Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). The Guilford Press.

Chapter 1 – Coming of Age

- recommendations to get ready to learn about SEM: (1) know your area (theoretically and empirically); (2) know your measures (their psychometric properties); (3) understand regression techniques, tests of statistical significance, and data screening and measure selection; (4) use your brain; (5) get a computer tools for SEM; (6) join the SEMNET community;

Definition of SEM

- synonymous: covariance structure analysis, covariance structure modeling, analysis of covariance structure;

- SEM is a causal inference method that takes three inputs (qualitative causal hypotheses, questions about causal, relations, and data) and generates three outputs (parameter estimates, logical implications of the model, the degree to which the testable implications of the model are supported by the data);

- a priori, but not exclusively confirmatory (e.g., testing of alternative models);

- model generation: seeks to discover a model that makes theoretical sense, it is reasonably parsimonious, and it has acceptably close correspondence to the data;

- models in SEM assume probabilistic causality;

- all latent variables in SEM are continuous;

- an observed variable used as an indirect measure of a construct is an *indicator*, and the statistical realization of a construct based on analyzing scores from its indicators is a *factor*;

- residual or error terms: another latent variable category;

- path analysis: SEM without latent variables;

- the basic datum of SEM is the covariance:

- it represents the strength of the linear association between and and their variabilities;

- SEM analysis has two goals: (1) to understand patterns of covariances among a set of observed variables and (2) to explain as much of their variance as possible with the researcher’s model;

- sample size considerations;

- *N*:*q* rule: ratio 20:1, where *N* = number of cases and *q* = number of parameters;

- SEM involves a shift from significance testing about individual effects to the evaluation of whole statistical models;

- the whole of the GLM can be seen as just a restricted case of SEM for analyzing observed variables;

- regression (statistical prediction) versus SEM (causal modeling);

- MacCallum and Austin (2000): reporting flaws in SEM research;

Chapter 2 – Regression Fundamentals

Bivariate Regression

- predicting Y from —also called regressing on —takes the form:

- the predicted scores above defined make up a composite, or a weighted linear combination of the predictor and the intercept. Parameters are estimated using the ordinary least squares (OLS) method, which seeks to minimize the sum of squared residuals, or ;

- coefficient is related to the Pearson correlation and the standard deviations of and :

- thus, corresponds to the *covariance* structure of the linear regression model;

- the intercept is related to both and the means of both variables:

- the term reprsents the *mean* structure of the linear regression model;

- regression residuals, or , sum to zero and are uncorrelated with the predictor, or:

- this is required in order for the computer to calculate unique values of the regression coefficient and intercept in a particular sample (i.e., independence of residuals);

- the equation for regressing on when both variables are standardized (i.e., their scores are normal deviates, ) is:

- thus, Pearson correlation is the standardized regression coefficient predicting -score of from -score on . This equation there is not intercept because the means of standardized variables equal zero (variance of standardized variables are 1.0);

- if , this means that a score 1 *SD* above the mean on predicts a score almost seven-tenths of 1 *SD* above the mean on ;

- there is a special relation between and the unstandardized predicted scores. If is regressed on , for example, then: (1) , that is, the bivariate correlation between and equals the bivariate correlation between and ; (2) the observed variance in can be represented as the exact sum of the variances of the predicted scores and the residuals, ; and (3) , which says that the squared correlation between and equals the ratio of the variance of the predicted scores over the variance of the observed socres on ;

- when replication data are available, it is actually better to compare unstandardized regression coefficients, such as , especially if those samples have different variances on or ;

Multiple Regression

- the form of the unstandardized equation for regressing on both and is:

- where and are the *unstandardized partial regression coefficients*, which estimates the change in , given a 1-point change in (or in ), while controlling for (or for ). The intercept equals the predicted score on when the scores on *both* predictors are zero, or ;

- the statistic () equals the proportion of variance explained in by both predictors, controlling for their intercorrelation;

