DOCUMENT RESUME

ED 379 355 TM 022 752

AUTHOR	Newman, Isadore; And Others
TITLE	Application of the Binomial Index of Model Fit: What Research Question Is Being Investigated?
PUB DATE	Oct 94
NOTE	21p.; Paper presented at the Annual Meeting of the Midwestern Educational Research Association (Chicago, IL, October 12-15, 1994).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Goodness of Fit; Heuristics; *Path Analysis; *Research Methodology; Statistical Analysis; *Structural Equation Models
IDENTIFIERS	*Binomial Distribution

ABSTRACT

In this paper the binomial index of model fit is applied to four path-analytic or structural-equation models to demonstrate how this goodness-of-fit measure can be used and interpreted. In addition, the conclusions derived from the results of the binomial index of goodness of fit were compared to the statements presented by the authors of the reviewed articles regarding the goodness of fit of the models, which were based on traditional goodness-of-fit indexes. The goodness-of-fit statements derived from the binomial index of model fit were not always in agreement with the conclusions drawn from the traditional goodness-of-fit measures. The binomial index of model fit is heuristically different from the more traditional goodness-of-fit measures, and its use will provide additional information to a researcher regarding the degree to which the data support the theoretical model. (Contains 20 references.) (Author)

Running head: APPLICATION OF THE BINOMIAL INDEX

Application of the Binomial Index of Model Fit:

What Research Question is Being Investigated?

A Paper Presented at the Meeting of the

Mid-Western Educational Research Association,

Chicago, Illinois

October 12-15, 1994

Isadore Newman
The University of Akron

John W. Fraas
Ashland University

Timothy Norfolk
The University of Akron

Abstract

In this paper the binomial index of model fit is applied to four path analytic or structural equation models to demonstrate how this goodness-of-fit measure can be used and interpreted. In addition, the conclusions derived from the results of the binomial index of goodness of fit were compared to the statements presented by the authors of the reviewed articles regarding the goodness of fit of the models, which were based on traditional goodness-of-fit indexes. The goodness of fit statements derived from the binomial index of model fit were not always in agreement with the conclusions drawn from the traditional goodness-of-fit measures. The binomial index of model fit is heuristically different from the more traditional goodness-of-fit measures, and its use will provide additional information to a researcher regarding the degree to which the data support the theoretical model.

Introduction

Numerous goodness-of-fit measures have been developed to test model fit (Bentler & Bonett, 1980; Joreskog & Sorbom, 1986). Although measures such as the Tucker-Lewis index and Hoelter's critical N have gained in popularity, the most commonly reported goodness-of-fit measures are the likelihood-ratio chi-square statistic, the goodness-of-fit index, the adjusted goodness-of-fit index, and the Bentler-Bonett normed fit index (Schumacker, 1992).

In previous papers we have presented the concept of using the number of paths in a nomological net that was supported by the data to estimate the goodness of fit of the theoretical model (Newman, 1990; Fraas & Newman, 1994; and Newman, Fraas, & Norfolk, in press). In this paper we will refer to this goodness-of-fit approach as the binomial <u>index</u> of model fit rather than the binomial <u>test</u> of model fit, which was the terminology we used in our earlier papers. We believe that this terminology better described the function of this approach, which is to measure model fit.

This approach relates to a different research question than do the more traditionally used goodness-of-fit indexes. The binomial index of model fit value measures the proportion of all subsets of supported paths that agree with the model at least as well as the given sample. Whereas, Hair, Anderson, Tatham & Black (1992) describe the more traditional approaches of estimating the goodness of fit of the model as: "... a measure of the correspondence of the actual or observed (covariance or correlation) matrix with that predicted from the proposed model" (p. 447).

The purpose of this paper is twofold. First, this paper illustrates how the binomial index of model fit is calculated for a path analytic or structural equation model. The binomial index of model fit is calculated for each of four models that were published in various education and business journals. Second, this paper demonstrates that the use of the binomial index of model fit will, when applied to certain analytical results, lead to conclusions regarding model fit that contradict the results obtained from the traditional goodness-of-fit measures.

