived by visiting the no-replay unit by many times, this remedy measure is the most general and most successful method; third, extract random subsample from no-reply unit, but this improvement is only applicable to specific survey method as mailing survey, etc; fourth, delete the no-reply sample, this method can effectively reduce the no-reply error, but the dropped sample quantity shall incur the drop of sample precision, therefore the sample size shall also be considered during selection and utilization thereof.

Control and remedy of measurement error: First, the fundamental survey work shall be done well, to derive the real data of survey unit to the greatest extent; second, the data type-in and review process shall be controlled, to avoid the type-in error of data.

Reference

[1] William G. Cochran. Sampling Techniques[M] John Wiley & Sons; Third edition, 1977. 08.

[2] Shiyong Feng, Guohua Zou. Sample Retotation Method Based on Auxiliary Information[J]. Statistical Research, 1996,V13(3): 62-68

[3] Zhenyong Jiao. The Manifestation of Population Variance in Sampling[J]. Statistics & Information Tribune, 2003, 18(2).

[4] Zonghui Lu, Renen Chen. On Sample Rotation in Social Economics Survey[J]. Journal of RTV University (Philosophy & Social Sciences, 2005, (1).

[5] Yan Wang. Sample Evaluation Method in Sample Survey[J]. China Statistics, 2007, (12).