- equations for the unstandardized partial regression coefficients for each of two continuous predictors are:

and

- where and for and are, respectively, their *standardized partial regression coefficients*, also known as *beta weights*. Their formulas are listed next:

and

- in the numerators, the bivariate correlation of each predictor with the criterion is adjusted for the correlation of the other predictor with the criterion and for correlation between the two predictors. The denomitators adjust the total standardized variance by removing the proportion shared by the two predictors;

- the intercept is equals to:

- and the regression equation for standardized variable is:

- because all variables have the same metric in the standardized solution, we can directly compare values of with in a same sample (but unstandardized coefficients are preferred for comparing results for the same predictor over different samples);

Corrections for Bias

- Wheery correction of the estimate of ;

- assumptions of regression: (1) regression coefficients reflect unconditional linear relations only; (2) all predictors are perfectly reliable (no measurement error; criterion variable need not be measured without error, although an instrument with poor psychometric properties may can reduce the value of ); (3) significance tests in regression assume that the residuals are normally distributed and homoscedastic; (4) there are no causal effects among the predictors (i.e., there is a single equation); (5) there is no specification error;

- overestimation due to omoission of a predictor may occur more often than underestimation (suppression);

Supression

- suppression occurs when either (1) the absolute value of a predictor’s beta weight is greater than that of its bivariate correlation with the criterion or (2) thw two have different signs. So defined, suppression implies that the estimated relation between a predictor and a criterion while controlling for other predictors is a “surprise,” given the bivariate correlations. There are negative suppression, classical suppression, and reciprocal supression;

- entry methods in multiple regression include sequential and simultaneous methods. Stepwise methods capitalize on chance. The findings are unlikely to replicate;

- once a final set of rationally selected predictors has been entered into the equation, they should *not* be subsequently removed if their regression coefficients are not statistically significant. […] If you had good reason for including a predictor, then it is better to leave it in the equation until replication indicates that the predictor does not appreciably relate to the criterion;

Part and Partial Correlations

- the concept of partial correlation concerns the idea of *spuriousness*: If the observed relation between two variables is wholly due to one or more common cause(s), their association is spurious. The *first-order partial correlation* removes the influence of a third variable from both and . The formula is:

- the *second-order partial correlation*, estimates the association between and controlling for both and ;

- *part correlation* or *semipartial correlation* that controls for external variables out of either of two other variables, but not both (e.g., it measures the association between and controlling for the association between and , but not for the association between and ). The formula for *first-order part correlation* is:

Logistic Regression and Probit Regression

- the prediction equation in logistic regression is a *logistic function*, or a sigmoid function with an “S” shape. It is a type of *link function*, or a transformation that relates the observed outcomes to the predicted outcomes in a regression analysis;

- in standard regression, the link function is the *identity link*, which says that observed scores on the criterion are in the same units as , the predicted scores;

Chapter 3 – Significance Testing and Bootstrapping

Standard Errors

- the *SE* is a *SD* in a *sampling distribution*, the probability distribution for a sample statistic based on all possible random samples selected from the same population and each based on the same *N*. The *SE* estimates the *sampling error*, or the difference between sample statistics and the corresponding population parameter. The is estimated by:

- assumptions of the *SEM*: (1) the method of sampling is random; (2) there is no source of error besides sampling error; (3) *SE*s for parametric statistics often assume normality (to population distributions) or homoscedasticity;

Critical Ratios

- the basic form of a significance test is the *critical ratio*, the ratio of a statistic over its *SE*. Assuming large samples and normality, a critical ratio is interpreted as a deviate in a normal curve () with a mean of zero and a *SD* that equals the *SE*;

- the *p* value for an unstandardized estimate does not automatically apply to its standardized counterpart;

- in small samples, the ratio, approximates a distribution, which necessitates the use of special tables to determine the critical values of for the .05 or .01 levels;

Power and Types of Null Hypotheses

- the failure to reject a null hypothesis such as when testing at the .05 level, is meaningful only if (1) the power of the test is adequate and (2) the null hypothesis is at least plausible to some degree;

- power is affected by sample size, the alpha level and the directionality of the *H*1, design type (between vs. within), the test statistic used, and the reliability of the scores;