Calculation of the Binomial Index of Model Fit Value

As proposed by Fraas and Newman (1994), two steps are required to calculate the binomial index of model fit value for a nomological net. Since a binomial index requires that any event be classified into one of two categories (Seigel & Castellan, 1988), the first step in implementing the binomial index as a goodness-of-fit measure requires the researcher to determine the criterion that will be used to judge whether a given path in the nomological net is supported by the data. It should be noted that both "causal" and "noncausal" parameter estimates are judged. In this paper both types of parameter estimates will simply be referred to as path estimates.

There are several criteria that could be used: (a) The parameter estimate for a path exceeds an a priori effect size, (b) the parameter estimate is statistically significant, (c) the parameter estimate reflects the hypothesized sign, or (d) a combination of these criteria. It is important to note that the binomial index value may differ for a given model and data set. The value, and therefore, its interpretation, is dependent on the criteria used to determine whether the paths are supported by the data. This situation is similar to that encountered by researchers when reviewing statistical test results. That is, a researcher may obtain different results if the alpha level is set at .05 rather than .01 or if the effect size is considered along with statistical significance rather than either criterion alone. Thus, it is important for the researcher to state what criteria were used to determine whether the path estimates were supported by the data when the binomial index of model fit is reported. The binomial index of model fit value, as well as all measurement and statistical values, must be interpreted in the context of the criteria used to calculate and evaluate the value.

Once the criterion of whether a given path in the nomological net is supported by the data has been selected, the second step in the binomial index of model fit approach requires the researcher to calculate the proportion of all subsets that agree with the model at least as well as the given sample. The following formula can be used to calculate this proportion:

$$p(x) = \frac{n!}{x!(n-x)!} (.5)^{x} (.5)^{(n-x)} \qquad \text{Equation 1}$$

where:

1. $\underline{p}$ is equal to the proportion of all subsets that agree with the model at least as well as the given sample.
2. $\underline{x}$ is equal to a series of numbers ranging in value from the number of paths supported by the data to the total number of paths represented by the model, inclusive.
3. $\underline{n}$ is equal to the number of paths in the model.

It is important to note the reason why we have proposed the use of Equation 1.

The proportion of all subsets of supported paths that agree with the model at least as well as the given sample could also be calculated using the following equation:

$$Q_n(k) = \frac{1}{2^n} \sum_{i=k}^{n} \frac{n!}{i!(n-i)!} \qquad \text{Equation 2}$$

where:

1. n is equal to the number of paths in the model.
2. k is equal to the number of supported paths in the model.

In Equation 2, the value 2^n indicates the total number of subsets of significant paths that could exist for a given model. Since we want the proportion of all subsets that agree with the model at least as well as the given sample, 2^n must be the value for the denominator of the proportion of all subsets that agree with the model at least as well as the given sample. The value of .5 is required in Equation 1 if the equation is to provide the proportion of all subsets of significant paths that agree with the model at least as well as the given sample. We have presented our approach of estimating model fit through Equation 1, which is a binomial cumulative distribution value, rather than Equation 2 because researchers can readily obtain such values from a binomial cumulative distribution tables found in numerous statistical textbooks or calculated by various statistical software.

Interpreting of the Binomial Index of Model Fit Value

When interpreting the binomial index value, it is important to remember that we are proposing that the binomial index of model fit be used to indicate a good fit between the data and the nomological net when a substantial number of the paths are supported by the data. The binomial index of model fit is designed so that as the number of supported paths in the nomological net increases, the binomial index of model fit value decreases. Thus, a small binomial index of model fit value indicates a good fit between the data and the theoretical model.

This use of the binomial index of model fit should also direct the attention of the researcher to the nonsignificant paths. A reexamination of the theories on which those paths were based may be in order. If a researcher wishes to make a statement regarding whether the data do or do not support the theoretical model using the binomial index value, the researcher can compare the binomial test value to some established threshold level, such as .05. The researcher should consider this threshold level to be a value used to assist in separating noise from a signal rather than implying any "confidence level" (Wheeler, 1990). Based on our initial applications of the binomial index of model fit, meaningful results occur when the binomial index value is less than the threshold of .05. That is, when the binomial index value is less than .05, a large proportion of the paths in the nomological net is supported by the data.

