BAYESIAN ANALYSIS OF STOCHASTIC CONSTRAINTS IN STRUCTURAL EQUATION MODEL WITH POLYTOMOUS VARIABLES IN SEVERAL GROUPS

by

Tung-lok NG

A Thesis
submitted to
the Graduate School of
The Chinese University of Hong Kong
(Division of Statistics)

In Partial Fulfillment
of the Requirements for the Degree of
Master of Philosophy (M. Phil.)

Hong Kong
June 1990

THE CHINESE UNIVERSITY OF HONG KONG

GRADUATE SCHOOL

The undersigned certify that we have read a thesis, entitled "Bayesian Analysis of Stochastic Constraints in Structural Equation Model with Polytomous Variables in Several Groups" submitted to the Graduate School by Tung-lok NG (吳東樂) in partial fulfillment of the requirement for the degree of Master of Philosophy in Statistics. We recommend that it be accepted.

Dr. S. Y. Lee,
Supervisor

W. Y. Poon

Dr. W. Y. Poon

N. N. Chan

Dr. N. N. Chan

Professor P. M. Bentler,
External Examinar

DECLARATION

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institution of learning.

ACKNOWLEDGMENT

I would like to express my sincere thanks to Dr. S.Y. Lee for his encouragement and his supervision on my thesis. At the same time, the kind assistance of Dr. W. Y. Poon and Mr. K. H. Leung in technical consultation is gratefully acknowledged.

ABSTRACT

In this thesis, structural equation model with polytomous variables in several groups is analyzed in the presence of stochastic constraints. Prior distributions of the structural parameters are considered based on a Bayesian point of view. An iterative procedure is implemented to produce the various Bayes estimates. It is shown via a simulation study that the Bayesian approach are more flexible as well as more accurate than the ordinary maximum likelihood approach.

Key words: Structural equation model, prior distributions, conjugate family, Bayesian approach.

CONTENT

Chapter 1

Introduction

The analysis of structural equation models, also known as covariance structure analysis, is an applied multivariate technique in analyzing causations and correlations among latent and observed random variables. Due to its distinctive features, this method of analysis allows researchers to effectively study ample problems that could not be easily solved using alternative approaches, see Newcomb and Bentler (1988).

With the development of highly sophisticated package programs such as LISREL VII (Jöreskog & Sörbom, 1988) and EQS (Bentler 1989), the method has been widely employed in many branches of study especially in behavioral and social sciences researches. Nevertheless, the applicability of these packages relied heavily on the assumption that the observed random variables are continuous. In real life experience, however, we frequently encounter variables of dichotomous or polytomous form. For instance, suppose in an opinion survey, an respondent is asked to answer question concerning their attitudes towards a particular issue on scale like

strongly favor	favor	neutral	unfavor	strongly unfavor .

In such circumstance, the usefulness of these packages may be greatly reduced.

To overcome such deficiencies, one direction of recent development is to extend the basic theory to handle data of dichotomous or polytomous form. In the literature, Bock and Lieberman (1970), Christoffersson (1975), and Muthen (1978) had respectively considered dichotomous factor analysis using either the maximum likelihood approach or the generalized least squares approach. More recently, Lee, Poon and Bentler (1989a) have developed theory for analyzing general covariance structure models with polytomous variables using the maximum likelihood approach.

In the meantime, another exciting development in structural equation modeling is the incorporation of auxiliary prior information. The provision of prior information in the form of exact equality constraints provide researchers more flexibility in defining appropriate structures to model many realistic problems. The widely publicized package programs LISREL and EQS have provided the option to allow users to impose simple exact equality constraints as well. Lee and Bentler (1980) have developed theory in analyzing general structural equation models in the presence of exact prior information. In addition, Lee (1988a,b) extended the previous work to consider prior information of stochastic nature. Clearly, such advancement provides more freedom in studying the functional relationship among parameters in the model.

