Latent Classification of Survey Respondents, Based on Respondents' and Interviewers' Evaluations

Meta Gnidovec and Samo Kropivnik[1]

Abstract

In general, respondents have different beliefs about the level of applicability and accuracy of social science surveys. On the other hand, interviewers have different feelings about the level of participants' cooperation and understanding. In the Slovene "Victim Survey" conducted in 1992, the participants' evaluations of the level of survey's applicability and accuracy, and the interviewers' evaluations of the level of respondent's cooperation and understanding were measured by means of four categorical variables. A significant level of association among four variables was found. Therefore, these variables were considered to be part of the same latent "complex" and they were used to develop a typology of survey respondents. An exploratory latent class model was used to identify three latent classes. On the basis of the conditional response probabilities, the obtained latent classes were recognized as the "Ideal Respondents Class", the "Skeptics Class" and the "Delicate Respondents Class". Furthermore, each respondent was assigned to one of the three classes on the basis of the modal classification probabilities. In the final step, the characteristics of the obtained classes (groups) of respondents were analysed with respect to several "external" variables, e.g., gender, age, level of education, interview method, number of contacts.

Keywords: Association between categorical variables; Exploratory latent class model; Latent typology; Social science survey respondents.

1 Introduction

In general, respondents have different beliefs about social science surveys. Some of them believe that surveys serve a good purpose, and others do not. Some of them believe that survey results can be trusted and others do not.

---

[1]Faculty of Social Sciences, University of Ljubljana, P.O. Box 47, Ljubljana, Slovenia

On the other hand interviewers have different feelings about the respondents' level of cooperation and understanding. Interviewers say that some respondents are more willing to cooperate in surveys than others. Interviewers also say that it is easy for some people to understand the questions, and not so easy for others. Some respondents are easy to deal with and others are not.

There are different believes among respondents about surveys as well as different feelings among interviewers about respondents. The question to be investigated here is whether there is any connection among the respondents' evaluations of surveys or among the interviewers' feelings about respondents and among both? And: if there is a relationship, how can it be described?

2   Respondents' and interviewers' evaluations

In Slovene "Victim Survey" conducted in September 1992 (Ferligoj et al., 1992) participants' evaluations of the level of survey's applicability and accuracy, and interviewers' evaluations of the level of respondent's cooperation and understanding were measured using four categorical variables.

These variables were four out of the 286 variables that were included in this survey. Sample consisted of 1000 people 700 from which were interviewed by phone and 300 face to face.

To measure the *respondents' evaluations* the following two questions were asked:

   1. APPLICABILITY : In general, do you feel that surveys usually serve a good purpose, or do you feel that they are usually a waste of time and money ?
     1. Good purpose
     2. Depends
     3. Waste of time and money

   2. ACCURACY : How often do you think you can trust the results of surveys ?
     1. Almost always
     2. Most of the time
     3. Some of the time
     4. Hardly ever

To measure the *interviewers' evaluations,* the interviewers answered the following two questions:

   3. COOPERATION : In general, what was the respondent's attitude toward the interview ?
     1. Friendly and interested
     2. Cooperative but not particularly interested
     3. Impatient and restless
     4. Hostile

   4. UNDERSTANDING : Was the respondent's understanding of questions good, fair, or poor ?

1. Good
2. Fair
3. Poor

From these four variables (and after recoding) we obtain a 36 cell cross-tabulation of observed frequencies.

Table 1: Cross-Tabulation of Observed Variables

			COOPERATION		
			Interested	Cooperative	Impatient, Hostile
APPLICABILITY	ACCURACY	UNDERSTANDING			
Good	Mostly True	Good	295	78	2
		Fair, Poor	25	11	0
	Not True	Good	90	34	1
		Fair, Poor	10	4	3
Depends	Mostly True	Good	92	46	1
		Fair, Poor	8	6	2
	Not True	Good	146	59	2
		Fair, Poor	6	9	3
Waste	Mostly True	Good	2	0	0
		Fair, Poor	0	2	0
	Not True	Good	9	7	1
		Fair, Poor	1	1	2

Due to some low frequencies and because of the close/identical meaning of some categories, we have collapsed several categories. Thus, we have merged the categories *Almost always* and *Most of the time* of the variable ACCURACY into the new category *Mostly True* and the categories *Some of the time* and *Hardly ever* into the new category *Not True*. In the case of the variables UNDERSTANDING and COOPERATION, we have recoded the last two categories in two new categories ("Fair, Poor" and "Impatient, Hostile").

One third of respondents (295) believed that survey results can be trusted and that surveys serve a good purpose. These respondents were also interested in surveys and they understand questions well. 146 respondents are also interested in surveys and understand questions, but they don't trust survey results and doubt about their applicability. Other cells in Table 1 are smaller and only 5 of them are empty.