This interpretation of the size of binomial index of model fit value is similar to the interpretation that Joreskog and Sorbom (1986) suggest for the likelihood-ratio chi-square test. According to Joreskog et al. the assumptions of the chi-square test are seldom met, thus, they state: "Instead of regarding χ^2 as a test statistic, one should regard it as a goodness (or badness) of fit measure in the sense that large χ^2-values correspond to bad fit and small χ^2-values to good fit" (p. I-39). With the binomial index of model fit, low values indicate good fit and high values reflect a poor fit between the data and the nomological net.

When examining this binomial index of model fit approach to gauging how well the data fit a nomological net, one may wonder why the researcher would not just consider using the proportion of paths supported by the data as the measurement of model fit. For example, if 4 of 5 paths were supported by the data in a model, the researcher would simply report that .8 of the paths were supported. If 16 of 20 paths in a model were supported by the data, again, the researcher would report that .8 of the paths were supported. Using the proportion of paths supported by the data as an index of model fit would produce the same value for these two studies. The binomial index of model fit values for these studies, however, would be .188 and .006, respectively. We believe that the model with 16 of 20 paths supported is intuitively "better" than 4 of 5, which the binomial index of model fit values would indicate, but the simple proportions of paths supported would not.

Selection of the Journal Articles

Four journal articles from various fields of study that contained structural equation models were selected to demonstrate two points. First, the articles were used to illustrate how the binomial index of model fit could be used to evaluate the models. Second, the articles were selected to show that the binomial index supports the conclusions derived from the more traditional goodness-of-fit measures in some, but not all, studies.

Specifically, the criteria used to select the four articles used in this paper were that the articles must: (a) analyze at least one structural equation model or nomological net, (b) report at least one traditional goodness-of-fit measure, (c) list the statistical test results of the path estimates, (d) be published during the period 1992-94. Of the studies that met these criteria, two were selected to illustrate that the conclusions drawn from the binomial index of model fit may agree with the conclusions drawn from the more traditional goodness-of-fit measures; and two articles were specifically chosen to demonstrate that the conclusions drawn from the two approaches may be at odds with each other.

Goodness-of-Fit Measures Utilized in the Articles

Each of the four articles included one or more of the following traditional measures of model fit: (a) likelihood-ratio chi-square statistic, (b) goodness-of-fit index (GFI), (c) root mean square residual (RMSR), (d) adjusted goodness-of-fit index (AGFI), and (e) normed fit index (NFI). Before we compare the interpretations of traditional goodness-of-fit indexes reported in the articles to the interpretations of the binomial index of model fit results calculated in this paper, it will be helpful to review what these five traditional goodness-of-fit indexes measure.

According to Hair, et al. (1992), each of the traditional goodness-of-fit measures can be placed under one of the following three classifications: (a) absolute fit measures, (b) incremental fit measures, and (c) parsimonious fit measures. The five traditional goodness-of-fit measures reported in the four articles used in this paper were classified as being either absolute fit measures or parsimonious fit measures.

Absolute Fit Measures

The likelihood-ratio chi-square statistic, the GFI, and the RMSR are classified under a group of goodness-of-fit measures referred to as measures of absolute fit. The goodness-of-fit measures placed under in this category attempt to determine the degree to which the model predicts the observed covariance or correlation matrix. The likelihood-ratio chi-square statistic is used to test the null hypothesis that the specified model leads to an exact reproduction of the population covariance matrix of the observed variables (Bollen & Long, 1993). This likelihood-ratio chi-square statistic is the only statistically based goodness-of-fit measure (Hair et al., 1992).

A chi-square value that results in a probability value of less than .05 (or .01) would indicate that the actual and predicted matrices are statistically significantly different (Joreskog & Sorbom, 1986). Fornell (1983) suggests, however, that significance levels of the chi-square test should exceed .1 or .2 before one fails to reject the null hypothesis. It should be noted, as previously discussed in this paper, that Joreskog and Sorbom (1986) recommend that the chi-square value be interpreted as a goodness-of-fit measure rather than a statistical test.

