A Structural Equation Modeling Approach to E-Readiness

Amir Teimour Payandeh Najafabadi[1], Maryam Omidi Najafabadi[2]
[1] Department of Mathematical Sciences, Shahid Beheshti
University, G.C. Evin, 1983963113, Tehran, Iran
amirtpayandeh@sbu.ac.ir
[2] Islamic Azad University, Science and Research Branch, Tehran, Iran
m.omidi@srbiau.ac.ir

ABSTRACT: *Many authors considered measuring (e-) readiness of their educational training systems. Then, they employed either regression or factor analysis methods to model impact of factors or variables on (e-) readiness of their educational systems. This article suggests employing the sophisticate and well developed structural equation modeling (SEM) to model impact of other factors on (e-) readiness of a given educational training system. An application for a hybrid training system for Probability and Statistics for Engineers has been given.*

Keywords: E-readiness; Regression; Factor analysis; structural equation modeling (SEM).

Received: 19 May 2010, Revised 29 June 2010, Accepted 4 July 2010

© DLINE. All rights reserved

1. Introduction

Merriam-Webster's collegiate dictionary (2003) defines readiness as being prepared, mentally or physically, for some experience or actions. Borotis & Poulymenakou (2004) defined e-learning readiness of an organization as preparedness, mentally or physically, for some e-learning experiences or actions.

Readiness or e-readiness of an educational training system usually measure by several questions. There are the following statistical approaches to study (e-) readiness of an educational system: (i) The (e-) readiness' questions using an appropriate central tendency index (such mean, median, mode) summarize into a single response variable. Then, they employ an appropriate regression model to study impact of other independent variables on the (e-) readiness; (ii) The (e-) readiness' questions as well as other questions using confirmatory or explanatory factor analysis (FA) summarizes into a several factors. Then, such developed factors order based upon their explanation on the common variances. Each statistical approach has several advantages and disadvantages. For instance, the regression approach may be criticize by the fact that "How one may summarize several different variables into a single variable". More precisely, from statistical viewpoint such approach cannot be justified theatrically. The regression's approach may be admired since it can be answer to research hypothesizes. The FA eliminates such weakness of regression method. But, it cannot answer to any research hypothesizes.

The Structural equation modeling (SEM) translates of hypothesized relationships into testable mathematical models. Then, such hypothesizes tested against empirical data. The SEM is a well known technique which employs to estimate, analyze, and test models that specify relationships among (observed and latent) variables, say, theoretical framework.

This article suggests employing the SEM technique to study (e-) readiness of an educational training system. It structured as the following. Section 2 provides a review on mathematical concept of the SEM. In Section 3, an application is given for an e-readiness study of a hybrid course for "probability and statistics for engineers". Conclusion and suggestions are provided in Section 4.

2. Structural equation modeling (SEM)

The Structural equation modeling (SEM) is a well known technique which employs to estimate, analyze, and test models that specify relationships among (observed and latent) variables, say, theoretical framework. The term of “latent variable models” refer to classes of hypothetical or theoretical variables (constructs) that cannot be observed directly, and treat observed variables as indicators of underlying constructs rather than perfectly measured representations of these same constructs. The SEM models are well recognized as the most important statistical method to establish an appropriate model to evaluate a series of simultaneous hypotheses about the impacts of latent variables and observed variables on the other variables (Shipley, 2000). Once a theoretical framework has been proposed, it can then be tested against empirical data. The relationships are described by parameters, say factor loadings, that indicate the magnitude of the effect (direct or indirect) that independent variables (either observed or latent) have on dependent variables (either observed or latent).

The theoretical framework of a SEM may be decomposed into two kind of mathematical relationships: (i) The confirmatory factor analysis (CFA); (ii) and SEM relationships. Such mathematical relationships study through an example in the next section. Appropriateness and goodness of a SEM are measured by several statistical indices and hypothesis tests. The chisquare goodness of fit test with its null hypothesis indicates that the model is valid (i.e., p-value>0.05 supports goodness of the model) and Root Mean Square Error of Approximation (RMSEA) index are two well known indices to measure appropriateness of a SEM model. The RMSEA takes into account the error of approximation in the population and asks the question “How well would the model fit the population covariance matrix if it were available?” The RMSEA less than 0.05 indicates good fit, and higher than 0.08 represents reasonable errors of approximation in the population, more detail on the RMSEA as well as other indices may be found in Shipley (2000) among others.