Finally, there is an increasing trend to consider the general model in several populations. Multiple populations models frequently arise when we consider data coming from different sex groups, ethnic groups, treatment groups or the like. Major interests rest on the comparison of covariance structures across populations. In the literature, Jöreskog (1971) considered simultaneous factor analysis in several populations. Sörbom (1974) proposed a general method for studying differences in factor means and factor structure between groups. Lee and Tsui (1982) generalized the basic results in Jöreskog (1971) to general covariance structure models with functional constraints. However, it is worth to note that all the cited work were restricted to continuous variables only. Although Muthen and Christoffersson (1981) has worked out the simultaneous factor analysis model with categorical variables in several groups, the results were still restricted to the simpliest dichotomous case only. Later on, Poon, Lee, Bentler and Afifi (1989) developed a computationally efficient multi-stage estimation procedure to analyze general covariance structure model with polytomous variables in several groups.

The primary objective of this paper is to extend Lee (1988b)'s work to consider the Bayesian analysis of stochastic prior information in structural equation model with data of polytomous form and coming from several populations or groups. At the same time, we are also interested to study the performance of the approach as compared to the classical maximum likelihood method. The order of presentation is as follows : In Chapter 2, the general structural equation model proposed in Poon, Lee, Bentler and Afifi (1989) is presented and the full maximum likelihood estimation procedure based on Lee, Poon and Bentler (1989a) is introduced. An artificial example is given to illustrate the implementation of the procedure. In Chapter 3, stochastic prior information in the form of stochastic constraints has been incorporated into the general model and the estimation technique based on the Bayesian approach is studied. As before, an artificial example would be given to illustrate the method. A series of simulation studies have been conducted to examine the effectiveness of the stochastic prior information by comparing the accuracy of the various Bayes estimates to the ordinary maximum likelihood estimates. The results are reported in Chapter 4. A brief discussion of the findings and the final conclusions are presented in Chapter 5.

Chapter 2

Full Maximum Likelihood Estimation of the General Model

§ 2.1 Introduction

In this chapter, we will present the general structural equation model proposed in Poon, Lee, Bentler and Afifi(1989). Instead of analyzing the model by the multi-stage estimation procedure, the classical maximum likelihood approach based on Lee, Poon and Bentler (1989a) is employed. Finally, as the general solutions of the ML estimates cannot be solved in closed form, the iterative scoring algorithm has been used to obtain the estimates.

§ 2.2 Model

Consider a set of G independent populations or groups arising from different culture groups, sex groups, etc. Assume p is the common number of variables in each group. For the g^{th} group, let $\underset{\sim}{Z}^{(g)} = (Z_1^{(g)},\ldots,Z_p^{(g)})'$ denote a vector of p observed polytomous variables, $g = 1,\ldots,G$. Suppose the corresponding latent continuous random vector $\underset{\sim}{Y}^{(g)} = (Y_1^{(g)},\ldots,Y_p^{(g)})$ is multivariately normally distributed with mean $\underset{\sim}{0}$ and covariance matrix $\underset{\sim}{\Sigma}^{(g)}(\underset{\sim}{\theta}^{(g)}) = \left[\sigma_{ij}^{(g)}(\underset{\sim}{\theta}^{(g)})\right]$, where $\sigma_{ij}^{(g)}$ are functions of the unknown structural parameter vector $\underset{\sim}{\theta}^{(g)}$. $\underset{\sim}{Z}^{(g)}$ and $\underset{\sim}{Y}^{(g)}$ are related by

$$Z_i^{(g)} = k(i) \quad \text{if} \quad \alpha_{i,k(i)}^{(g)} \leq Y_i^{(g)} < \alpha_{i,k(i)+1}^{(g)} \qquad (1)$$

for $i=1,\ldots,p$ and $k(i)=1,\ldots,m(i)$. Here, $m(i)$ denotes the number of categories corresponding to the i^{th} variable. To simplify matter further, we do not consider the dichotomous case and assume $m(i) \geq 3$ as well as invariant over the groups. These categories are defined by a set of thresholds,

$$\underset{\sim}{\alpha}_i^{(g)} = \left\{\alpha_{i,1}^{(g)},\ldots,\alpha_{i,m(i)+1}^{(g)}\right\} \tag{2}$$

with $\alpha_{i,1}^{(g)} = -\infty$ and $\alpha_{i,m(i)+1}^{(g)} = \infty$ for all g.

