

## Alternative Estimation Methods

### ML

Remember that the usual approach to estimating fit and coefficients in SEM is the maximum likelihood (ML) approach. ML uses derivatives to minimize the following fit function:

$$F_{ML} = \log|\Sigma(\theta)| + tr(S\Sigma^{-1}(\theta)) - \log|S| - (p + q)$$

The ML estimator assumes that the variables in the model are multivariate normal (i.e., the joint distribution of the variables is distributed normally).

### GLS

Generalized least squares is an alternative fitting function. The GLS fit function also minimizes the discrepancy between  $S$  and  $\Sigma$ , but uses a weight matrix for the residuals, designated  $W$ .

$$F_{GLS} = \left(\frac{1}{2}\right) tr\left(\left\{\left[S - \Sigma(\theta)W^{-1}\right]\right\}^2\right)$$

Notice that this is a much simpler function (e.g., no logs), and it is clear that the discrepancy between the obtained covariance matrix and the covariance matrix implied by the model ( $S - \Sigma$ ) is minimized after weighting it by  $W$ . Although any  $W$  can be chosen for the weight matrix, most commonly, the inverse of the covariance matrix,  $S$ , is used in SEM packages.  $F_{GLS}$  is asymptotically equivalent to  $F_{ML}$ , meaning that as sample sizes increase, they are approximately equal.  $F_{GLS}$  is based on the same assumptions as  $F_{ML}$  and would be used under the same conditions. It is thought to perform less well, however, in small samples, so  $F_{ML}$  is usually chosen instead of  $F_{GLS}$ . The simplicity of the function, however, means that other weight matrices could be used in an attempt to correct for violations of distributional assumptions.

### ADF

The asymptotic distribution free function was developed by Browne (1984). It is described as arbitrary generalized least squares (AGLS) by Bentler in the EQS package and weighted least squares (WLS) by Joreskog and Sorbom in Lisrel. The main advantage of the ADF estimator is that it does not require multivariate normality. The ADF estimator is based on the  $F_{GLS}$ , except a different  $W$  is chosen. It can be written in a general form that encompasses GLS, ML, and ULS (not discussed here) where the difference depends on the choice of  $W$ :

$$F_{ADF} = F_{AGLS} = F_{WLS} = (s - \sigma)' W^{-1} (s - \sigma)$$

$W$  used in  $F_{ADF}$  is based on a covariance of all of the elements of the covariance matrix,  $S$ . That is, a covariance matrix is constructed that estimates the covariances between each  $s_{ij}$  element of  $S$ , and is therefore a  $\frac{1}{2}[v(v+1)]$  by  $\frac{1}{2}[v(v+1)]$  matrix. The reason for this is that these "covariances of covariances" are related to kurtosis estimates (so called "fourth-order moments"). So, the GLS fit function is weighted by variances and kurtosis in attempt to correct for violations of the normality assumption. Another way of saying this is that when the data are normal, the ADF estimator reduces to GLS because there is no kurtosis. The large weight matrix causes serious practical difficulties when there is a large number of variables in the model (e.g., more than 20 or so), and computer packages (e.g., EQS) do not allow estimation unless the number of cases is equal or greater than number of elements in the weight matrix (i.e.,  $\frac{1}{2}[v(v+1)]$  times  $\frac{1}{2}[v(v+1)]$  divided by 2). Simulation studies suggest that chi-square values are severely overestimated with small samples and that sample sizes of about 5000 are necessary for good estimates. A recent study by Olsson, Foss, Troye, and Howell (2000) suggests that ADF estimation performs poorly when the model is misspecified. Combined with the limitation of variables, this is usually seen as an unattractive approach when nonnormality exists.

### Muthen's CVM

Muthen (1993) suggested a categorical variable model (CVM) for use when models measured variables are categorical (either dichotomous or ordered categorical). Models with categorical variables are always considered to be in violation of the normality assumption and, thus, the usual  $F_{ML}$  estimator is not recommended. The CVM approach uses the general ADF function (which Muthen and Lisrel call WLS), but

does not have a practical limit on the number of parameters nor require such large samples, because it avoids inversion of the large weight matrix (using something called “Taylor expansion”). The idea behind the method is that categorical variables have an underlying continuous latent variable, called  $y^*$ .  $y^*$  is estimated by *polychoric* correlations which correct for loss of information in Pearson correlations due to categorization of a continuous variable (See MacCallum, Zhang, Preacher, & Rucker, 2002). *Tetrachoric* correlations are a special case of polychoric correlations involving only binary variables, and *polyserial* correlations are those involving the correlation between a binary and a continuous variable. The polychoric correlations are then used to estimate the model using the  $F_{WLS}$  estimator. Mplus has special features that implement the CVM approach. A similar approach is available in Lisrel by creating a polychoric correlation matrix in Prelis (the Lisrel preprocessor) and then analyzing the new matrix in Lisrel with WLS (not ML). The CVM approach with Mplus is a simpler process and is able to avoid inversion of the large  $W$  matrix.

