This module is part of the

Memobust Handbook

on Methodology of Modern Business Statistics

26 March 2014

Theme: Measuring Coding Quality

Contents

General section

1. Summary

Two indicators are usually adopted to measure quality of coding: coding rate (percentage of coded texts on the total of texts to be coded) and precision rate (percentage correct coded texts on the total of coded texts). The quality analysis is usually based on coding texts multiple times and reconciling different codes assigned to the same texts, usually based on a sample of the coded descriptions. The expected values of these rates are different depending on some factors like the complexity of the classification and the detail level of the codes to be assigned.

2. General description

The quality of coding can be measured with two indicators:

- Coding rate (efficacy) → percentage of coded texts on the total of texts to be coded;
- Precision rate (accuracy) → percentage correct coded texts on the total of coded texts.

These rates are suitable either if coding is made automatically (both AUC or CAC) or manually.

The results of the analysis of coding quality requires different approaches depending on which of these two is selected: in the first case, when the results do not fulfil the expectations, the software application and/or the informative base must be updated, while in the second one the further training of interviewers/coders can be necessary.

The quality analysis is usually based on the verification of the coding, which means coding again texts and reconciling different codes assigned to the same texts. Naturally:

- if automatic coding was used, texts will be coded again by human coders (manually or with assisted coding);
- if texts were coded by human coders (manually or with assisted coding) they will be coded again automatically or with the intervention of different coders.

For the precision rate, it is assumed that when the original code and the verification code are equal, the code is correct, otherwise the reconciliation process must be performed by a different expert coder.

Concerning the expected values of these rates, different factors must be considered such as the complexity of the classification and the detail level of the codes to be assigned (Macchia and Murgia, 2002). On the other hand, it also has been noticed that different types of respondents, using the same classification, can have an impact on the final coding rate. For instance, in the experience of Istat, the coding rate of the economic activity responses has always been higher in business surveys than in households or individuals. This is due to the fact that the concept of economic activity is closer to respondents of the first type of surveys than to the latter one; as a result, less precise responses are given (Colasanti et al., 2009).

Finally, the quality analysis is usually conducted on a sample of texts. The sample for verification of coding can be selected in different ways.

Statistics Sweden, for instance, conducts the verification process for at least five percent of the coded records (this threshold of five percent is not statistically motivated, but a requirement for fulfilment of ISO 20252) (Svensson, 2012).

This quality control is made in each relevant survey for data coded through a computer-assisted manual procedure, while once every three year for data coded through an automatic coding procedure.

In Istat a different method is used for the verification of automated coding results, when the amount of processed texts is big: a sample of ‘different’ texts is checked (D’Orazio and Macchia, 2002). In practice, in order to avoid analysing more than once the same texts, “different” texts are identified through a kind of “raw normalisation”, so to delete from descriptions the articles, the conjunctions, the prepositions and the suffixes (in practice all the elements that determine the gender of words, the singular/plural, etc.). Then the occurrence of ‘equal’ texts is calculated and classes of occurrences are defined (texts are considered ‘equal’ after a process of raw normalisation). Then texts are stratified according to their frequency of occurrence; then, within each stratum, a simple random sample (without replacement) of texts is selected. The strata coincide with the previously defined classes of occurrences. This implies that the sample contains only different texts, but each of them has a different weight according to its class of occurrence. In this way the work of expert coders is reduced because they will never analyse more than once a text, which, on the other hand, could correspond to a certain number of collected responses.

3. Design issues

4. Available software tools

5. Decision tree of methods

6. Glossary

For definitions of terms used in this module, please refer to the separate “Glossary” provided as part of the handbook.

7. References

Colasanti, C., Macchia, S., and Vicari, P. (2009), The automatic coding of Economic Activities descriptions for Web users. NTTS 2009.

D’Orazio, M. and Macchia, S. (2002), A system to monitor the quality of automated coding of textual answers to open questions. *RESEARCH IN OFFICIAL STATISTICS (ROS)*, N.2 2002.

Macchia, S. and Murgia, M. (2002), Coding of textual responses: various issues on automated coding and computer assisted coding. Journée d’Analyse des Données Textuelles JADT, Saint Malo.

Svensson, J. (2012), Quality control of coding of survey responses in Statistics Sweden. European Conference on Quality in Official Statistics Q2012.

Interconnections with other modules

8. Related themes described in other modules

1.

9. Methods explicitly referred to in this module

1.

10. Mathematical techniques explicitly referred to in this module

1.

11. GSBPM phases explicitly referred to in this module

1. GSBPM sub-process 5.2

12. Tools explicitly referred to in this module

1.

13. Process steps explicitly referred to in this module

1.

Administrative section

14. Module code

Coding-T-Measuring Coding Quality

15. Version history

Version	Date	Description of changes	Author	Institute
0.1	20-07-2012	first version	Stefania Macchia	Istat (Italy)
0.2	21-11-2012	second version (following first revision)	Stefania Macchia	Istat (Italy)
0.3	25-10-2013	third version (following EB review 04-10-2013)	Stefania Macchia	Istat (Italy)
0.3.1	29-10-2013	preliminary release		
1.0	26-03-2014	final version within the Memobust project		

16. Template version and print date

Template version used	1.0 p 4 d.d. 22-11-2012
Print date	21-3-2014 18:08