The vector $\underset{\sim}{Y}^{(g)}$ is unobservable and we only have a random sample of $\underset{\sim}{Z}^{(g)}$ with size N_g. Therefore, altogether we have the frequencies of G independent p-way contingency tables which are obtained based on the value of $\underset{\sim}{Z}^{(g)}$. The observed frequency of the $\underset{\sim}{k} = (k(1),\ldots,k(p))^{th}$ cell in the g^{th} group is denoted by $f_{\underset{\sim}{k}}^{(g)}$. The probability that an observation in g^{th} group falls into the $\underset{\sim}{k}^{th}$ cell is given by

$$\xi_{\underset{\sim}{k}}^{(g)} = \Pr\left\{ Z_1^{(g)} = k(1),\ldots,Z_p^{(g)} = k(p) \right\}$$

$$= (-1)^p \sum_{i(1)=0}^{1} \cdots \sum_{i(p)=0}^{1} (-1)^{\sum_{j=1}^{p} i(j)} \times \phi_p\left[\alpha_{1,v(1)}^{(g)},\ldots,\alpha_{p,v(p)}^{(g)};\ \underset{\sim}{\Sigma}^{(g)}(\underset{\sim}{\theta}^{(g)})\right] \tag{3}$$

where $v(j) = k(j) + i(j)$; and

$$\phi_p(\alpha_1,\ldots,\alpha_p;\ \underset{\sim}{\Sigma}) = \int_{-\infty}^{\alpha_1}\cdots\int_{-\infty}^{\alpha_p} (2\pi)^{-\frac{p}{2}}|\underset{\sim}{\Sigma}|^{-\frac{1}{2}}\exp(-\underset{\sim}{y}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{y}/2)\ dy_1\ldots dy_p\ . \tag{4}$$

§ 2.3 Identification of the model

Suppose $\underset{\sim}{D}$ is any diagonal matrix with diagonal elements $d_{ii} > 0$. For any set of $(\alpha_1,\ldots,\alpha_p)$,

$$\Phi_p(\alpha_1,\ldots,\alpha_p;\ \underset{\sim}{\Sigma}) = \Phi_p(\alpha_1^*,\ldots,\alpha_p^*;\ \underset{\sim}{\Sigma}^*) \tag{5}$$

where

$$\underset{\sim}{\Sigma}^* = \underset{\sim}{D}\,\underset{\sim}{\Sigma}\,\underset{\sim}{D} \quad \text{and} \quad \alpha_i^* = \alpha_i d_{ii} \quad . \tag{6}$$

Therefore, for any group g, the parameters $\underset{\sim}{\alpha}^{(g)}$ and $\underset{\sim}{\theta}^{(g)}$ are not identified unless suitable constraints concerning the parameters are imposed. Unfortunately, it is difficult to give the general sufficient conditions for identification. Nevertheless, by Lee, Poon, and Bentler (1989a), it has been shown that we may identify all the parameters in any single group if we fix either the variances or the thresholds of the variables in that group to some specified values. The former identification condition, according to Lee and Poon (1985), will impose additional nonlinear restrictions on $\underset{\sim}{\theta}^{(g)}$ and hence may demand a great deal of computational effort in obtaining the various estimates. As a matter of convenience, we therefore adopt the following identification conditions :

(i) Consider an arbitrary group, say group r, fix the thresholds $\alpha_{1,2}^{(r)}$, $\alpha_{1,m(1)}^{(r)}$, $\alpha_{2,2}^{(r)}$, $\alpha_{3,2}^{(r)}$,..., $\alpha_{p,2}^{(r)}$ to some constants.

(ii) For any other group $g \neq r$, fix $\underset{\sim}{\alpha}_i^{(g)} = a_i \underset{\sim}{\alpha}_i^{(r)}$, i=1,...,p where a_i's are some fixed constants (in most cases, a_i=1).