The likelihood-ratio chi-square test is sensitive to the sample size. Joreskog (1969) noted that small differences between the actual and predicted matrices for a large data set can lead the researcher to reject the null hypothesis. Thus, researchers have been encouraged to use other goodness-of-fit measures along with the likelihood-ratio chi-square test (Hair et al., 1992).

The GFI measures the relative amount of variances and covariances jointly accounted for by the model (Joreskog & Sorbom, 1986). That is, the GFI compares the prediction data to the actual data. The value for the GFI will fall between 0 and 1, with a higher value indicating a better fit. Volkan (1991) suggest that GFI values of .9 or higher are generally considered to indicate an adequate fit of the model.

The RMSR is the square root of the mean of the residuals between the observed and estimated matrices. The lower the RMSR value, the better the data fit the model. The RMSR does not have an established threshold value (Hair, et al., 1992).

Parsimonious Fit Measures

The adjusted goodness-of-fit index (AGFI) and the normed chi-square are two measures of goodness of fit that belong to a third classification, which is referred to as parsimonious fit measures. The various parsimonious fit measures adjust the goodness-of-fit of the model for the number of estimated coefficients that was required to obtain the observed level of fit. As indicated by Joreskog and Sorbom (1986), the AGFI is calculated as follows:

$$AGFI = 1 - [k(k+1)/2d](1-GFI) \qquad \text{Equation 3}$$

where:

1. $\underline{k}$ is equal to the number of observed variables analyzed.
2. $\underline{d}$ is the degrees of freedom for the proposed model.

Tanaka (1993) noted that this AGFI is a shrunken version of the GFI.

The normed chi-square goodness-of-fit measurement is the ratio of the chi-square of model fit divided by the degrees of freedom (Joreskog, 1970). Wheaton, Muthen, Alwain, and Summers (1977) suggest that the normed chi-square value should be less than 5.0 before the researcher concludes that there is a good fit between the data and the theoretical model.

Comparison of the Goodness-of-Fit Results

Two of the four studies utilized in this paper were selected to demonstrate that the binomial index of model fit may support the conclusions drawn from the traditional goodness-of-fit measures. The other two studies were selected to illustrate that the conclusions drawn from the traditional goodness of fit measures and the binomial index of model fit will not always be in agreement.

Consistent Conclusions

In the first study Kaplan, Peck, and Kaplan (1994) used structural equation modeling to evaluate the linkage between self-derogation in a school context and subsequent academic failure that is mediated by a disposition to deviate from conventional expectations. The model specified 18 paths among the five exogenous variables and four endogenous variables included in the model.

The results of analyses obtained form LISREL VII indicated that 16 our of the 18 paths were statistically significant. Using Equation 1, the binomial index of model fit value for a situation in which 16 our of 18 paths were supported by the data was calculated to be .0007. For this study, the value for $\underline{n}$ was 18 and the values for $\underline{x}$ were 16, 17, and 18. Since the binomial index of model fit value of .0007 is less than .05, we were willing to conclude that the data provided substantial support for the proposed nomological net.

The three traditional goodness-of-fit measures reported in the Kaplan et al. (1994) study also indicated that the theoretical model was supported by the data. The authors state that they considered the ratio of chi-square to its degrees of freedom value (5.64) to be adequate. In addition, they referred to the values of the GFI (.992) and the AGFI (.963) as being extremely large. Thus, the conclusion drawn from the traditional goodness of fit measures was consistent with the conclusion obtained from the binomial of model fit value.

In the second study Calantone, di Benedetto, and Bhoovaraghavan (1994) analyzed the data from United States business firms who were involved in product innovation. The authors constructed a nomological net that contained three exogenous variables and four endogenous variables. The EQS statistical computer package, version 3, was used to analyze the set of equations constructed to represent the relationships depicted in a nomological net. Eight of the nine path estimates in the nomological net were statistically significant. Thus, in the Equation 1, the value for n was equal to 9 and the values for x were 8 and 9. The binomial index value was calculated to be .02. Since the binomial index of model fit value for this study (.02) was less than .05, we believe that the results indicate that the data fit the proposed nomological net quite well.

