some time to conclude—for example, a divorce, serious illness, or going on welfare. Near the middle of the spectrum is the category of daily hassles. While hassles are not major stressors individually, their accumulation from day to day may represent an important stress source. More continuous in nature is the ongoing absence of an expected or desired social role, or nonevent. Stressors of this type include joblessness and childlessness. Chronic stress is the most continuous type; examples include living in a dangerous neighborhood, poverty, and living with a disability.

Eventful and chronic stress may be related through a process of stress proliferation. One example is when a worker loses a job because macroeconomic conditions led to the closure of a plant. Soon the loss of the individual's worker role and his or her source of income precipitate a financial crisis and increased conflict in the marital and parent roles—the "event" of job loss has proliferated stressful experience in a whole constellation of life domains. Social roles, and hence role-related stressors, do not occur in isolation.

Sometimes, the stressful meaning of a normally undesirable life event is negated by the context within which it occurs. Consider the separation or divorce of a person whose marital role history had been fraught with disappointment, conflict, and unhappiness—the event in such a case does not demand the kind of adaptation that is a threat to the person's well-being.

Stress as a Social Process

Stress may be viewed as the central means by which the structural arrangements of society create differential health outcomes for the people who occupy different social statuses and roles. Stress theory does not treat stress exposure as a health determinant in isolation: Stress is one element within a process that is closely linked to the social system. The amount of stress experienced is largely determined by an individual's social location. So are the social and personal resources that are available to forestall or cope with stressful events and circumstances as they occur. Stressful experiences may motivate social support (if it is available), which can mitigate their deleterious consequences. Successful resolution of a stressful event, such the loss of a home in a natural disaster, can build confidence that one can cope with future losses. Or it could be devastating to a person who had limited access to coping resources in the first place. Social inequality in the exposure to stress and in the availability of protective factors amplifies the production and reproduction of health disparities. That is because social structural arrangements are systematically related both to the amount a person is exposed to stress and access to the resources needed to mitigate its ill-health effects.

Stress arises from the social context of people's lives. There is systematic variation in the level of stress and coping resources across social status dimensions. Greater exposure to stress is associated with low education and poverty, unmarried status, minority group membership, and youth. The existence of a gender difference is less clear. Since stress and coping resources are important determinants of health and illness outcomes, stress functions as an epidemiological link between a society's structure and the health outcomes of its members.

—Donald A. Lloyd

See also Geographical and Social Influences on Health; Health Disparities; Social Capital and Health; Social Epidemiology; Social Hierarchy and Health

Further Readings

Aneshensel, C. S. (1992). Social stress: Theory and research. *Annual Review of Sociology*, *18*, 15–38.

Cohen, S., Kessler, R. C., & Gordon, L. U. (1995). *Measuring stress: A guide for health and social scientists.* New York: Oxford University Press.

Dohrenwend, B. P. (1998). *Adversity, stress, and psychopathology.* New York: Oxford University Press.

Pearlin, L. I., Lieberman, M., Menaghan, E., & Mullan, J. T. (1981). The stress process. *Journal of Health and Social Behavior*, *22*, 337–356.

Thoits, P. A. (1995). Stress, coping, and social support processes: Where are we? What next? [Extra issue]. *Journal of Health and Social Behavior*, 53–79.

Turner, R. J., Wheaton, B., & Lloyd, D. A. (1995). The epidemiology of social stress. *American Sociological Review*, *60*, 104–125.

Structural Equation Modeling

The roots of structural equation modeling (SEM) begin with the invention of least squares about 200 years ago, the invention of factor analysis about 100 years ago, the invention of path analysis about 75 years ago, and the invention of simultaneous equation models about 50 years ago. The primary focus with SEM is on testing causal processes inherent in our theories. Before SEM, measurement error was assessed separately and not explicitly included in tests of theory. This separation has been one of the primary obstacles to advancing theory. With SEM, measurement error is estimated and theoretical parameters are adjusted accordingly—that is, it is subtracted from parameter estimates. Thus, SEM is a fundamental advancement in theory construction because it integrates measurement with substantive theory. It is a general statistical methodology, extending correlation, regression, factor analysis, and path analysis.

SEM is sometimes referred to as "latent variable modeling" because it reconstructs relationships between observed variables to infer latent variables. Many variables in epidemiological research are observable and can be measured directly (e.g., weight, pathogens, mortality). However, many variables are also inherently unobservable or *latent*, such as well-being, health, socioeconomic status, addiction, and quality of life. Measuring and interpreting latent variables requires a measurement theory. Latent variables and its respective measurement theory can be tested using an SEM technique called "confirmatory factor analysis." This involves specifying which latent variables are affected by which observed variables and which latent variables are correlated with each other.