From condition (i), for the reference group r, the only transformation $\underset{\sim}{D}$ in (6) that preserve restriction (5) is the identity matrix, $\underset{\sim}{I}$. Thus, the reference group is identified. Consider the possible transformations with arbitrary $\underset{\sim}{D}$ for other distinct groups. From condition (ii), for any group g, since $\underset{\sim}{\alpha}_i^{(g)} = a_i \underset{\sim}{\alpha}_i^{(r)}$, i=1,...,p are identified, this implies $\underset{\sim}{D} = \underset{\sim}{I}$ and hence that group is also identified. Following the similar argument, it is easy to show that the whole model is identified.

It is worth to note that the difference between the fixed parameters $\alpha^{(r)}_{1,m(r)}$ and $\alpha^{(r)}_{1,2}$ specified in condition (i) provides a standard for the measure of dispersion for other variables in the reference group. Thus, the condition is not restrictive at all. Moreover, if the reference model is scale invariant, the choice of the fixed thresholds only changes the scale of the covariance matrix but not its structure, see Lee, Poon and Bentler (1989a). As a result, the essential interpretation of the covariance structure will not be affected. On the contrary, condition (ii) restricts the thresholds of the other groups to satisfy a linear relationship with those in the reference group. Clearly, it is quite restrictive but still be acceptable especially when similar instruments have been administered to all groups. For instance, respondents from different groups are asked to answer questions on the same scale in a questionnaire. As the relations over groups are unchanged, the statistical inferences are unaffected by the choice of reference group. For simplicity, we let the first group as the reference group, i.e. r=1.

Before closing this section, it is important to mention that the method suggested here is not the only way to solve the identification problems. Clearly, different methods may lead to different special cases of the general model and hence to different interpretations of the parameters. Throughout this thesis, we will apply conditions (i) and (ii) to identify the general model.

§ 2.4 Maximum Likelihood Estimation

Basically, there are two kinds of parameters in the model, namely, the thresholds and the structural parameters. Let $\underset{\sim}{\alpha} = (\underset{\sim}{\alpha}^{(1)'},\ldots,\underset{\sim}{\alpha}^{(G)'})'$ with $\underset{\sim}{\alpha}^{(i)} = \left\{\alpha^{(i)}_{1,3},\ldots,\alpha^{(i)}_{1,m(1)-1},\alpha^{(i)}_{2,3},\ldots,\alpha^{(i)}_{2,m(2)},\ldots,\alpha^{(i)}_{p,3},\ldots,\alpha^{(i)}_{p,m(p)}\right\}'$

be the vector of all unknown thresholds, and let $\underset{\sim}{\theta} = (\underset{\sim}{\theta}^{(1)\prime}, \ldots, \underset{\sim}{\theta}^{(G)\prime})$ be the vector of all unknown structural parameters. Then the overall parameter vector is defined by

$$\underset{\sim}{\gamma} = (\underset{\sim}{a}', \underset{\sim}{\theta}')' \quad . \tag{7}$$

Suppose $\underset{\sim}{f} = \left\{ f_{\underset{\sim}{k}}^{(g)} \; ; \; g = 1,\ldots,G \; ; \; k(i) = 1,\ldots,m(i) \; ; \; i = 1,\ldots,p \right\}$ denotes the overall vector of the observed frequency counts for the G independent p-way contingency tables. The negative of the log-likelihood function for $\underset{\sim}{f}$ is given by

$$L(\underset{\sim}{\gamma}) = -\sum_{g=1}^{G} \sum_{k(1)=1}^{m(1)} \cdots \sum_{k(p)=1}^{m(p)} f_{\underset{\sim}{k}}^{(g)} \ln \xi_{\underset{\sim}{k}}^{(g)} \quad . \tag{8}$$