3. An application to e-readiness of a hybrid course

Probability and statistics for engineers is one of the most challenging courses for both instructors and students. Overloading of the course content, time limitation, and simultaneous offering the course with several difficult courses (such as fundamental of physics, multivariate calculus, differential equations) transform an interesting course to a difficult one. Many instructors suggest dropping some less important materials of the course, and teaching the rest with more care. But majority of them believe that the course contents have been chosen based upon students’ needs in other courses and their research. Therefore, it makes sense to teach the course using a training system which have no time limitation and can be adapted based upon learner abilities. To eliminate such barriers, Payandeh & Omidi (2010) designed a hybrid training system which combines some elements of both traditional and e-learning training systems to reduce disadvantage of both systems. Using the ordinal logistic regression approach, they study e-readiness of such training system, see Payandeh & Omidi (2010) for more detail.

Figure 1, based upon Payandeh & Omidi (2010), provides a theoretical framework for factor may impact on e-readiness of the hybrid training system which introduced by Payandeh & Omidi (2010).

The above theoretical framework may reduce to two mathematical relations: (**i**) among observed variables and latent variables, say CFA relations:

$$\begin{cases} X_1 = 1'\text{Skills of users`} + e_1 \\ X_2 = \lambda_1'\text{Skills of users`} + e_2 \\ X_3 = \lambda_2'\text{Skills of users`} + e_3 \\ X_4 = \lambda_3'\text{Skills of users`} + e_4 \\ X_5 = \lambda_4'\text{Skills of users`} + e_5 \end{cases} \qquad \begin{cases} X_1 = 1'\text{Skills of users`} + e_1 \\ X_2 = \lambda_1'\text{Skills of users`} + e_2 \\ X_3 = \lambda_2'\text{Skills of users`} + e_3 \\ X_4 = \lambda_3'\text{Skills of users`} + e_4 \\ X_5 = \lambda_4'\text{Skills of users`} + e_5 \end{cases}$$

$$\begin{cases} X_{11} = 1'\text{Learner attitude`} + e_{11} \\ X_{12} = \lambda_9'\text{Learner attitude`} + e_{12} \\ X_{13} = 1'\text{Equipments`} + e_{13} \\ X_{14} = \lambda_{10}'\text{Equipments`} + e_{14} \\ X_{15} = \lambda_{11}'\text{Equipments`} + e_{15} \end{cases} \qquad \begin{cases} Y_1 = 1'\text{E} - \text{Readiness`} + e_{16} \\ Y_2 = \lambda_{12}'\text{E} - \text{Readiness`} + e_{17} \\ Y_3 = \lambda_{13}'\text{E} - \text{Readiness`} + e_{18} \\ Y_4 = \lambda_{14}'\text{E} - \text{Readiness`} + e_{19} \\ Y_5 = \lambda_{15}'\text{E} - \text{Readiness`} + e_{20} \\ Y_6 = \lambda_{16}'\text{E} - \text{Readiness`} + e_{21} \\ Y_7 = \lambda_{16}'\text{E} - \text{Readiness`} + e_{22} \end{cases}$$

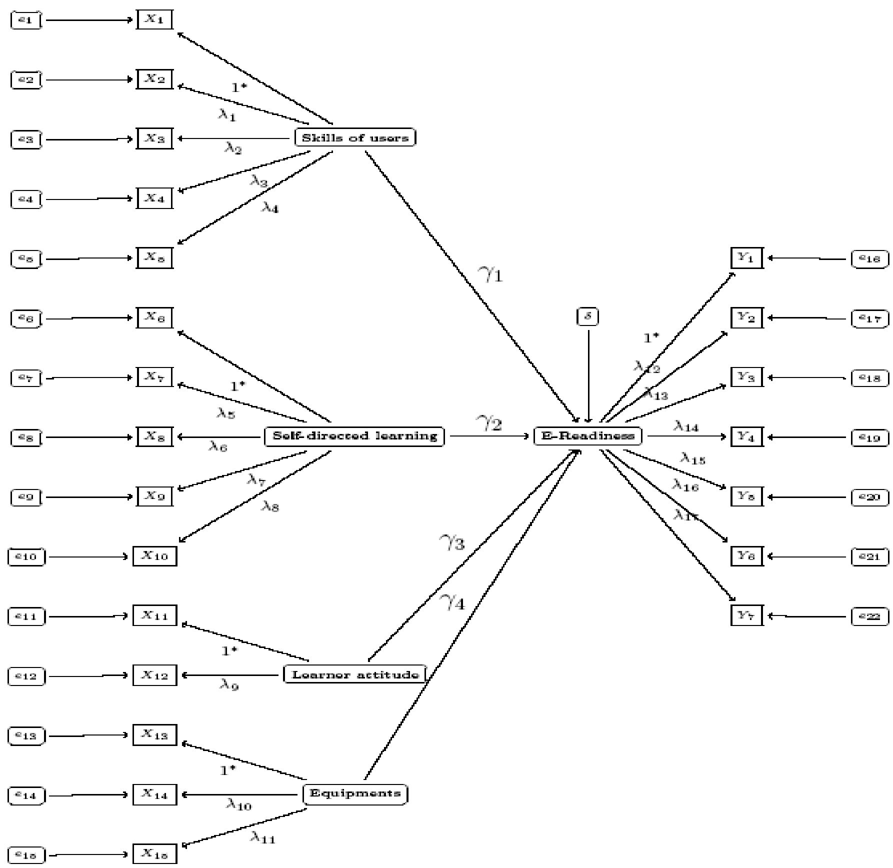

Figure 1. Theoretical Framework of E-readiness of the Hybrid Course for Probability and Statistics