SEM also provides a way of systematically examining reliability and validity. Reliability is the consistency of measurement and represents the part of a measure that is free from random error. In SEM, reliability is assessed as the magnitude of the direct relations that all variables except random ones have on an observed variable. This capability of SEM to assess the reliability of each observed variable and simultaneously estimate theoretical and measurement parameters is a fundamental methodological advancement. The potential for distortion in theoretical parameters is high when measurement error is ignored, and the more complicated the model the more important it becomes to take measurement error into account. Validity is the degree of direct structural relations (invariant) between latent and measured variables. SEM offers several ways of assessing validity. Validity differs from reliability because we can have consistent invalid measures. The R^2 value of an observed variable offers a straightforward measure of reliability. This R^2 sets an upper limit for validity because the validity of a measure cannot exceed its reliability.

Major Assumptions

Like other kinds of analyses, SEM is based on a number of assumptions. For example, it assumes that data represent a population. Unlike traditional methods, however, SEM tests models by comparing sample data with the implied population parameters. This is particularly important because the distinction between sample and population parameters has been often ignored in practice. SEM generally assumes that variables are measured at the interval or ratio level, and ordinal variables, if used at all, are truncated versions of interval or ratio variables. Hypothesized relationships are assumed to be linear in their parameters. All variables in a model are assumed to have a multivariate Gaussian or normal distribution. Therefore, careful data screening and cleaning are essential to successfully work with SEM.

SEM shares many assumptions with ordinary least squares regression and factor analysis. For example, the error of endogenous latent variables is uncorrelated with exogenous variables. The error of the endogenous observed variables is uncorrelated with the latent endogenous variables. The error of the exogenous observed variables is uncorrelated with the latent exogenous variables. The error terms of the endogenous latent variables and the observed endogenous and exogenous variables are mutually uncorrelated. This is the result of combining factor analysis and regression in one overall simultaneous estimation.

Steps in SEM

Specification

Models are constructed by defining concepts, clarifying the dimensions of each concept, forming measures of the dimensions, and specifying the expected empirical relationships between the measures and the construct. The accuracy of parameter estimates is partly dependent on the correctness of the theory and partly dependent on the validity of the measurement. There is always more than one model that fits the data, and thinking about these alternative models and testing them helps refine theory. Depicted in Figure 1 is a path diagram—a common way to represent models. The circles represent latent variables, squares represent observed variables, double-headed arrows represent correlations, and single-headed arrows represent causal effects. The one-to-one correspondence between path diagrams and sets of structural equations facilitates communication and clarification of all

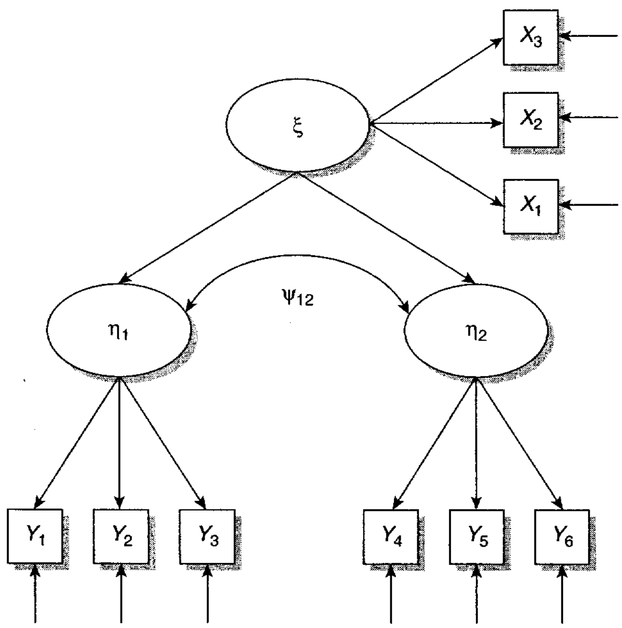

Figure 1 Example of a Path Diagram in SEM

parameters and their interrelationships. Model parameters are fully specified, which means stating a hypothesis for every parameter.

Identification

Models are composed of a set of equations with known and unknown parameters. Identification is the problem of determining whether there is a unique solution for each unknown parameter in the model. It is a mathematical problem involving population parameters, not sample size. A model can fail to be identified even with a large sample. There are a number of rules that if followed ensure identification. The most common is the *t* rule. The *t* in the *t* rule refers to the number of free parameters specified in the model. Specifically, a model is identified if the *t* value is equal to or smaller than half the number of observed variables multiplied by the number of observed variables plus 1 [$t \leq (1/2)(p)(p+1)$]. The *t* rule is a necessary but not sufficient condition for identification. Other rules are the scaling rule, three-indicator rule, null-β rule, recursive rule, and rank-and-order rules.