The data reported in the Table 1 suggest that variables are not independent of one another. In the further analysis they were considered as parts of the common latent complex and were used to develop a typology of survey respondents(statistical test of independence follows).

3 Latent typology of respondents

To examine the relationships among four variables and to develop a typology, an exploratory latent class analysis was carried out. The latent class approach is the most appropriate for this purpose because of the categorical nature of the variables. We can say that Latent Class Analysis is analogous to modern factor analysis and based on a probabilistic approach (see McCutcheon, 1987).

In its most general form, Latent Class Analysis makes possible the characterization of multidimensional discrete latent variable from a cross-classification of two or more observed categorical variables (see McCutcheon, 1987:8). Latent Class Analysis can be also applied to analyze typologies, as a method for empirical characterization of a set of latent types within a set of observed indicators, what was the case in our study.

The first decision to be made concerns the number of latent classes to identify. An exploratory latent class analysis of data reported in Table 1 allows for identification of a latent variable with up to five classes (i.e., five latent types), since the number of degrees of freedom $((3*2*2*3)-(2+3+3+2-3)*T$; T is number of latent classes) must be positive.

The fit among each latent class model and the observed data is reported in Table 2, where Pearson chi-square ($\chi 2$), the likelihood ratio chi-square (L2) and degrees of freedom are reported. The decision criterion is set at the $p > .05$ alpha level, which is the standard practice for statistical modelling.

Table 2: Exploratory Latent Class Model

Model	L2	$\chi 2$	Degrees of Freedom	Decision at .05 Alpha Level
Complete Independence	206.335	449.744	36	reject
Two-class Model	60.485	115.511	22	reject
Three-class Model	19.319	16.564	16	accept
Four-class Model	11.309	8.893	11	accept
Five-class Model	9.384	8.322	6	accept

We rejected the "Complete Independence" model and accepted "Three-class model". Three-class model fits the observed data and lends itself best to a meaningful interpretation.

But, how large are those classes of respondents and what are their characteristics according to the observed variables? Results, obtained with MLLSA program (written by C.C. Clogg, 1977) are presented in Table 3.

Table 3 : Conditional Probabilities and Latent Class Probabilities for Three Classes

Manifest Variables		Ideal (I)	Respondent Type Delicate (II)	Skeptics (III)
APPLICABILITY	Good	.7974	.3093	.0978
	Depends	.2026	.5304	.8403
	Waste	.0000	.1603	.0620
ACCURACY	Mostly True	.7873	.3504	.1870
	Not True	.2163	.6496	.8130
COOPERATION	Interested	.7731	.1162	.6930
	Cooperative	.2269	.5690	.3070
	Impatient,Hostile	.0000	.3148	.0000
UNDERSTANDING	Good	.9151	.4144	.9733
	Fair,Poor	.0849	.5856	.0267
Latent Class Probabilities		.6682	.0564	.2754

Let us consider the latent class probabilities first. Latent class probabilities tell us what proportion of population is associated with each class. As we can see 67% of population is located in the first class, 27.5 % of population is located in the third class and only 5.6 % in the second class, which means that only one respondent out of twenty is of Type 2.

Conditional response probabilities enable us to associate the categories of the observed variables and the latent categories. We can associate the first latent category (class) with the category *Good* of the variable *APPLICABILITY*, with the category *Mostly True* of the variable *ACCURACY* and the category *Interested* of the variable *COOPERATION*. The highest probability is in the case of the category *Good* of the variable *UNDERSTANDING*. On the basis of these conditional probabilities the obtained class can be labelled as *Ideal* since the members of this class evaluate applicability and accuracy of surveys high and their understanding of questions and interest in surveys are evaluated high as well.

The third latent class can be associated with the category *Depends* of the variable *APPLICABILITY* and the category *Not True* of the variable *ACCURACY*. High probability can be also found in the case of the category *Interested* of the variable *COOPERATION* and the highest conditional probability at the category *Good* of the variable *UNDERSTANDING*. On the basis of these conditional probabilities the obtained class was named *Skeptics* since the members of this class evaluate applicability and accuracy low, but their understanding and interes are evaluated high.

And finally the second class. High probability can be found in the case of the category *Depends* and *Waste* of the variable *APPLICABILITY* and the category *Not True* of the variable *ACCURACY*. There is also high probability at the categories *Cooperative* and *Impatient, Hostile* of the variable *COOPERATION* and the category *Fair, Poor* of the variable *UNDERSTANDING*. On the basis of these conditional probabilities the obtained class was named *Delicate* since the members

of this class evaluate applicability and accuracy rather low and their understanding and cooperativness is evaluated as low.

In the next step of the analysis, each respondent was assigned to the appropriate latent class (according to the procedure described bellow).

