Revisiting a public opinion survey

SUGIMAN Toshio

Graduate School of Human and Environmental Studies, Kyoto University, Sakyo-ku, Kyoto, 606-8501 Japan (sugiman.toshio.7a@kyoto-u.ac.jp)

Abstract: Public opinion surveys present several problems to researchers which merit discussion. Specifically, these problems are (1) difficulties of securing representation of collected samples for statistical population, (2) importance of response patterns among plural questions, (3) nature of intermediate alternatives and DK (don't know) responses, (4) the arbitrary quantification of responses and (5) two ways of investigating social norms.

Keywords: public opinion survey; representativeness of samples; response pattern; intermediate response; social norms

1 Introduction

We often see results of public opinion surveys that were administered by various research institutes, public or private, and media companies. They are useful to discover what opinions people have and what actions people take in various areas of life and society. The results of Japanese samples are sometimes compared with the ones of foreign countries to distinguish and illuminate the characteristics of Japanese people.

Researchers must, however, avoid simple-minded reliance on the results. We should pay attention to the way that data were collected and analyzed. Also, it is necessary to understand that respondents have general tendencies to answer questions beyond the differences of targeted areas.

2 Representation of data

We assume that survey data consists of responses to a questionnaire by samples that were selected randomly from a large number of people, *i.e.*, statistical population. Several sampling methods have been developed to secure representation of those statistical populations although perfect random sampling is usually difficult to be attained. For example, in a nationwide survey on Japanese national character that has been carried out each five years since 1953 by the Japanese Institute for Statistical Mathematics, several thousand samples are chosen by stratified multistage sampling method [1]. In the method, (1) all municipalities in Japan are stratified into six layers that are different in population; (2) a total of 400 towns (sampling points) are selected from each layer with probability of selection being proportionate to population of each layer; (3) samples assigned to each sampling point (16 samples on average) are selected from the basic register of residents where a single sample is selected by skipping the fixed number of persons on the register. The person who is selected by the above method is interviewed. An interviewer reads each question in front of a respondent and fills in a questionnaire sheet while showing a large card indicating both the question and the set of alternatives.

Such a sampling method as described above has become difficult over these thirty years because of several reasons. First, the increased concern with protection of personal information these days has made local governments more reluctant to permit perusal of the basic register of residents and the voting register. Both were the most reliable sources to grasp an entire statistical population. Local governments tend to restrict perusal to academic institutes.

Second, it has become difficult for an interviewer to meet a predetermined respondent and spend a long time, at least one hour, to complete filling in a questionnaire partly because of the diverse life style of respondents. Another method has become popular in which the interviewer visits a respondent's house and asks him/her to fill in a questionnaire sheet but does not remain with the respondent. Several days later, the interviewer visits the respondent's house again to retrieve the questionnaire sheet and gives an honorarium to the respondent. But, even this method has become costly because often the interviewer has to

Received date: May 23, 2012

visit a respondent's house several times to make the initial acquaintance. In addition, many people hesitate to spend a long time with a person who suddenly appears and wants to ask questions.

Telephone surveys have become prevalent due to the situation described above over the past twenty to thirty years. In telephone surveys, dialing is made to a phone number generated randomly by a computer. But, the phone number is restricted to a desk telephones in Japan, which means people, especially young people, who only use a mobile phone are excluded from the sampling procedure from the beginning. A theoretical problem of an automatic dialing method is that statistical population cannot be determined, which makes it difficult to depend on traditional statistical theories in which sampling from population is always assumed.

In addition, telephone surveys change the content of questions. In telephone surveys, verbal interactions without face-to-face communication do not allow as long an interaction as a standard in-person interview. Researchers should avoid telephone surveys combined with computer dialing if possible but there are no other options in many cases. Questions, regardless of the method used, should be as short and simple as possible.

It is easy to imagine what bias is brought about by prevalent telephone surveys. Respondents are most likely to be people who stay at home and have free time to talk with a computer-operated interviewer.

3 Response pattern

Let's turn from sampling into data analysis. The most important result of data analysis is response distribution among a set of alternatives within a single question. For example, it is most important to see what percent of respondents select each alternative of a single question.

However, we tend to combine the results of response distributions of two different questions with the use of 'and.' For example, suppose that we have two questions, Q1 and Q2, where, in both, two alternatives, yes or no, are given. Data of 200 respondents, 100 male and 100 female, is collected and analyzed. As a result, in both males and females, 50% answered 'yes' and 50% answered 'no' for both Q1 and Q2. Having read this report, you might conclude that there were no differences between male and female as far as these two questions were concerned. But, is this conclusion right?

