Analysis Conditional on Covariates

Beginning with Mplus Version 6.1, TYPE=GENERAL has been changed for models with all continuous outcomes and maximum-likelihood estimation so that the likelihood is considered for the outcomes conditioned on the covariates (x variables, exogenous observed variables). This is in line with regression analysis where the distribution of the covariates is not part of the model, but the model is expressed for the outcomes given the covariates. As in regression, the covariates are assumed to be freely correlated, but the parameters of the covariate part of the model are not estimated. In a model with all continuous-normal outcomes, the unconditional approach of not conditioning on the covariates and letting the parameters of the covariate part of the model be estimated was used initially in SEM. This leads to the same estimates, standard errors, and test of model fit as in the conditional approach. The unconditional approach, however, does not generalize to other model settings, such as with categorical outcomes and mixture modeling, without adding unnecessary assumptions. Using the conditional approach therefore makes Mplus take a consistent approach across model types. Following are some implications of this change.

Because the loglikelihood is reported for the outcomes given the covariates, its metric is not changed by changing the covariates, so that models can be compared with respect to BIC as long as they have the same outcomes.

The default parameterization does not covary covariates with other variables such as exogenous factors. If a user wants such variables to be related, this has to be specified using either ON or WITH. Because covariates are characterized as exogenous variables, using ON with the covariates on the right-hand side is a natural approach.

Starting with Version 6, Mplus deletes individuals who have missing data on one or more covariates as the default. If a user does not want this to happen, a distributional assumption such as normality has to be added for the covariates. Using maximum-likelihood estimation in Mplus this is accomplished by including the covariates in the model. This is done by mentioning means, thresholds, or variances of the covariates in the MODEL command. An alternative is to use multiple imputation in Mplus, where data sets are created from a Bayesian analysis imputing values for the covariates. In either case, the modeling has been extended to make assumptions about the covariates. The effects of such assumptions has been studied in Chapters 9 and 10 of the 2016 book Regression and Mediation Analysis using Mplus by Muthen, Muthen, and Asparouhov.