

Meaningful latent class analysis: Essential statistical frameworks and best practices for the applied researcher

Johan Lyrvall^a, Jennifer Oser^b and Zsuzsa Bakk^c

^aDepartment of Economics and Business, University of Catania, Catania, Italy, ORCID 0000-0002-1863-8147; ^bDepartment of Politics and Government, Ben-Gurion University of the Negev, Beer Sheva, Israel, ORCID 0000-0002-1531-4606; ^cDepartment of Methodology and Statistics, Leiden University, Leiden, Netherlands, ORCID 0000-0001-9352-4812

ARTICLE HISTORY

Compiled June 5, 2026

ABSTRACT

This article integrates existing theoretical and applied teaching resources to provide an updated overview of best practices for latent class analysis (LCA), focusing on applied researchers. We first clarify the unique usefulness of LCA compared to related latent variable models. We then focus on three fundamental research questions relevant for the applied researcher: (1) *do I have data with a possible clustering structure that can be analyzed using LCA?* (2) *what optimal number of LCs describe the possible clustering structure in my data?* and (3) *can the identified LCs in my data be predicted based on external covariates?* We conclude with summary guidance regarding modeling extensions for more complex LCA topics that are highly relevant for applied researchers, namely differential item functioning (DIF) and multilevel modeling. The topics are illustrated through empirical examples, with provision of replication code. An overview of software options for LCA is provided.

KEYWORDS

Latent class analysis; model selection; stepwise estimation; multilevel modeling; differential item functioning; software