By (3), $\xi_{\underset{\sim}{k}}^{(g)}$ is a function of $\underset{\sim}{\theta}^{(g)}$ and $\underset{\sim}{a}^{(g)}$ and hence $L(\underset{\sim}{\gamma})$ is a function of $\underset{\sim}{\gamma}$ only. The ML estimate of $\underset{\sim}{\gamma}$ is the vector $\hat{\underset{\sim}{\gamma}}$ that minimizes $L(\underset{\sim}{\gamma})$. Under mild regularity conditions, it can be shown that $\hat{\underset{\sim}{\gamma}}$ possesses the following desirable statistical properties : (i) It is consistent; and (ii) its asymptotic distribution is multivariate normal with mean vector $\underset{\sim}{\gamma}$ and covariance matrix equal to the inverse of the information matrix. The gradient vector and the information matrix of L, $\dot{\underset{\sim}{L}}(\underset{\sim}{\gamma})$ and $\underset{\sim}{I}(\underset{\sim}{\gamma})$ are respectively defined as follow :

$$\dot{\underset{\sim}{L}}(\underset{\sim}{\gamma}) = \partial L / \partial \underset{\sim}{\gamma} \tag{9}$$

$$\underset{\sim}{I}(\underset{\sim}{\gamma}) = \{E(\partial L/\partial \underset{\sim}{\gamma})(\partial L/\partial \underset{\sim}{\gamma})'\} \quad . \tag{10}$$

Adopting the results in Lee, Poon & Bentler (1989b), we have

$$\dot{\underset{\sim}{L}}(\underset{\sim}{\gamma})_c = -\sum_{g=1}^{G} \sum_{k(1)=1}^{m(1)} \cdots \sum_{k(p)=1}^{m(p)} \frac{f_{\underset{\sim}{k}}}{\xi_{\underset{\sim}{k}}^{(g)}} \frac{\partial \xi_{\underset{\sim}{k}}^{(g)}}{\partial \gamma_c} \tag{11}$$

and

$$\underset{\sim}{I}(\underset{\sim}{\gamma})_{cd} = \sum_{g=1}^{G} N_g \sum_{k(1)=1}^{m(1)} \cdots \sum_{k(p)=1}^{m(p)} \frac{1}{\xi_{\underset{\sim}{k}}^{(g)}} \frac{\partial \xi_{\underset{\sim}{k}}^{(g)}}{\partial \gamma_c} \frac{\partial \xi_{\underset{\sim}{k}}^{(g)}}{\partial \gamma_d} \tag{12}$$

where

$$\frac{\partial \xi_{\underset{\sim}{k}}^{(g)}}{\partial \gamma_c} = (-1)^p \sum_{i(1)=0}^{1} \cdots \sum_{i(p)=0}^{1} (-1)^{\sum_{j=1}^{p} i(j)} \times \frac{\partial}{\partial \gamma_c} \Phi_p\left[a_{1,v(1)}^{(g)}, \ldots, a_{p,v(p)}^{(g)}; \underset{\sim}{\Sigma}^{(g)}(\underset{\sim}{\theta}^{(g)})\right] \quad . \tag{13}$$

Hence, the derivatives of the normal distribution function with respect to its parameters are required in computing the gradient vector and the information matrix.

For the reference group, suppose $\underset{\sim}{D}$ is a diagonal matrix with diagonal elements $\sigma_{jj}^{(1)^{\frac{1}{2}}}$. Define

$$\underset{\sim}{R} = \underset{\sim}{D}^{-1} \underset{\sim}{\Sigma}^{(1)} \underset{\sim}{D}^{-1} \qquad \text{and} \qquad a_j^* = \sigma_{jj}^{(1)^{-\frac{1}{2}}} a_{j,v(j)}^{(1)} \tag{14}$$

Note $\underset{\sim}{R}$ is a correlation matrix and we have

$$\Phi_p\left[a_{1,v(1)}^{(1)}, \ldots, a_{p,v(p)}^{(1)}; \underset{\sim}{\Sigma}^{(1)}\right] = \Phi_p(a_1^*, \ldots, a_p^*; \underset{\sim}{R}) \quad . \tag{15}$$