- the type of null hypothesis tested most often is a *nil hypothesis*, which says that the value of a parameter or the difference between two or more parameters is zero;

- nil hypotheses may be appropriate in new research areas where it is unkown whether effects exist at all, but such hypotheses are less suitable in more established areas where it is already known that certain effects are not zero;

Significance Testing Controversy

- some authors in *statistics reform* suggest that overreliance on significance testing can led to *trained incapacity*, or the inability of researchers to understand their own results due to inherent limitations of significance tests and myriad associated cognitive distortions;

- criticisms to significance testing: (1) *p* values are wrong in most studies; (2) researchers do not understand *p* values; (3) most applications of significance testing are incorrect; (4) significance tests do not tell researchers what they want to know;

- “Big Five” misinterpretations of *p* values: (1) the *odds against chance fallacy*: is the false belief that *p* indicates the probability that a particular result happened by chance (i.e., due to sampling error); (2) the *local Type I error fallacy*: says that the likelihood that the decision just taken to reject the null hypothesis ia s Type I error is less than 5%; (3) the *inverse probability fallacy* is the false belief that *p* is the probability that the null hypothesis is true; (4) the *valid research hypothesis fallacy* is the false belief that 1 – *p* is the probability that the alternative hypothesis is true; (5) the *replicability fallacy* is that 1 – *p* is the likelihood of finding the same result in another random sample;

Confidence Intervals and Noncentral Test Distributions

- interval estimation is an alternative; the confidence interval is:

- the interval specifies a range of values considered equivalent to the observed mean within the limits of sampling error at the 95% CI;

- null hypotheses are required for significance tests, but not for CIs, and many null hypotheses have little scientific value;

- a *noncentrality parameter* indicates the degree to which the null hypothesis is false;

Bootstrapping

- it is a computer-based method of *resampling* that combines the cases in a data set in different ways to estimate statistical precision. *Nonparametric* *bootstrapping* treats your sample (i.e., data file) as a pseudo-population in that cases are randomly selected *with replacement*  to generate other data sets, usually of the same size as the original. It enables de construction of an *empirical sampling distribution*. Nonparametric bootstrapped confidence intervals are calculated in the empirical distribution;

Chapter 4 – Data Preparation and Psychometrics Review

Forms of Input Data

- most estimation methods in SEM, including default maximum likelihood estimation, assume that the variables are unstandardized. Analyses using correlation matrix may derive incorrect *SE*s for standardized estimates if special methods for standardized variables are not used;

Positive Definiteness

- the data matrix that you submit should be *positive definite*. A positive definite data matrix has the properties summarized next: (1) the matrix is *nonsingular* or has an inverse. A matrix with no inverse is *singular*; (2) all eigenvalues of the matrix are positive (> 0), which also says that the matrix determinant is positive; and (3) there are no out-of-bounds correlations or covariances;

- in most kinds of multivariate analyses (SEM included), the computer needs to derive the inverse of the data matrix as part of its linear algebra operations. An *eigenvalue* is the variance of an *eigenvector*, and both are from a principal components analysis of the data matrix, or *eigendecomposition*, that creates a total of *v* orthogonal linear combinations, or eigenvectors, of the observed variables, where *v* is the total number of those variables. The maximum number of eigenvectors for a data matrix equals *v*, and the set of all possible eigenvectors explains all the variance of the original variables;

- perfect colllinearity means that some denominators in matrix calculations will be zero, which results in illegal (undefined) fractions (estimation fails);

- negative eigenvalues (< 0) may indicate a data matrix element—a correlation or covariance—that is *out of bounds*;

- equation 4.1 explain that the value of the Pearson correlation between two variables *X* and *Y* is limited by the correlations between these variables and a third variable *W*;

- in a positive definite data matrix, the maximum absolute value of covariance between *X* and *Y* must be less than or equal to the square root of the product of their variances;

- the *determinant* of the data matrix is the serial product (the first times the second times the third, and so on) of the eigenvalues. Assuming that all eigenvalues are positive, the determinant is a kind of matrix variance;

Extreme Collinearity

- researchers can inadvertenly cause extreme collinearity when composites and their constituents are analyzed together;