Calantone et al. (1994) report that the chi-square statistic for the overall model fit was 27.054 (p = .008). The authors, who were faced with the problem often encountered with the use of the model fit chi-square statistic in a study that utilized a relatively large sample, i.e., small discrepancies between the actual and predicted matrices can cast doubt on the model specification, were willing to state that: "As the model is barely (italics added) significant at the .01 level, it is reasonably well fit by the data" (p. 146).

If we are willing to accept this "liberal" interpretation of the chi-square statistic value of model fit, the conclusion drawn by Calantone et al. (1994) would be consistent with the conclusion drawn from the binomial index of model fit value of .02. That is, the data support the nomological net. In addition, we believe that the binomial index of model fit value is more representative of the degree of model fit.

Inconsistent Conclusions

In the third study, which was conducted by McCarty and Shrum (1994), structural equation modeling was used to investigate the relationships of personal values, value orientations, and attitudes about recycling of waste with the frequency of recycling of waste behaviors. The structural equation model specified four exogenous variables and three endogenous variables with 17 specified paths.

Using LISREL VII, McCarty et al. (1994) found 8 of the 17 paths to be statistically significant. Utilizing Equation 1 with the value of n equal to 17 and the values for x ranging from 8 to 17, inclusive, the binomial index of model fit value was calculated to be .69. Since the binomial index of model fit value (.69) far exceeds .05, we would question the accuracy of the proposed nomological net or the theoretical model based on the analysis of this data set.

The conclusion concerning model fit drawn by the authors of the article, which was based on three traditional goodness-of-fit measures, was quite different from the conclusion that we were willing to state based on the binomial index of model fit value. The likelihood-ratio chi-square statistic of 182.29 (p = .281) as not statistically significant. And, even though, the GFI was .869 was slightly below the heuristic value of .90, McCarty and Shrum (1994) stated that: "The measures of overall goodness of fit for the entire model were acceptable" (p. 57).

We do not conceptually accept the position that the data support the proposed model when only 8 of 17 paths were supported by the data. The binomial index of model fit value, which was .69, is consistent with our view that such a statement of support would be questionable.

In the fourth study, Cabrera, Castaneda, Nora and Hengstler (1992) analyzed various models of college student departure decisions. One model analyzed in their study, which was referred to as the Student Attrition Model, contained three exogenous variables and seven endogenous variables with 19 paths depicting the relationships among these variables. Ten of the 19 paths had statistically significant path estimates. Using Equation 1 with the value of n set equal to 19 and the values of x ranging from 10 to 19, inclusive, the binomial index value was calculated to be .50. Since the binomial index of model fit value (.50) exceeds .05, we would conclude that the data do not support the nomological net.

The chi-square for overall model fit, the GFI, and the AGFI, and RMSR values for this model were 88.17 (p = .002), .981, .963, and .056, respectively.

Cabrera et al. (1992) used these traditional goodness-of-fit values to conclude that: "Although the Chi-square was significant (p = .002), this model can be seen to fit quite well based on the GFI, the AGFI, and the RMR" (p. 152).

Again, we would disagree with the conclusion that the data support the nomological net when 8 of the 19 paths, which is nearly one half of the paths in the model, were not supported by the data. The binomial index value of .50, which far exceeds .05, would once again lead the researcher to question the view that the data fit the proposed theoretical model.

Summary

The binomial index of model fit is designed to provide the researcher with a goodness of fit value that is based on the number of path estimates and/or relationships supported by the data in the nomological net. As the number of supported parameter estimates increases relative to the total number of paths in the model, the binomial index of model fit value decreases. If the binomial index value is less than .05, which is a value selected to assist the researcher in separating a signal from noise, the researcher could state that the data were supportive of the proposed theoretical model.

As demonstrated in this paper, the conclusions regarding model fit that are drawn form the binomial index value may not agree with those obtained from the more traditional goodness-of-fit measures. For example, it is possible to obtain high values for the GFI and the AGFI, which indicate a good fit between the data and the theoretical model, when less than one half of the path estimates are supported by the data. In such a case, the binomial index of model fit value would be high, which would indicate a questionable fit of the model.