Since $\Phi_p(a_1^*, \ldots, a_p^*; \underset{\sim}{R})$ can be expressed as

$$\int_{-\infty}^{a_p^*} \phi(x_j)\, \Phi_{p-1}\left[\ldots, (a_h^* - \rho_{jh} x_j)/(1 - \rho_{jh}^2)^{\frac{1}{2}}, \ldots; \underset{\sim}{R}_{\cdot j}\right] dx_j$$

where j,h = 1,...,p and h ≠ j; $\phi(.)$ is the univariate standardized normal density function; ρ_{jh} is the $(j,h)^{th}$ element of $\underset{\sim}{R}$; and $\underset{\sim}{R}_{\cdot j}$ is the partial correlation matrix with the j^{th} variable partialled out. Following the fundamental theorem of calculus, we have

$$\partial\Phi_p(\alpha_1^*,\ldots,\alpha_p^*;\underset{\sim}{R})/\partial\alpha_j^* = \phi(\alpha_j^*)\Phi_{p-1}\left[\ldots,(\alpha_h^* - \rho_{jh}\alpha_j^*)/(1-\rho_{jh}^2)^{\frac{1}{2}},\ldots;\underset{\sim}{R}_{\cdot j}\right] . \quad (16)$$

Also from Johnson and Kotz (1972),

$$\partial\Phi_p(\alpha_1^*,\ldots,\alpha_p^*;\underset{\sim}{R})/\partial\rho_{jh} = \phi_2(\alpha_j^*,\alpha_h^*,\rho_{jh})\Phi_{p-2}(\ldots,\alpha_m^* - \tau_m,\ldots;\underset{\sim}{R}_{\cdot jh}) \quad (17)$$

where m ≠ j,h and j ≠ h, $\phi_2(.)$ is the standardized bivariate normal density function with correlation ρ_{jh}, $\underset{\sim}{R}_{\cdot jh}$ is the partial correlation matrix with the j^{th} and h^{th} variables partialled out, and

$$\tau_m = (1-\rho_{jh}^2)^{-\frac{1}{2}}\left\{(\rho_{jm} - \rho_{jh}\rho_{hm})\alpha_j^* + (\rho_{hm} - \rho_{jh}\rho_{jm})\alpha_h^*\right\} . \quad (18)$$

Now, by Lee, Poon & Bentler (1989a,b), we have

$$\frac{\partial\Phi_p(\alpha_{1,v(1)}^{(1)},\ldots,\alpha_{p,v(p)}^{(1)};\underset{\sim}{\Sigma}^{(1)})}{\partial\alpha_{j,v(j)}^{(1)}} = \left[\sigma_{jj}^{(1)}\right]^{-\frac{1}{2}} \frac{\partial\Phi_p(\alpha_1^*,\ldots,\alpha_p^*;\ \underset{\sim}{R})}{\partial\alpha_j^*} \quad (19)$$

$$\frac{\partial\Phi_p(\alpha_{1,v(1)}^{(1)},\ldots,\alpha_{p,v(p)}^{(1)};\underset{\sim}{\Sigma}^{(1)})}{\partial\theta_a} = \sum_{i=1}^{p}\sum_{j=1}^{i}\frac{\partial\sigma_{ij}^{(1)}}{\partial\theta_a}\frac{\partial\Phi_p(\alpha_{1,v(1)}^{(1)},\ldots,\alpha_{p,v(p)}^{(1)};\underset{\sim}{\Sigma}^{(1)})}{\partial\sigma_{ij}^{(1)}} \quad (20)$$

with

$$\frac{\partial \phi_p(\alpha^{(1)}_{1,v(1)},\ldots,\alpha^{(1)}_{p,v(p)};\underset{\sim}{\Sigma}^{(1)})}{\partial \sigma^{(1)}_{jj}} = - \frac{\alpha^{(1)}_{j,v(j)}}{2\left[\sigma^{(1)}_{jj}\right]^{\frac{3}{2}}} \frac{\partial \phi_p(\alpha^*_1,\ldots,\alpha^*_p;\ \underset{\sim}{R})}{\partial \alpha^*_j} \tag{21}$$

and

$$\frac{\partial \phi_p(\alpha^{(1)}_{1,v(1)},\ldots,\alpha^{(1)}_{p,v(p)};\underset{\sim}{\Sigma}^{(1)})}{\partial \sigma^{(1)}_{ij}} = \frac{1}{\left[\sigma^{(1)}_{ii}\ \ \sigma^{(1)}_{jj}\right]^{\frac{1}{2}}} \frac{\partial \phi_p(\alpha^*_1,\ldots,\alpha^*_p;\ \underset{\sim}{R})}{\partial \rho_{ij}} \quad . \tag{22}$$