- methods to detect collinearity: (1) suggests extreme multivariate collinearity; (2) *tolerance*, defined as , indicates the proportion of total standardized variance that is unique. Tolerance values may indicate extreme multivariate collinearity; (3) *variance inflation factor* (VIF), or , is the ratio of the total standardized variance over the proportion of unique variance (tolerance). The variable in question may be redundant if ;

Outliers

- xxx

Normality

- maximum likelihood estimation assumes multivariate normality;

Transformations

- the effect of applying normalizing transformations is to compress one part of a distribution more than another, thereby changing its shape but not the rank of the scores (this describes a *monotonic transformation*);

Chapter 7 – Identification of Observed Variable (Path) Models

General Requirements

- two general requirementos for identifying any type of model in SEM: (1) the model *df*s must be at least zero (the counting rule); (2) every latent variable—including disturbances or error terms—must be assigned a scale;

- most structural equation models with zero *df* that are also identified can perfectly reproduce the data (sample covariances), but such models test no particular hypothesis (*just-identifed model*);

- an *underidentified structural equation model* is one for which it is not possible to uniquely estimate all of its free parameters;

- deleted paths reflect strng causal assumptions and thus are more interesting than the specified (freely estimated) paths, which reflect weak hypotheses;

Chapter 9 – Specification and Identification of Confirmatory Factor Analysis Models

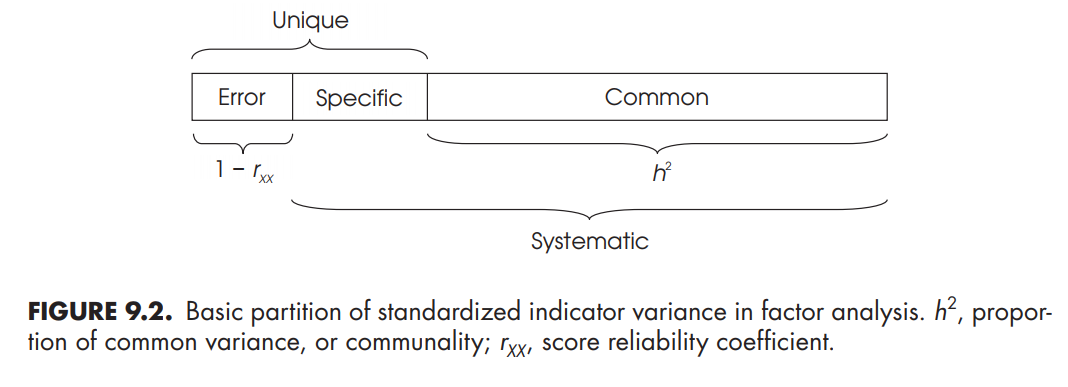
Latent Variables in CFA

- unrestricted measurement models are analyzed in EFA, but CFA deals with restricted measurement models;

- a latent variable is an explanatory variable that corresponds to the *local independence* assumption that (1) one or more latent variables create the association between observed variables, and (2) when the latent variables are held constant, the indicators are independent, if their error variances are independent of both one another and of the latent variables as well;

Factor Analysis

- all factor-analytic methods partition standardized indicator variance in the way shown in Figure 9.2. *Common variance* is shared among the indicators and is a basis for observed covariances among them that depart appreciably from zero. The proportion of total variance that is shared is called *communality*, *h*². The rest is *unique variance*, which consists of specific variance and random measurement error. *Specific variance* is systematic variance that is not explained by any factor in the model;



Characteristics of EFA Models

- linearity is assumed plus a method of factor extraction, such as the principal-axis method, that replaces the 1.0s in the main diagonal of the sample correlation matrix with estimated communalities;

- Kline calls factor loadings as *pattern coefficients*;

Characteristics of CFA Models

- each indicator in a standard CFA model has two unrelated causes, the factor and error. This specification is consistent with the view in classical reliability theory that observed scores are determined by a true (systematic) component and an error component. This also describes *reflective measurement*, where latent variables are assumed to cause observed variables;