Since the respondents with identical scores on the observed variables are considered to be in the same latent class, the assignment was carried out on cell by cell basis from the crosstabulation of the observed variables. For this purpose the differential contribution of each of the three latent classes to each of the cells of the observed contingency table was calculated and observations in each cell were assigned to the latent class with the largest conditional probability (see McCutcheon, 1987:36).

Since the assignment is probabilistic, we can expect that some units are "missclassified". In our case the percentage of correctly classified observations was 87.18.

4 Some characteristics of the obtained classes

In the final step the characteristics of the obtained classes of respondents were analyzed with respect to the variables presented in Table 4.

The significance of association was tested with χ^2 statistics (phi) and in the case of the variable EDUCATION also analysis of variance was used.

5 Conclusions and discussion

On the basis of respondents' and interviewers' evaluations we have obtained a typology with three latent classes: the "Ideal Respondents Class" which shares high positive level of evaluations of surveys, interes in surveys and understanding; the "Skeptic Respondents Class" which shares medium or low evaluations of surveys and high positive evaluations of interes in surveys and understanding, and the "Delicate Respondents Class" with medium or low evaluations of surveys and cooperation and understanding.

If we compare our results with the results of a similar study that was done in 1982 in the USA (McCutcheon, 1987:28-35) we can see that there were obtained three classes as well. American Ideal class is about the same size as ours. They obtained the Skeptic class as well, but that is their smallest class and it's approximately one third smaller than ours. The third class they found was named Believers since it is characterized by high respondents' evaluations of surveys and high evaluations of respondent's cooperation but low evaluations of respondents understanding. The similarities and differences between the USA typology and the

Slovene typology make sense to us, if cultural differences between USA and Slovenia are taken into account.

Table 4

VARIABLE		ASSOCIATION
SOCIO-DEMOGRAPHIC:		
Gender	1. Male	
	2. Female	NO
Age	1. 16-24	
	2. 25-34	
	3. 35-49	NO
	4. 50-64	
	5. 65 and more	
Education	1. No education	
	2. Less than primary	
	3. Primary	
	4. Secondary	YES
	5. College	
	6. Higher	
	7. University	
SURVEY DESIGN :		
Survey method	1. telephone	
	2. face to face	NO
Code of interviewer : A code was assigned to each interviewer		NO
Number of contacts : How many contacts had to be established before the interview was satisfactory completed ?		NO
SURVEY TOPIC :		
"Crime experience" : Has respondent been victimised in the last five years ?		
	1. Yes	
	2. No	NO
Q8 : Has the topic of crime came up in any conversation you have had with family,friends or colleagues in the last week ?		
	1. Yes	
	2. No	NO
Q50 : People have different ideas about the sentence which should be given to offenders. Take for instance the case of man 25 years old who is found guilty for burglary for the second time. The last he had stolen colour TV. Which of following sentences do you consider most appropriate for such case ?		
	1. Fine	
	2. Prison	
	3. Community sentence	NO
	4. Suspended sentence	
	5. Any other (please specify)	
RESPONDENTS' EVALUATIONS OF THIS PARTICULAR SURVEY:		
Difficulty : How difficult do you find this interview ?		
	1. Very easy	
	2. Easy	
	3. Neither easy, nor difficult	NO
	4. Difficult	
	5. Very difficult	
Usefulness : Do you believe that the results of this survey could serve a good purpose ?		
	1. Yes	YES
	2. No	
Participation : Would you participate in survey like this again if you are asked to ?		
	1. Yes	YES
	2. No	

We have found associations between the variables Usefulness and Participation and the obtained classes. These two variables can be considered as "control" variables, since they are in correspondence with the variables that were used to develop typology.

Thus class membership seems to be associated only with the level of education (χ^2 = 25.994, Sig. .00105). Average level of education is the lowest in the Delicate Class and the highest in the Sceptics Class. When we are dealing with respondents with lower education, the probability that they will evaluate surveys negatively and express lower level of cooperation and understanding is higher. But so far that is not a reason to exclude them from our sample. It does not mean that their answers are less valuable. Interviewers must be prepared for such respondents and should invest stronger effort to obtain answers from them.

In further analyses the effect of class membership on given answers should be studied - particularly the number of nonresponses given in each class.

References

[1] Ferligoj A. et al. (1992): *Poročilo o prvi fazi raziskave o kriminaliteti v Ljubljani*, Ljubljana: IDV-FDV.

[2] Gnidovec M. (1993): *Anketa in anketiranec*. Ljubljana: FDV (Diplomska naloga).

[3] Hagenaars J.A. (1989): Searching for ideal types: the potentialities of latent class analysis. *European Sociological Review*, 5, 81-86.

[4] McCutcheon A.L. (1987): *Latent Class Analysis*. Beverly Hills: Sage Publication.

[5] McKinney J.C. (1966): *Constructive Typology and Social Theory*. New York: Appleton Century Crofts.