And (**ii**) among latent variables, say SEM relation:

$$\text{E} - \text{Readiness} = \gamma_1' \text{Skills of users}' + \gamma_2' \text{Self} - \text{directed learning}' + \gamma_3' \text{Learner attitude}' + \gamma_4' \text{Equipments}' + \delta,$$

where respectively $X_1, \dots, X_{15}$ are Email, Hardware and Software, Internet, and Audio skills; Ability to: learn without assistance, resist on a given instruction, keep up with tasks, manage time, complete tasks on time; Learner's believe on e-learning and learner's computer anxiety; Learner's hardware, software, and internet equipments; and $Y_1, \dots, Y_7$ are Learner's IT skills, support of learner's parents on e-learning, learner's believe on his/her readiness for e-tasks (reviewing course slides, quizzes, assignments, and communication), and learner's believe on advantage of the course, respectively.

The above theoretical framework translates into the following hypothesizes:

Hypothesis 1. Skills of users doesn't influence on learners' e-readiness for the course.

Hypothesis 2. Self-directed ability of learners doesn't influence learners' e-readine

Hypothesis 3. Learners' attitude toward e-learning doesn't influences on their e-readiness for the course.

Hypothesis 4. Equipments doesn't influence on learners' e-readiness for the course.

Now, using the 109 collected observations explained in Payandeh & Omidi (2010), against the above theoretical framework the following conceptual framework has been produced by Lisrel 8.72.

The goodness of the model has been verified by chi-square test (with p-value=0.25>0.05) and the RMSEA (with value=0.033<0.05).

From the above model, one can conclude that: (i) Linear relationship among e-readiness, as an independent latent variable and other dependent latent variables can be stated as:

$$E - \text{Readiness} = 3.89 *' \text{Skills of users}` + 1.66 *' \text{Self} - \text{directed learning}` + 0.26 *' \text{Learner attitude}' + 0.14 * '\text{Equipments}` + \delta.$$

(ii) Using estimated factor loading (regression approach to e-readiness), effect of such independent factors on e-readiness can be ordered as "Skill of users", "Self-directed Learning", "Leaner attitude toward e-learning", and "Leaner's equipments"; (iii) Since error of the above model is 0.230, therefore, about 90% of data are described by the model; (iv) "ability of learners to manage their time" (x_9) and "Audio skill" of learners (x_5), respectively, provide more indirect impact on the e-learning, among variables included in Learner factor; (v) all the above hypothesizes are significant at level $\alpha = 0.05$; (vi) Using explained common variance (factor analysis approach to e-readiness), effect of each factors can be ordered as "Skill of users" (35.06%), "Self-directed Learning" (9.77%), "Leaner attitude toward e-learning" (9.01%), and "Leaner's equipments" (6.93%); and (vii) 60.77% of total common variance has been explained by the model. Conclusions (i) to (v), may also obtain from a regression approach to the problem, while conclusions (vi) and (vii) may arrive also from a factor analysis approach.