It is our position that when a significant number of the path estimates are not supported by the data, the data do not support the proposed nomological net. Likewise, if many of the path estimates are supported by the data, the data do support at least a substantial portion of the proposed nomological net. We have proposed the binomial index of model fit in order to provide researchers with a goodness-of-fit measure that is based on the number of supported paths in the theoretical model.

The greater the number of perspectives from which a researcher can evaluate their model, the more likely it is that they will have an accurate picture of the situation. We agree with Schumacker (1992), who stated that: "Structural equation modeling requires the use of various GOF [goodness-of-fit] criteria in combination to interpret a model" (p. 13). The use of the binomial index of model fit along with the more traditional goodness-of-fit measures will provide additional insight into the degree to which the data fit the model.

References

Bentler, P. M., & Bonett, D. C. (1980). Significance test of goodness of fit in the analysis of covariance structures. Psychological Bulletin, 88, 588-606.

Bollen, K. A., & Long, J. S. (1993). Introduction. In K. A. Bollen & J. S. Long (Eds.), Testing structural equation models (pp. 1-9). Newbury park, CA: Sage.

Cabrera, A. F., Castaneda, M. B., Nora, A., & Hengstler, D. (1992). The convergence between two theories of college persistence. Journal of Higher Education, 63(2), 143-164.

Calantone, R. J., di Benedetto, C. A., & Bhoovaraghavan, S. (1994). Examining the relationship between degree of innovation and new product success. Journal of Business Research 30(2), 143-148.

Fornell, C. (1983). Issues in the application of covariance structure analysis: A comment. Journal of Consumer Research, 9, 443-448.

Fraas, J. W. & Newman, I. (1994). A binomial test of model fit. Structural Equation Modeling: A Multidisciplinary Journal, 1(3), 268-273.

Hair, J. F., Jr., Anderson, R. E., Tatham, R. L., & Black, W. C. (1992). Multivariate data analysis with readings, (3rd ed.) New York: MacMillan.

Joreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. Psychometrika, 34, 183-202.

Joreskog, K. G. (1970). A general method for analysis of covariance structures. Biometrika 57, 239-251.

Joreskog, K. C. & Sorbom, D. (1986). Lisrel VI: Analysis of linear structural relationships by the method of maximum likelihood (3rd ed.). Morrsville, IN: Scientific Software.

Kaplan, D. S., Peck, B. M., & Kaplan, H. B. (1994). Structural relations model of self-rejection, disposition to deviance, and academic failure. Journal of Educational Research, 87(3), 166-173.

McCarty, J. A. & Shrum, L. J. (1994). The recycling of solid wastes: Personal values, value orientations, and attitudes about recycling as antecedents of recycling behavior. Journal of Business Research, 30(1), 53-62.

Newman, I. (1990, April). To path analyze or not to path analyze: Is there an alternative approach? Paper presented at the meeting of the American Educational Research Association, Chicago.

Newman, I.; Fraas, J. W. & Norfolk, T. (in press). The binomial test of model fit: An elaboration and a response to comments. Journal of Structural Equation Modeling: A Multidisciplinary Journal.

Schumacker, R. E. (1992, April). Goodness-of-fit criteria in structural equation models: A summary with recommendations. Paper presented at the meeting of the American Educational Research Association, San Francisco.

Seigel, S. & Castellan, N. J. (1988). Nonparametric statistics for the behavioral sciences (2nd ed.). New York: McGraw-hill.

Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. In K. A. Bollen & J. S. Long (Eds.), Testing structural equation models (pp. 10-39). Newbury Park, CA: Sage.

Volkan, K. (1990, April). Using LISREL: Some notes on technique and interpretation. Paper presented at the meeting of the American Educational Research Association, Chicago.

Wheaton, B., Muthen, D., Alwin, D., & Summers, G. (1977). Assessing reliability and stability in panel models. In D. Heise (Ed.), Sociological Methodology. San Francisco: Jossey-Bass.

Wheeler, G. (1990). Understanding industrial experiments (2nd ed.). Knoxville, TN: SPC Press.