It is possible to obtain different cross-tabulation tables for males and females as shown in Table 1. Large differences are found between males and females in Table 1, which shows that males who answered 'yes' to Q1 were also likely to answer 'yes' to Q2 and males who answered 'no' to Q1 were also likely to answer 'no' to Q2. In contrast, females who answered 'yes' to Q1 were likely to answer 'no' to Q2 and females who answered 'no' to Q1 were likely to answer 'yes' to Q2. Marginal distributions are the same for the two questions in both males and females, but it is different how answers to the two questions are related with each other.

Table 1 Two cross-tabulation tables resulting in the same marginal distribution

Male

		Q2		
		Yes	No	Total
Q1	Yes	40	10	50
	No	10	40	50
	Total	50	50	100

Female

		Q2		
		Yes	No	Total
Q1	Yes	10	40	50
	No	40	10	50
	Total	50	50	100

A concept of 'response pattern' proposed by Hayashi is useful to understand how responses to two or more questions are related. Table 1 shows there are two major groups that differ in response patterns in males: one group responds to both Q1 and Q2 by 'yes' and another group responds to the both by 'no.' In contrast, there are also two major groups that differ in response pattern in females but females' response patterns are different from males': one group responds to Q1 by 'yes' but to Q2 by 'no' and another group responds to Q1 by 'no' but to Q2 by 'yes.'

It is easy to discover major response patterns when you have just two questions as we do in the above. But, how can you discover these patterns when you have ten, twenty, or more questions? For this, the quantification method for classification of response patterns, developed by Hayashi and often called Hayashi's quantification III, is useful [2].

4 Intermediate responses

When you make a cross-cultural comparison based on a questionnaire type survey, you have to take into account Japanese characteristics of response to a question in which a set of alternatives differing in degree is asked. For example, a set of alternatives such as 'strongly agree,' 'agree,' 'neutral,' 'disagree,' and 'strongly disagree' is given to answer a question for respondents to show how much they agree or disagree to a certain political judgment.

Japanese people are likely to hesitate showing an extreme response, 'strongly agree' (or 'strongly disagree') and prefer a moderate response, 'agree' (or 'disagree') even when they strongly agree (or strongly disagree) [3]. This makes an average score of Japanese respondents lower than foreign respondents when data is processed by quantifying each alternative, like scoring 'strongly agree' as 5, 'agree' as 4 and so on.

It might be interesting to explore how the cultural background is related to the Japanese preference for a moderate response, but language was found to be one of major reasons. For example, it was found that Japanese university students tended to answer 'strongly agree' or 'strongly disagree' more often when given an English version of a questionnaire than the Japanese version of the same questionnaire. The tendency was confirmed by a different study in which respondents were asked to assign a total of five tokens to two different alternatives of a single question according to how much they agree. The results showed that an assignment of '2-3' or '3-2' was made more often when they were asked in Japanese than in English [3].

A question often includes an alternative, 'I don't know,' which is sometimes referred to as 'DK' (abbreviation of 'Don't Know') response. Two different interpretations are possible for the DK response, especially when special knowledge is required to answer a question. A typical example might be a question concerning acceptance of technology such as biotechnology and nuclear power generation. It was found, by Hibino with the use of Hayashi's quantification method mentioned above, that the DK response is interpreted as showing an intermediate attitude which is between two different clear attitudes, *e.g.*, agree or disagree, on some occasions while it is interpreted as lack of knowledge to answer a question on the other occasions [4].

5 Arbitrary quantification

Numerical analysis of questionnaire survey data is often reported in academic journals. In the analysis, responses to verbal alternatives are quantified with numbers *i.e.* 'strongly agree' = 5, 'agree' = 4, 'neutral' = 3, 'disagree' = 2 and 'strongly disagree' = 1. After such quantification, various statistics such as mean values, standard deviations, and correlation coefficients are computed. Furthermore, data are often submitted to such statistical analyses as analysis of variance or multivariate analyses such as regression analysis and factor analysis.

However, it should be noted that such quantification is made arbitrarily by those who analyze data. Respondents never gave a numerical value to a question but they chose an alternative described verbally. A mean value of quantified response is sometimes useful to see an overall tendency if frequency distribution among alternatives is similar to normal distribution. A Pearson's correlation coefficient is sometimes useful to see overall relationships between two different questions if the relation is restricted to a linear one. However, we should be cautious about going beyond these and using seemingly more sophisticated methods, like multivariate analyses, depending on arbitrary quantification. This is because one can obtain totally different results if one depends on different quantifications. Equal intervals between two adjacent alternatives are not guaranteed theoretically.

