

# Detecting heterogeneity in logistic models

Katalin Balázs, István Hidegkuti, Paul De Boeck

Katholieke Universiteit Leuven, Tiensestraat 102, B-3000 Leuven, Belgium

e-mail: [Katalin.Balazs@psy.kuleuven.ac.be](mailto:Katalin.Balazs@psy.kuleuven.ac.be), [Istvan.Hidegkuti@psy.kuleuven.ac.be](mailto:Istvan.Hidegkuti@psy.kuleuven.ac.be),

[Paul.DeBoeck@psy.kuleuven.ac.be](mailto:Paul.DeBoeck@psy.kuleuven.ac.be)

**Keywords:** heterogeneity, logistic models

## Abstract

The data we are interested in are binary person-by-item data. The basic model we start from is a logistic regression model with fixed parameters (intercept and slopes) and identical parameter values for all persons. The problem is that in most cases the data are correlated, or in other words stated, that there is extra binomial variation. Heterogeneity of the parameters can explain this extra variation. For example in an item-response model, the intercept is assumed to vary over persons.

We will study different sources of heterogeneity: fixed group differences in the intercept and slopes, and random intercepts and random slopes. We will investigate several methods to explore the heterogeneity: cluster analysis of the raw data, deviance-based statistics, individual analyses (one per person), and principal component analysis of the raw data residuals and of the cross-product ratio residuals. For model estimation we used SAS NLMIXED and GENMOD. We will report on a simulation study with data generated from various types of models, differing as to the kind and degree of heterogeneity. The methods just mentioned are then compared as to their diagnostic performance.