# A new framework for goodness of fit testing in IRT

Maydeu-Olivares, Albert
Dept. of Psychology. University of Barcelona.
P. Valle de Hebron, 171
08035 Barcelona (Spain)
amaydeu@psi.ub.es
Joe, Harry
Dept. of Statistics. University of British Columbia.
333-6356 Agricultural Rd.
Vancouver, BC V6T 122 (Canada)
harry@stat.ubc.ca

## Abstract

High-dimensional binary contingency tables tend to be sparse and standard goodness-of-fit statistics such as $X^2$ cannot be used without pooling cells. As an improvement on arbitrary pooling, for goodness-of-fit of large $2^n$ contingency tables, we propose a class of quadratic form statistics, $M_r$, based on the residuals of margins or multivariate moments up to order $r$. This class includes from $M_1$ to $M_n$, where $M_1$ can be used to assess how well an IRT model reproduces the means of the binary variables to $M_n$ which is a full information test statistic. All of the members of $M_r$ are asymptotically chi-square with known degrees of freedom for any consistent and asymptotically normal estimator, including limited information estimators such as those implemented in LISREL or NOHARM. Thus, one can now compare the fit of a model estimated using LISREL against the fit of another model estimated using BILOG using either a limited information test such as $M_2$ or a full information test, $M_n$. Furthermore, we show that in the case of BAN estimators such as full information maximum likelihood, $M_n = X^2$.
For small $r$, $M_r$ have better small sample properties and are more powerful than $X^2$ and $G^2$. This is illustrated with a two-parameter logistic model.