Testing for serial correlation in hierarchical linear models

Javier Alejo, Gabriel Montes-Rojas, Walter Sosa-Escudero

PII: S0047-259X(17)30728-5
DOI: https://doi.org/10.1016/j.jmva.2017.11.007
Reference: YJMVA 4308

To appear in: *Journal of Multivariate Analysis*

Received date : 15 October 2016

Testing for serial correlation in hierarchical linear models

Javier Alejo[a], Gabriel Montes-Rojas[b], Walter Sosa-Escudero[c,*]

[a] Universidad Nacional de La Plata and CONICET
[b] CONICET-IIEP-BAIRES-Universidad de Buenos Aires and Universitat Autònoma de Barcelona
[c] Universidad de San Andrés and CONICET

Abstract

This paper proposes a simple hierarchical model and a testing strategy to identify intra-cluster correlations, in the form of nested random effects and serially correlated error components. We focus on intra-cluster serial correlation at different nested levels, a topic that has not been studied in the literature before. A Neyman's $C(\alpha)$ framework is used to derive LM-type tests that allow researchers to identify the appropriate level of clustering as well as the type of intra-group correlation. An extensive Monte Carlo exercise shows that the proposed tests perform well in finite samples and under non-Gaussian distributions.

Keywords: Clusters, random effects, serial correlation.
2010 MSC: 62P20, 62H15, 62J10

1. Introduction

Intra-group correlation has received considerable interest. When the data can be grouped in clusters, observations within each group are typically dependent. As highlighted by Bertrand et al. [13], failure to take such dependence into account can lead to misleading statistical inferences; this concern dates back to Moulton's seminal paper [27]. Consequently, the problem of whether and how to cluster observations is related to identifying: (a) the "finest" grouping structure that leaves out more independent groups and, (b) the type of intra-cluster correlation, in the form of either random effects, serial correlation or both.

Practitioners typically rely on "cluster robust methods", e.g., on estimates of standard errors that explicitly allow for correlations among observations within a group. However, the consistency of this approach depends on the number of independent groups growing large. This is problematic when grouping obeys a nested structure, as would be the case of students in a given class, school, etc. In such a scenario a safer strategy that allows for arbitrary correlations at a higher level (e.g., at the school instead of the class level) comes at the price of leaving fewer independent groups, making asymptotic approximations less reliable. In their recent survey, Cameron and Miller [15] point out that "there is no general solution to this trade-off, and there is no formal test of the level at which to cluster. The consensus is to be conservative and avoid bias and use bigger and more aggregated clusters when possible, up to and including the point at which there is concern about having too few clusters" [p. 321].

We are thus concerned with the appropriate level of clustering in a hierarchical linear model. Proper identification of the source of intra-group correlation is important to decide how to estimate the parameters of interest and their standard deviation. For example, when only random effects cause intra-cluster correlation, feasible generalized least-squares (GLS) strategies as in [6] might offer a simple and convenient alternative over cluster robust methods in the few groups scenario. The most obvious source of intra-group correlation arises when all observations within a group share an unobserved common factor, hence all observations in a group are equicorrelated in the sense that all pairwise correlations are the same. Tests for nested random effects have been studied in [8]. Another source of intra-cluster correlation that has received particular consideration in [13] is time, i.e., cluster correlation is induced when

*Corresponding author. Universidad de San Andrés, Vito Dumas 284, B1644BID Victoria, Buenos Aires, Argentina.
Email addresses: javieralejo@gmail.com (Javier Alejo), gabriel.montes@fce.uba.ar (Gabriel Montes-Rojas), wsosa@udesa.edu.ar (Walter Sosa-Escudero)