1

Mplus Estimation Strategies

Mplus offers a wide range of estimation methods. Maximum likelihood (option ML in Mplus) is the traditional approach used in many SEM applications, but it assumes multivariate normality and requires larger sample sizes because it is based on asymptotic theory. Mplus invokes default methods that are dependent on the characteristics of the model tested and the nature of the data. They can be overridden. In this primer, I first provide a general summary of major Mplus estimation options and then I discuss some of them in more depth. I do not discuss Bayes estimation because I cover it in Chapters X and X. Technical descriptions of the approaches described below are in the Mplus Technical Manual that is distributed with the Mplus software.

ESTIMATION ALGORITHMS IN MPLUS: GENERAL DESCRIPTION

Here are the major forms of estimation offered in Mplus:

ML: Maximum likelihood parameter estimates with conventional standard errors and chi-square test statistic.

MLM: Maximum likelihood parameter estimates with standard errors and a mean-adjusted chi-square test statistic that are robust to certain forms of non-normality. The MLM chi-square test statistic is also called the Satorra-Bentler chi-square.

MLMV: Maximum likelihood parameter estimates with standard errors and a mean- and variance-adjusted chi-square test statistic that are robust to certain forms of non-normality.

MLR: Maximum likelihood parameter estimates with standard errors and a chi-square test statistic that are robust to certain forms of non-normality and certain forms of non-independence of observations when used with complex samples with clustering and/or weights (TYPE=COMPLEX). The MLR standard errors are computed using a sandwich estimator. The MLR chi-square test statistic is asymptotically equivalent to the Yuan-Bentler T2* test statistic.

MLF: Maximum likelihood parameter estimates with standard errors approximated by first-order derivatives and a conventional chi-square test statistic

MUML: Muthén's limited information parameter estimates, standard errors, and chi-square test statistic

WLS: Weighted least square parameter estimates with conventional standard errors and chi-square test statistic that use a full weight matrix. WLS is also referred to as ADF (arbitrary distribution function) when all outcome variables are continuous, in which case it requires large sample sizes (e.g., more than 4,000).

WLSM: Weighted least square parameter estimates using a diagonal weight matrix with standard errors and mean-adjusted chi-square test statistic that use a full weight matrix. WLSM generally outperforms WLS.

WLSMV: Weighted least square parameter estimates using a diagonal weight matrix with standard errors and mean- and variance-adjusted chi-square test statistic that uses a full weight matrix. Standard errors and parameter estimates are generally the same as WLSM. WLSMV uses specialized methods to calculate the model degrees of freedom (see below). It tends to work better than WLSM in many contexts.

ULS: Unweighted least squares parameter estimates. This also is called arbitrary generalized least squares function.

ULSMV: Unweighted least squares parameter estimates with standard errors and a mean- and variance-adjusted chi-square test statistic that use a full weight matrix

GLS: Generalized least squares parameter estimates with conventional standard errors and chi-square test statistic based on a normal-theory based weight matrix. The weights are derived from the inverse of the sample covariance matrix.

BAYES: Bayesian posterior parameter estimates with credibility intervals and posterior predictive checking.

ESTIMATION ALGORITHMS IN MPLUS: SOME SPECIFICS

In this section, I focus first on maximum likelihood based Mplus estimation methods that are appropriate for both normal and non-normal, continuous data. The methods are MLM, MLMV, MLR, and MLR with INFORMATION=EXPECTED. In traditional maximum likelihood analysis under normality, a type of likelihood ratio test is applied that yields a statistic (the traditional chi square test) with a sampling distribution that is chi square distributed with a specified degrees of freedom (df). Under non-normality, the sampling distribution is no longer chi square distributed, but it can be close to being chi square distributed depending on the nature and degree of normality violations. The MLM estimator adjusts the statistic under non-normality so that the sampling distribution of the statistic coincides to the mean of a chi square distribution with appropriate degrees of freedom. The MLMV estimator introduces adjustments so that the test statistic has a sampling distribution that coincides in both the mean and the variance of a chi square distribution with appropriate degrees of freedom (Asparouhov & Muthén, 2010). It applies a classic Satterthwaite variance correction to MLM. The two methods yield the same parameter estimates and standard errors but they have different values of their chi square test statistic. Since many global fit indices use the chi square test statistic in their calculation, it follows that the methods will often produce different values of them as well. Simulation research suggests that the MLMV estimator generally performs better than the MLM method in terms of omnibus model Type I error rates. The MLMV test statistic does reasonably well in a wide range of scenarios, with the exception of large models coupled with extreme non-normality and small sample sizes (Maydeu-Olivares, 2017b). The MLM and MLMV methods require complete data or the use of listwise deletion. This can result in reduced power if the reduced sample size that results from listwise deletion is non-trivial. In this sense, both MLM and MLMV are less desirable than the MLR options.

