

TEACHING OF SURVEY STATISTICS: COMPLEX SURVEY ANALYSIS AND STRUCTURAL EQUATION MODELS

Maria Valaste
University of Helsinki, Finland, maria.valaste@helsinki.fi

Sample surveys are essential tools in a modern society to provide information of different areas. Surveys are used to produce Official Statistics but also e.g. to gain insight into population attitudes. At the moment emerging data sources and developments of big data and open data provides new opportunities (Ridgway 2016) but this data revolution also challenges the data users.

Often the survey data are collected by a complex sampling design e.g. involving stratification, clustering and unequal inclusion probabilities. Thus, the estimators should be constructed so that the complexities of the sampling design are accounted for. Structural equation models are often used to assess unobservable latent constructs (Bollen 1989). Structural equation modelling (SEM) is a powerful technique that is used to analyse structural relationships. SEM includes e.g. confirmatory factor analysis, path analysis, and latent growth modelling.

This paper considers some implications of the data revolutions for survey statistics education and describes the methodological aspects of Structural Equation Models when sampling design is complex. Practical examples, that are produced for the forthcoming Survey Statistics courses, are presented. In the examples the design variables are included in SEM with R package lavaan.survey (Oberski 2019; Lumley 2010, 2004).

References

- Bollen K (1989). Structural Equations with Latent Variables. John Wiley & Sons, New York.
Lumley T (2004). Analysis of Complex Survey Samples. Journal of Statistical Software, 9(8).
Lumley T (2010). Complex Surveys: A Guide to Analysis Using R. John Wiley & Sons, New York.
Oberski D. (2019) lavaan.survey. Complex Survey Structural Equation Modeling (SEM). R package version 1.1.3.1, URL: <https://cran.r-project.org/web/packages/lavaan.survey/index.html>
Ridgway, J. 2016, Implications of the Data Revolution for Statistics Education, International Statistical Review, 84(3), s. 528-549. DOI 10.1111/insr.12110.