Estimation

SEM estimation procedures use a particular fitting function to minimize the difference between the population and the sample. Basically, this is a recipe to transform data into an estimate. The data matrix for SEM must be positive definite, a mathematical requirement for the estimation algorithms. Maximum likelihood is the default estimator in most SEM programs. Maximum likelihood is based on the idea that the sample is more likely to have come from a population of one particular set of parameter values than from a population of any other set of values. Maximum likelihood estimation is the vector of values that creates the greatest probability of having obtained the sample in question. This method of estimation is asymptotically unbiased, consistent, and asymptotically efficient, and its distribution asymptotically normal. If the sample is large, no other estimator has a smaller variance. There are two drawbacks with maximum likelihood. First, it assumes a normal distribution of error terms, which is problematic for many measures in the health and social sciences fields. Second, the assumption of multinormality is even more problematic, again because of the extensive use of crude measures.

In choosing estimators, the main choice is between maximum likelihood and weighted least squares. The weighted least squares estimator is used when multivariate normality is lacking and, especially, when some of the variables are ordinal. Although weighted least squares is computationally demanding, it is important to have a large sample size when some variables are ordinal. Other choices in estimators include generalized least squares and unweighted least squares. Maximum likelihood and generalized least squares are very similar. The generalized least squares estimator weights observations to correct for unequal variances or nonzero covariance of the disturbance terms. It is used when variable distributions are heteroscedastic or when there are autocorrelated error terms. An unweighted least square is used with variables that have low reliability. This estimator is less sensitive to measurement error than maximum likelihood or generalized least squares. Research shows estimates from the unweighted least square to be similar in models with and without error, while maximum likelihood estimates without and without errors are very different.

Fitting

After a model is estimated, its fit must be assessed. There are more than 20 different fit measures to assess misfit and goodness of fit. They are based on six different criteria: (1) the discrepancy between the sample covariance matrix and the fitted (population) covariance, (2) accounting for observed variances and covariance, (3) maximizing the fit of a cross-validated model, (4) including a penalty for unnecessarily estimating parameters or creating fewer degrees of freedom, (5) the amount of improvement over a baseline model, and (6) separating the measurement model from the latent variable model.

Most of the existing fit measures are tied directly or indirectly to the chi-square ratio. This chi-square statistic is based on the same general idea as the familiar chi-square comparison between the observed and expected values. The difference is that in SEM our substantive interest or hypothesis is the null hypothesis. In traditional applications, we want to reject the hypothesis of no difference between observed and expected frequencies so that we can accept the alternative hypothesis of a difference. In contrast, with SEM, we want to find no difference between the expected and observed values. Therefore, the smaller the chi-square values the better because this leads to a failure to reject the null hypothesis, which is our substantive interest.

The chi-square statistic assumes that the variables in a model are multivariate normal, that the data are unstandardized (covariance as opposed to correlations matrixes), that sample sizes are at least $N > 100$ and preferably $N > 200$, and that the model holds exactly in the population. This chi-square has been found robust to skew violations but sensitive to Kurtosis violations. The interpretation of chi-square depends on adequate sample sizes. With large samples tiny deviations can lead one to reject the null hypothesis, which again in SEM is of substantive interest.

SEM fit measures are not applicable in exactly identified models. In exactly identified models (when degrees of freedom $= 0$), the sample variances and covariance always equal the estimates of the population variances and covariance because there is only enough information to calculate one estimate per parameter. A limitation of chi-square is that the closer the model is to being exactly identified, the higher the chi-square value. In other words, chi-square values always decrease when parameters are added to the model. With an overidentified model (degrees of freedom > 0), the overall fit can differ from the fit of different parts of the model. A poor overall fit does not help to detect areas of poor fit. The overall fit statistics also do not tell us how well the independent variables predict the dependent variables.

A good fit does not mean that model is "correct" or "best." Many models fit the same data well. Measurement parameters often outnumber theoretical parameters. Therefore, a "good fit" may reflect the measurement and not the theory. There is considerable discussion about fit measures. The best current advice in evaluating model fit is to seek a nonsignificant chi-square (at least $p > .05$ and preferably .10, .20, or better); an IFI (incremental fit index), RFI (relative fit index), or CFI (comparative fit index) greater than .90; low RMSR (root mean square residual) and RMSEA (root mean square of approximation) values, plus a 90% confidence interval for RMSEA $< .08$; and a parsimony index that show the proposed model as more parsimonious than alternative models.