Similarly, for other group $g = 2,\ldots,G$, the expressions for the various derivatives are as follows :

$$\frac{\partial \phi_p(a_1\alpha^{(1)}_{1,v(1)},\ldots,a_p\alpha^{(1)}_{p,v(p)};\underset{\sim}{\Sigma}^{(g)})}{\partial \alpha^{(1)}_{j,v(j)}} = a_j\left[\sigma^{(g)}_{jj}\right]^{-\frac{1}{2}} \frac{\partial \phi_p(\alpha^*_1,\ldots,\alpha^*_p;\ \underset{\sim}{R})}{\partial \alpha^*_j} \tag{23}$$

$$\partial \phi_p(a_1\alpha^{(1)}_{1,v(1)},\ldots,a_p\alpha^{(1)}_{p,v(p)};\ \underset{\sim}{\Sigma}^{(g)})/\partial\theta_a$$

$$= \sum_{i=1}^{p}\sum_{j=1}^{i} \frac{\partial \sigma^{(g)}_{ij}}{\partial \theta_a} \frac{\partial \phi_p(a_1\alpha^{(1)}_{1,v(1)},\ldots,a_p\alpha^{(1)}_{p,v(p)};\underset{\sim}{\Sigma}^{(g)})}{\partial \sigma^{(g)}_{ij}} \tag{24}$$

with

$$\frac{\partial \phi_p(a_1\alpha^{(1)}_{1,v(1)},\ldots,a_p\alpha^{(1)}_{p,v(p)};\underset{\sim}{\Sigma}^{(g)})}{\partial \sigma^{(g)}_{jj}} = - \frac{a_j\,\alpha^{(1)}_{j,v(j)}}{2\left[\sigma^{(g)}_{jj}\right]^{\frac{3}{2}}} \frac{\partial \phi_p(\alpha^*_1,\ldots,\alpha^*_p;\ \underset{\sim}{R})}{\partial \alpha^*_j} \tag{25}$$

$$\frac{\partial \phi_p(a_1\alpha^{(1)}_{1,v(1)},\ldots,a_p\alpha^{(1)}_{p,v(p)};\underset{\sim}{\Sigma}^{(g)})}{\partial \sigma^{(g)}_{ij}} = \frac{1}{\left[\sigma^{(g)}_{ii}\ \ \sigma^{(g)}_{jj}\right]^{\frac{1}{2}}} \frac{\partial \phi_p(\alpha^*_1,\ldots,\alpha^*_p;\ \underset{\sim}{R})}{\partial \rho_{ij}} \tag{26}$$

where $\underset{\sim}{D}$ is a diagonal matrix with diagonal elements $\sigma^{(g)\frac{1}{2}}_{jj}$;

$$\underset{\sim}{R} = \underset{\sim}{D}^{-1}\underset{\sim}{\Sigma}^{(g)}\underset{\sim}{D}^{-1} \quad \text{and} \quad \alpha_j^* = a_j \sigma_{jj}^{(g)-\frac{1}{2}} \alpha_{j,v(j)}^{(1)} \tag{27}$$

From (11), (12) and (13), expressions for the gradient vector and the information matrix can be obtained via (14) to (27).

