Non-response in Surveys

DSA2101 Tutorial 07

13/10/2021

Dealing with Non-response Bias

Raking

There are two main types of non-response in survey responses. The first type occurs when the respondent leaves gaps in his/her questionnaire; most questions would be filled in, but some might be empty, either by choice or by accident. This sort of non-response is dealt with using *multiple imputation*. Typically, 5 - 10 full datasets are created, and the results from these datasets are averaged.

The second type of non-response occurs when some people do not respond at all. The typical method for dealing with such situations is to weight the respondents based on certain auxiliary variables. One such technique is known as *raking*. In our case, we shall use nationality and on/off-campus stay as the auxiliary variables.

Suppose that our population consists of 1500 individuals, but we have only managed to obtain measurements for 1000 of them. Fortunately, we have also recorded auxiliary variables A and B regarding these 1000 respondents. The tables below break the respondents down according to A and B, in the sample and in the population.

Sample

	B1	B2	B3	Total
A1	20	40	40	100
A2	50	140	310	500
A3	100	50	50	200
A4	30	100	70	200
Total	200	330	470	1,000

Population

	B1	B2	B3	Total
A1	80	40	55	175
A2	60	150	340	550
A3	170	60	200	430
A4	55	165	125	345
Total	365	415	720	1,500

The table indicates that there were 30 respondents from the A4-B1 category, but there were in fact 55 individuals in total (in the population) in this category.

It seems intuitive that we should weight the readings of the 20 respondents in A1-B1 by 4, and the readings of the 40 respondents in A1-B2 by 1. However, this method, known as *cell weighting*, would lead to a highly unstable set of weights, which would manifest as high variability in the final score estimates.

Moreover, cell weighting requires us to know the cross-classification of each respondent. When we only know the marginal categorisation of the sample or the population, then we use raking. Raking iteratively fits weights row-wise and then column-wise, until the column margins and row margins are close to the population values.

Step 1

	B1	B2	B3	Total
A1	35.00	70.00	70.00	175.00
A2	55.00	154.00	341.00	550.00
A3	215.00	107.50	107.50	430.00
A4	51.75	172.50	120.75	345.00
Total	356.75	504.00	639.25	1,500.00

Step 2

	B1	B2	B3	Total
A1	35.81	57.64	78.84	172.29
A2	56.27	126.81	384.08	567.16
A3	219.97	88.52	121.08	429.57
A4	52.95	142.04	136.00	330.99
Total	365.00	415.00	720.00	1,500.00

In the first step of raking, the rows are weighted to match the A population marginals. In the second step of raking, the columns are weighted to match the B population marginal totals. This continues until the weights converge:

A: Cell weighting

	B1	B2	B3
A1	4.00	1.00	1.38
A2	1.20	1.07	1.10
A3	1.70	1.20	4.00
A4	1.83	1.65	1.79

B: Raking

	B1	B2	B3
A1	1.81	1.45	2.02
A2	1.08	0.87	1.21
A3	2.20	1.76	2.45
A4	1.83	1.47	2.04

Consider the set of weights from raking for row A1:

$$1.81 \times 20 + 1.45 \times 40 + 2.02 \times 40 = 175$$

Meanwhile, for column B1:

$$1.81 \times 20 + 1.08 \times 50 + 2.2 \times 100 + 1.83 \times 30 = 365.1$$

The population totals in the first row and first column are very similar. Compare the raking weights with the directly computed cell weights - the former are smaller in magnitude.

For more information about weighting methods, raking and the package we use, please refer to Kulas et al. (2018), Kalton and Flores-Cervantes (2003), and Lumley (2011).

Example Correction for Non-response

Our dataset is based on the Oxford Happiness Questionnaire. There are 29 questions. The scoring can be obtained and interpreted from the accompanying pdf.

```
library(survey)
respondents <- readRDS("../data/survey_example.rds")
des1 <- svydesign(id=~id, data=respondents, fpc = ~fpc)
pop.gender <- data.frame(gender=c("M", "F"), Freq=c(70, 30))

svymean(~happ, design=des1)
##        mean     SE
## happ 4.1851 0.0903

des1.raked <- rake(des1, sample=list(~gender), population=list(pop.gender))
svymean(~happ, design=des1.raked)
##        mean    SE
## happ 3.9068 0.027
```

Tasks:

1. Compute the Happiness Score for each individual.
2. Correct for non-response. Summarise the overall happiness score of our class.
3. Present 2 plots based on the data, along with your explanations/comments.
4. Here are some ideas for analysis:
    - Identify outlier responses.
    - Find clusters within the class (use hierarchical clustering from topic 06)
        - What dissimilarity measure to use?
    - Study spread of individual happiness scores

References

Kalton, Graham, and Ismael Flores-Cervantes. 2003. "Weighting Methods." *Journal of Official Statistics* 19 (2): 81.

Kulas, John T, David H Robinson, Jeffrey A Smith, and Donald Z Kellar. 2018. "Post-Stratification Weighting in Organizational Surveys: A Cross-Disciplinary Tutorial." *Human Resource Management* 57 (2): 419–36.

Lumley, Thomas. 2011. *Complex Surveys: A Guide to Analysis Using r*. Vol. 565. John Wiley & Sons.