BJPsych The British Journal of Psychiatry
1. doi: 10.1192/bjp.bp.112.119032

Online supplement

This online supplement contains further details of statistical methods and procedures for assessing goodness of fit.

Assessing the fit of confirmatory factor analysis (CFA) models

For the CFA models, the chi-squared test, the comparative fit index (CFI),[43] the Tucker–Lewis index (TLI)[44] and the root mean square error of approximation (RMSEA)[45] were used. Hoyle & Panter[46] recommend that a non-significant chi-squared test, along with TLI and CFI values of $\leqslant$0.95, and a RMSEA value of $\leqslant$0.05, indicates acceptable model fit. Bollen,[47] however, notes that the chi-squared statistic is highly sensitive to large sample sizes and may overestimate the lack of fit of a structural model. Thus, the chi-squared test should be viewed in conjunction with the other fit indices. A chi-squared difference test can be computed between nested factor models to examine whether a less stringent set of model constraints improves the model fit. It is important to note that in Mplus version 6, the chi-squared value obtained for WLSMV estimation (as conducted in this study) cannot be used for chi-squared difference testing in the usual manner and has to be adjusted using the DIFFTEST command.

Estimating and assessing the fit of latent class analysis (LCA) and factor mixture model analysis (FMMA)

One problem that may arise when using algorithms to produce maximum-likelihood estimation is the presence of *local maxima*. This means that during the estimation process, there are several solutions around which a model may converge (i.e. local maxima, in which a model fits the data in an apparently satisfactory way), but there is only one best solution (i.e. *the global maximum*). The algorithm stops when a maximum is reached, but it cannot distinguish the global maximum from a local maximum.[48] If a model converges around a particular local maximum, instead of the global maximum, the best-fitting solution can be missed.[49] To ensure successful convergence on the global maximum solution, LCA and FMMA models were estimated with different sets of random starting values (i.e. 500 random sets of starting values were used in the initial stage, and 20 optimisations were used in the final stage of convergence). All models were inspected to identify whether the log-likelihood value for each model was replicated several times, as this increases confidence that the solution obtained is not a local maximum.[50] We report circumstances where the log-likelihood was not replicated (in Table 5).

There is no single definitive method for deciding on the optimal number of latent classes,[51] and several statistical indices are conventionally used to assess the fit of the models. For the LCA and FMMA models, the log-likelihood, the Akaike information criterion (AIC),[52] the Bayesian information criterion (BIC)[53] and the sample-size adjusted BIC (SSABIC)[54] were used as goodness-of-fit indicators. A high log-likelihood value in conjunction with lower values on the AIC, BIC and the SSABIC reflect a good-fitting model. The BIC has been shown to be more reliable than the other information criteria.[50] For LCA models, the Lo, Mendell and Rubin likelihood ratio test (LMR-LRT)[55] and the entropy[56] can also be useful in determining the best-fitting model. The LMR-LRT compares models with different number of classes: a non-significant value suggests that the model with one fewer class is a better explanation of the data. The entropy statistic, which ranges from 0 to 1, is a standardised summary measure of the classification accuracy of placing participants into classes based on their model-based posterior probabilities. Higher entropy values reflect better classification of individuals.[56]

Additional references

43 Bentler PM. Comparative fit indices in structural models. *Psychol Bull* 1990; **107**: 238–46.

44 Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika* 1973; **38**: 1–10.

45 Steiger JH. Structural model evaluation and modification: an interval estimation approach. *Multivar Behav Res* 1990; **25**: 173–80.

46 Hoyle RH, Panter AT. Writing about structural equation models. In *Structural Equation Modeling, Concepts, Issues, and Applications* (ed. RH Hoyle): 158–76. Sage Publications, 1995.

47 Bollen KA. *Structural Equations with Latent Variables*. Wiley, 1989.

48 Neely-Barnes S. Latent class models in social work. *Soc Work Res* 2010; **34**: 114–21.

49 Vermunt JK, Magidson J. Latent class cluster analysis. In *Applied Latent Class Models* (eds JA Hagenaars, AL McCutcheon): 89–106. Cambridge University Press, 2002.

50 Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. *Struct Equ Modeling* 2007; **14**: 535–69.

51 Dunn KM, Jordan K, Croft PR. Characterizing the course of low back pain: a latent class analysis. *Am J Epidemiol* 2006; **163**: 754–61.

52 Akaike H. Factor analysis and the AIC. *Psychometrika* 1987; **52**: 317–33.

53 Schwartz G. Estimating the dimension of a model. *Ann Stat* 1978; **6**: 461–4.

54 Sclove SL. Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika* 1987; **52**: 333–43.

55 Lo Y, Mendell N, Rubin DB. Testing the number of components in a normal mixture. *Biometrika* 2001; **88**: 767–78.

56 Ramaswany V, De Sarbo W, Reibstein D, Robinson W. An empirical pooling approach for estimating marketing mix elasticities with PIMS data. *Marketing Science* 1993; **12**: 103–24.