Modification

It is not uncommon for models to exhibit a poor fit with the data. There are many potential sources of error, including an improperly specified theory, poor correspondence between the theory and the model, and causal heterogeneity in the sample. Modifications are typically made to poor-fitting models, and most SEM software packages provide modification indices that suggest which changes can improve model fit. However, using these indices in the absence of theory represents one of the main abuses of SEM. It is important that a systematic search for error is conducted and that modifications are based on theory or to generate new theory. A well-fitting respecified model does not represent a test. Respecified models must be tested on new data.

Specialized Techniques

SEM is a highly flexible methodology that allows for many special types of models to be examined. The most common models are those with unidirectional (recursive) causal effects, but SEM also allows for bidirectional (nonrecursive) effects to be tested. Stacked or multiple groups can also be examined, which facilitates interpretation and tests of interaction. Repeated measures designs can be analyzed using an SEM technique called "latent growth curves." This provides a way of examining both linear and nonlinear changes over time. Recent advances in software also provide a way of accounting for hierarchical or nested data structures, including survey weights.

Summary

SEM is a flexible and extensive method for testing theory. These models are best developed on the basis of substantive theory. Hypothesized theoretical relationships imply particular patterns of covariance or correlation. Statistical estimates of the hypothesized covariance indicate, within a margin of error, how well the models fit with data. The development and testing of these models advance theory by allowing latent variables, by including measurement error, by accepting multiple indicators, by accommodating reciprocal causation, and by estimating model parameters simultaneously. Structural equation models subsume factor analysis, multiple regression, and path analysis. The integration of these traditional types of analysis is an important advancement because it makes possible empirical specification of the linkages between imperfectly measured variables and theoretical constructs of interest.

The capabilities, technical features, and applications of SEM are continually expanding. Many of these advances are reported in the journal *Structural Equation Modeling* and communicated on the international and interdisciplinary SEM listserv called SEMNET. This listserv also archives its discussion and provides a forum for offering and receiving advice, which makes it an invaluable resource for epidemiologists and other social scientists learning and using SEM.

—*David F. Gillespie and Brian Perron*

See also Factor Analysis; Measurement; Regression

Further Readings

De Stavola, B. L., Nitsch, D., Silva, I. D. S., McCormack, V., Hardy, R., Mann, V., et al. (2005). Statistical issues in life course epidemiology. *American Journal of Epidemiology*, *163*, 84–96.

Der, G. (2002). Structural equation modelling in epidemiology: Some problems and prospects [Commentary]. *International Journal of Epidemiology*, *31*, 1199–1200.

Hoyle, R. H. (Ed.). (1995). *Structural equation modeling.* Thousand Oaks, CA: Sage.

Keller, S. D., Ware, J. E., Jr., Bentler, P. M., Aaronson, N. K., Alonso, J., Apolone, G., et al. (1998). Use of structural equation modeling to test the construct validity of the SF-36 Health Survey in ten countries: Results from the IQOLA project. *Journal of Clinical Epidemiology*, *51*, 1179–1188.

Kline, R. B. (2004). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford Press.

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Singh-Manous, A., Clarke, P., & Marmot, M. (2002). Multiple measures of socio-economic position and psychosocial health: proximal and distal measures. *International Journal* of Epidemiology, *31*, 1192–1199.

Study Design

Epidemiologic studies have traditionally been categorized as having "descriptive" or "analytic" designs. Descriptive studies are viewed primarily as hypothesis-generating studies and usually take advantage of routinely collected data to describe the distribution of a disease in a population in terms of the basic descriptors of person, place, and time. Analytic studies are further divided into "observational" and "experimental" study designs and are viewed as approaches suitable for testing specific hypotheses about disease etiology or the efficacy of disease prevention strategies. The main categories of observational studies are the cohort, case-control, nested case-control, case-cohort, case crossover, and cross-sectional designs. The most commonly employed experimental designs used in epidemiologic research include the classic randomized clinical trial and the quasi-experimental nonrandomized study design used to evaluate the effectiveness of population-based disease prevention approaches.

Descriptive Epidemiology

Data Sources

Descriptive epidemiologic studies are designed to determine the distribution of a disease in a population with regard to person, place, and time. The numbers of individuals in the population who are diagnosed with or die from various diseases are obtained from sources such as vital records files, disease registries, and surveys. Death certificates provide information on the underlying cause of death and provide basic sociodemographic data on the decedent such as age, gender, race/ethnicity, marital status, and place of residence at the time of death. Birth certificates are used to study the incidence of various birth outcomes such