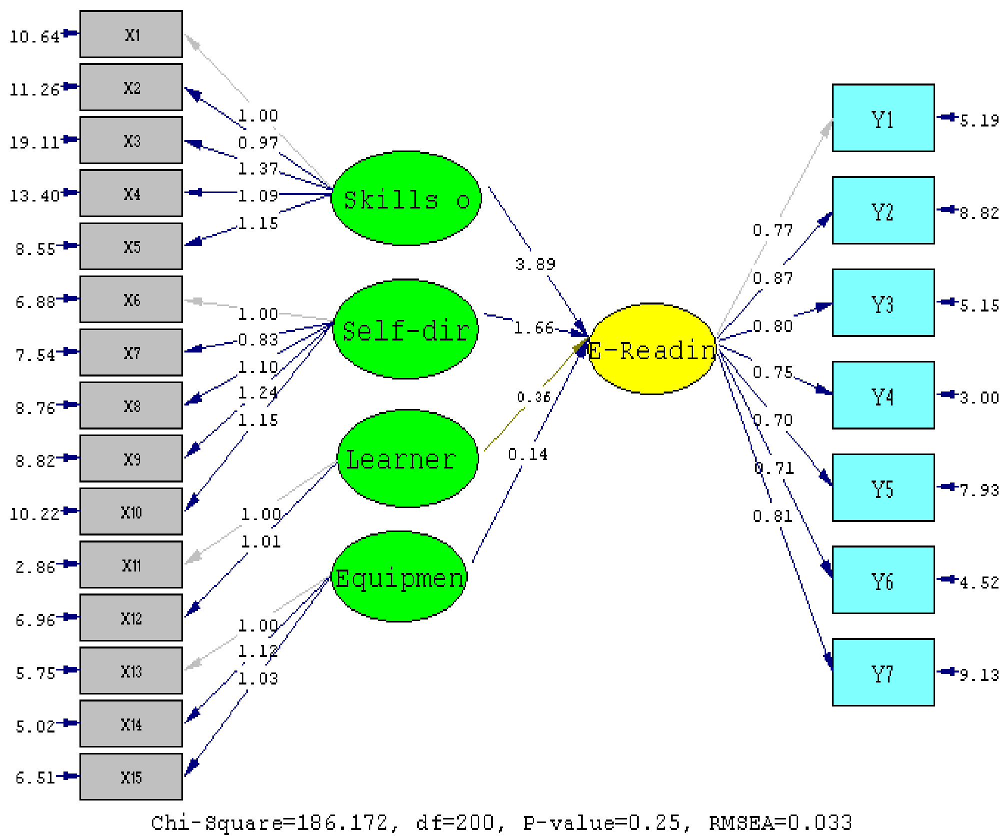

Figure 2. Conceptual Framework of E-readiness of the Hybrid Course for Probability and Statistics

4. Conclusion and suggestions

This article suggests using the structural equation modeling (SEM) as an alternative statistical method for both regression and factor analysis approaches to study (e-) readiness of an educational training system. The SEM combines benefits of both regression and factor analysis approaches and introduces an elegant statistical technique to study (e-) readiness of an educational training system.

The usual SEM employs the maximum likelihood (ML) method to estimate unknown parameters. It is well known that the statistical properties of the ML approach are asymptotic (Lehmann & Casella, 1998). Therefore, many of properties of the ML estimators have been oscillated for small sample size. In the context of some basic SEMs, many studies have been devoted to study the behaviors of the ML asymptotic properties with small sample sizes, see Lee 2007 for an excellent review. It was concluded by such researches that the properties of the statistics are not robust for small sample sizes, even for the multivariate normal distribution. The Bayesian approach to the SEM has ability: (i) to work properly for small sample size. Even small sample size, the posterior distributions of parameters and latent variables can be estimated by using a sufficiently large number of observations that are simulated from the posterior distribution of the unknown parameters through efficient tools in statistical computing such as the various Markov chain Monte Carlo (MCMC) methods (Lee, 2007); (ii) to utilize useful and prior information about the problem (which translated to a prior distribution) to achieve better results. For situations without accurate prior information, some type of non-informative prior distributions can be used. In these cases, the accuracy of the Bayesian estimates is close to that obtained from the classical SEM (Robert, 2001); and (iii) to treat the discrete variables (such as the Likert and rating scales) as the hidden continuous normal distribution with a specified threshold (or cut point). Clearly, such approach provide a powerful tool to analyze the discrete variables rather than using special, but less powerful, statistical technique to do so (see Lee, 2007).

5. Acknowledgements

This article have presented at the Third International Conference on the Applications Digital Information and Web Technologies (ICADIWT 2010). The authors would like to pass on their deep appreciation to conference's Committee and referee of this article for his constrictive comments.

References

[1]. Borotis, S. & Poulymenakou, A. (2004). E-learning readiness components: key issue to consider before adopting e-learning interventions. Available at: http://www.eltrun.gr/papers/ereadiness-elearn2004.pdf

[2]. Lee, S. Y (2007). Structural Equation Modeling: A Bayesian Approach. John Wiley.

[3]. Lehmann, E. L. & Casella, G. (1998). Theory of point estimation. 2nd ed., Springer-Verlag, New York.

[4]. Merriam-Webster, Inc (2003). Merriam-Webster's collegiate dictionary. Merriam-Webster, Inc

[5]. Payandeh Najafabadi A. T. & Omidi Najafabadi M. (2010). A Hybrid course for probability and statistics for engineers: An e-readiness study at Shahid Beheshti University. iJET International journal: emerging technologies in learning, 3, 18-26.

[6]. Robert, C. (2001). The Bayesian Choice. 2nd ed., Springer-Verlag, New York.

[7]. Shipley, B. (2000). Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference. Cambridge University Press, Cambridge UK.