Rather than depending on arbitrary quantification, we should use analysis methods in which selection of verbal alternatives are respected. For this, frequency distribution among alternatives of a question should

be reported rather than a mean value that is computed after arbitrary quantification. Also, a cross-tabulation table of two questions is more useful than a correlation coefficient to report data. For multivariate analysis, Hayashi's quantification method, mentioned above, is preferred to factor analysis because the method not only respects alternatives of a question but can explore linear and non-linear relations among plural questions.

When one is interested in finding major determinants of a criterion variable, comparison among possible cross-tabulation models with the use of Akaike's Information Criterion (AIC) is preferred to regression analysis. This is because arbitrary quantification is required for regression analysis and non-linear relationships are ignored [5].

6 Social norms

One of the major purposes of a public opinion survey is to investigate social norms. A social norm is defined as a set of shared assumable recognitions or actions, which are significant and meaningful to that group.

There are two different kinds of social norms, *i.e.*, valuational and cognitive. The difference in the two refers to what happens when non-assumable recognition or action takes place. In a valuational norm, a person who takes non-assumable recognition or action is pressured to conform to the norm while the norm itself is never changed. A norm that determines traffic rules on the road is an example of a valuational norm. In contrast, in cognitive norms, a non-assumable recognition or action does not pressure a person to change while the norm is revised or expanded to include the non-assumable recognition or action as assumable. Academic knowledge is a typical example of a cognitive norm [6].

Two different methods are possible to understand social norms, regardless of their valuational or cognitive character, by public opinion survey. One is a widely-used method in which respondents are asked to express their own opinions in answers to questions beginning with 'How do you think ----?' Obviously, responses to such questions suggest what social norm is prevalent and how much pressure it exerts on the individual

But, we can use another method in which respondents are asked to show their judgment about most people's opinion, like a question beginning with "Do you think most people around you agree ----?" Here, suppose that there is a group of ten persons working together and each of them is convinced that he/she works hard but the other people do not. In this situation, you can obtain quite different results when you ask them "Do you work hard?" rather than "Do you think most people around you work hard?".

The two methods are posited to measure two different aspects of social norms. Some social phenomena are determined by one of the two, but others are determined by both with different degrees of influence. Purchase or sale in the stock market is an example in which the both determine people's behavior. You buy an issue whose price you believe will rise soon, but at the same time you buy an issue that you predict most people will want to buy soon.

7 Impossibility of neutral measurement

Lastly, we have to remember that it is impossible to achieve a neutral relationship between a survey administrator and respondents. In other words, it is impossible for a public opinion survey to secure an objective measurement in which results of the measurement are the same regardless of who investigates the respondents. For example, if a survey involved questions about taxes, on a macroscopic level, there would be a great difference between an academic institution and a government agency conducting that survey.

A public survey is a kind of conversation between a survey administrator and respondents rather than an instrument of objective measurement. It is natural for a person to change an answer to a question depending on who asks the question. In fact, more microscopically, respondents might change their answers depending on the sex or age of a personal interviewer or what that interviewer looks like.

The context of an entire conversation has a great deal of influence on each remark made in the course of the conversation. In addition, the order of questions produces a certain context, which might affect

answers of each question. But, it is natural rather than problematic when we take public opinion survey as conversation between a survey administrator and respondents macroscopically and between an interviewer and respondents microscopically. We, as social scientists, should be cautious enough not to mistake data of public opinion surveys for objective data which is the type of data required in the natural sciences.

References

[1] INSTITUTE OF STATISTICAL MATHEMATICS: Study on the Japanese National Character. http://www.ism.ac.jp/kokuminsei/ks_e/index_e.html (May 10, 2012)

[2] HAYASHI, C.: Theory and Examples of Quantification (II), Proceedings of the Institute of Statistics and Mathematics, 1956, 4:19-30.

[3] HAYASHI, C., and HAYASHI, F.: Comparative Study of National Character, Proceedings of the Institute of Statistical Mathematics, 1995, 43 (1), 27-80.

[4] HIBINO, A.: Interpreting “Don’t Know” Answers in the Public Attitude to Science and Technology: Questionnaire Survey on Public Perception of Biotechnology in Japan, Japanese Sociological Review, 2010, 60 (4), 554-569.

[5] SAKAMOTO, Y.: Model Analysis of Categorical Data, Tokyo: Kyoritsu-Shuppan, 1885.

[6] SUGIMAN, T.: Part II-7 “A Theory of Construction of Norm and Meaning: Osawa’s Theory of Body”, In Sugiman T, Gergen K.J., Wagner, W., and Yamada Y. (Eds.): Meaning in Action: Constructions, Narratives and Representations, Tokyo Springer, 2008.