The default MLR option in Mplus uses sandwich estimation with observed information in the outer block of the sandwich and cross products in the inner block (Asparouhov & Muthén, 2005). It does not require listwise treatment of missing data, but the non-default option MLR with INFORMATION=EXPECTED does, making MLR the more general method for applied research, such as RETs. Several simulations suggest that the Mplus default MLR option generally performs better than the MLR with INFORMATION=EXPECTED in terms of yielding less biased standard errors for parameter tests in the face of non-normality (e.g., Maydeu-Olivares, 2017b). MLR tends to be conservative in terms of parameter Type I errors when sample sizes are smaller, near 100 (Bentler & Yuan, 1999).

The WLS method, when applied to all continuous outcomes, is the same as the classic ADF (Arbitrary Distribution Function) method, which offers reasonably good protection against non-normality but it is sample size demanding (usually requiring N > 1,000 or 1,500 if not more). It also requires listwise deletion of missing data. WLS can be used for ordinal outcomes. For continuous variables, MLR is generally preferable.

For ordinal outcomes, WLS and the WLSMV algorithms can be used in addition to MLR. Simulation studies suggest that WLSMV outperforms both WLS and WLSM (especially for smaller sample sizes) across a wide range of analytic scenarios. However, it too is somewhat sample size demanding, often requiring N > 500.

ULS (unweighted least squares) and its variants as well as GLS (generalized least squares) require listwise deletion of missing data A drawback of ULS is that it is not scale invariant; results can change with simple linear transformations of the measures. It also requires all variables be measured on the same metric. GLS is scale invariant and is similar to ML (they are asymptotically equivalent). GLS is computationally easier to apply and faster to compute, but these advantages are less important with modern, high speed computers.

With binary outcomes, using ML or a variant of it (typically MLR) yields a logistic regression model. Using WLS or a variant of it (typically WLSMV) yields a probit regression model. Declaring the binary variable as continuous and using MLR results in a modified linear probability model. For ordinal regression, using ML or MLR yields a logistic-based ordinal regression model. Using WLS or WLSMV yields an ordered probit model. For ordinal regression, MLR performs similarly to WLSMV with larger sample sizes, but MLR seems to work somewhat better for smaller sample sizes.

Given all the above, MLR and WLSMV will generally be your "workhorse" estimation methods, with more nuanced choices based on more intimate knowledge of one's substantive questions and the analytics of the situation. I outline these considerations in different chapters of my book, when we encounter analytic situations relevant to them. For a discussion of choosing an estimation method for categorical or ordinal outcomes, see Muthén and Asparouhov (2015).

Bootstrapping can be used in Mplus. The most common approach is to use ML estimation in conjunction with percentile bootstrapping. Mplus also offers biased corrected bootstrapping. Nevitt and Hancock (2001) found that bootstrapping with an ML estimator for a range of SEM factor models did not work as well as parameter tests using MLR for small samples near N=100, but that for N=200 or more, bootstrapping provided somewhat better Type I error control. There is controversy about the how large a sample size is needed for effective bootstrapping, especially with mediation models focused on indirect effects. Some studies suggest Ns can be as low as 100 and others suggest larger sample sizes are needed (Shrout & Bolger, 2002; Nevitt & Hancock, 2001). Although Koopman et al. (2015) caution against bootstrapping with N less than 100 based on their simulations, most of their complaints focus on low power that is true of small sample size scenarios to begin with. More work is needed in this area.

TWO STAGE LEAST SQUARES ESTIMATION

Although it is not available in Mplus, R offers an SEM package that uses two stage least squares (2SLS) estimation (Bollen, 2019). The 2SLS estimator is consistent, asymptotically unbiased, asymptotically normally distributed, and asymptotically efficient among limited information estimators. The version proposed by Bollen yields as an accurate asymptotic covariance matrix without normality assumptions. 2SLS is noniterative, which is a strength. Research is needed on its performance in representative RET contexts.