## Bootstrapping

Bootstrap is another approach to problems with nonnormality (but is not typically recommended for dichotomous variables). In the bootstrap a large number of samples (usually 500 or 1000 are recommended) are drawn from your data. The samples are drawn *with replacement*, so that the same cases may be drawn into the same bootstrap sample. These repeated samples create a mini sampling distribution, and based on the central limit theorem, it should have desirable distributional characteristics. There are a number of variations on bootstrapping with SEM, including “naïve” bootstrap, bias correction, and bias corrected accelerated, which I will not discuss here (but see Bollen & Stine, 1993; Yung & Bentler, 1996). Usually, the bootstrap samples are used to calculate new standard errors and can be used to correct the chi-square for fit. The z-tests or “critical ratio” uses the bootstrap standard errors, and are considered “approximate” significance tests. Bootstrapping is a relatively new approach in SEM and has not undergone thorough testing involving a wide range of sample sizes, model complexities, or correct vs. incorrect models. Recent evidence by Nevitt and Hancock (2001) suggest sample size of 500 or greater may be needed.

## The Satorra-Bentler Scaled Chi-square and Standard Errors

Satorra and Bentler developed a process whereby multivariate kurtosis estimates are used to “scale” or correct the chi-square value and standard errors. Chi-square is usually inflated with nonnormal samples and standard errors are usually too small. This approach is used for continuous nonnormal variables and appears to do fairly well with small samples (200-500 cases; Curran, West, & Finch, 1996).

## Other Estimators

Other possible estimators include two-stage least squares (2SLS), three-stage least squares (3SLS), ordinary least squares (OLS), diagonally weighted least squares (DWLS), and unweighted least squares (ULS). Most of these approaches are seldom used, because they provide poor estimation (e.g., ULS) or because they have not been very thoroughly investigated (e.g., DWLS). 2SLS has received more attention lately in statistical papers because it does not rely on normality assumptions and is one approach to moderator tests in SEM (more on this topic later; Bollen, 2002; Bollen & Paxton, 1998)

## References and Further Reading

- Bollen, K.A. (2001). Two-Stage Least Squares and Latent Variable Models: Simultaneous Estimation and Robustness to Misspecifications. Ch 7, pp 199-138 in R. Cudeck, S. du Toit and D. Sorbom (eds) *Structural Equation Modeling: Present and Future: A Festschrift in honor of Karl Joreskog*, Scientific Software International: Lincolnwood.
- Bollen, K. A., & Paxton, P. (1998). Two-stage least squares estimation of interaction effects. In R. E. S. G. A. Marcoulides (Eds.), *Interaction and nonlinear effects in structural equation modeling* (pp. 125-151). Mahwah, NJ, Lawrence Erlbaum Associates.
- Bollen, K.A., & Stine, R.A. (1993). Bootstrapping goodness-of-fit measures in structural equation modeling. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models*. Newbury Park: Sage.
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1, 16-29.
- Finney, S.J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G.R. Hancock & R.O. Mueller (Eds.), *Structural equation modeling: A second course*. Greenwich, CT: Information Age Publishing.
- MacCallum, R.C., Zhang, S., Preacher, K.J., & Rucker, D.D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7, 19-40.
- Muthen, B.O. (1993). Goodness of fit with categorical and other nonnormal variables. In K.A. Bollen, & J.S. Long (eds.), *Testing structural equation Models* (pp. 205-234). Newbury Park, CA: Sage.
- Nevitt, J., & Hancock, G. R. (2001). Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling. *Structural Equation Modeling*, 8, 353-377.
- Olsson, U.H., Foss, T., Troye, S. V., & Roy D. Howell (2000). The Performance of ML, GLS and WLS Estimation in Structural Equation Modeling Under Conditions of Misspecification and Nonnormality. *Structural Equation Modeling*, 7 (4), 557-595.
- Yung, Y.-F., & Bentler, P.M. (1996). Bootstrapping techniques in analysis of mean and covariance structures. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced Structural Equation Modeling. Issues and Techniques* (pp.125-157). Mahwah, NJ: Lawrence Erlbaum.