§ 2.5 Computational Procedure

Clearly, the minimum of the function L cannot be solved algebraically in closed form, hence one of the iterative methods for function optimization is required. In this paper, the following scoring algorithm (see Lee and Jennrich (1979)) will be employed :

$$\underset{\sim}{\gamma}^{j+1} = \underset{\sim}{\gamma}^{j} - \delta\, \underset{\sim}{I}(\underset{\sim}{\gamma}^{j})^{-1}\, \dot{\underset{\sim}{L}}(\underset{\sim}{\gamma}^{j}) \tag{28}$$

where δ is a step-halving parameter which may be chosen as the first value in the sequence $\{1,\frac{1}{2},\frac{1}{4},\ldots\}$ that reduces L. $I(\underset{\sim}{\gamma}^{j})$ and $\dot{\underset{\sim}{L}}(\underset{\sim}{\gamma}^{j})$ are the information matrix and the gradient vector of L respectively evaluated at $\underset{\sim}{\gamma}^{j}$, the j^{th} step estimate of $\underset{\sim}{\gamma}$. The algorithm has many nice features that need to be mentioned here : (i) From (11) and (12), we see that only the first derivatives of the cell probabilities are required to implement the algorithm. (ii) Since $\underset{\sim}{I}(\underset{\sim}{\gamma})$ is positive definite, the algorithm is robust to the choice of starting values as it always produce an acceptable step. (iii) The asymptotic covariance matrix of $\hat{\underset{\sim}{\gamma}}$ is estimated by $I(\hat{\underset{\sim}{\gamma}}^{j})^{-1}$ which is automatically produced as a by-product at the last iteration of the algorithm.

§ 2.6 Tests of Hypothesis

The goodness of fit of the proposed model can be assessed by the likelihood ratio criterion (see, e.g. Bock, 1975). The test statistics is given by

$$\chi_A^2 = 2(\hat{L} - \hat{L}_0) \tag{29}$$

where $\hat{L} = L(\hat{\gamma})$, and $\hat{L}_0$ is the final function value obtained by minimizing $L(\gamma)$ without any covariance structure imposed on $\Sigma^{(g)}$, but subject to the same identification constraints (see Poon & Lee, 1987). Under the null hypothesis, the asymptotic distribution of χ_A^2 is central chi-square with degrees of freedom equal to the difference between the number of unknown parameters in the multivariate polychoric model and in the proposed model. The proposed model is rejected if χ_A^2 is larger than the corresponding chi-square tabled value.

Once the proposed model is not rejected, various null hypothesis concerning the covariance structures across groups can be tested in terms of appropriate equality constraints. For instance, consider the simultaneous confirmatory factor analysis model proposed in Jöreskog (1971), we would like to test : (a) the invariance of factor loadings across groups; (b) the covariance matrices of the factors across groups are equal; (c) the covariance matrices of the error measurements across groups are equal; and/or combinations of (a), (b) and (c), etc. This is done by estimating the model subject to the interesting constraints and compares its function value with the basic function value without the constraints.

More explicitly, let $\hat{L}_1$ and $\hat{L}_2$ be the function values obtained with and without the constraints. Then the likelihood ratio test statistic for the null hypothesis is given by

$$\chi_B^2 = 2(\hat{L}_1 - \hat{L}_2) \tag{30}$$

Similar to the previous argument, under the null hypothesis, the asymptotic distribution of χ_B^2 is central chi-square with degrees of freedom equal to the number of independent restrictions on the unknown parameters specified by the constraints.

§ 2.7 Example

To illustrate the theory developed so far, a computer program written in FORTRAN IV with double precision has been implemented to obtain the ML estimate, $\hat{\underset{\sim}{\gamma}}$. The subroutine developed by Schervish (1984) was employed to compute the distribution function of the multivariate normal distribution. The following simultaneous confirmatory factor analysis model proposed in Jöreskog (1971) is used :

$$\underset{\sim}{\Sigma}^{(g)} = \underset{\sim}{F}^{(g)}\underset{\sim}{M}^{(g)}\underset{\sim}{F}^{(g)\prime} + \underset{\sim}{E}^{(g)} \tag{31}$$

where $\underset{\sim}{F}^{(g)}$ is the factor loading matrix, $\underset{\sim}{M}^{(g)}$ and $\underset{\sim}{E}^{(g)}$ are the covariance matrices of the factors and error measurements respectively. The parameter vector $\underset{\sim}{\theta}^{(g)}$ consists of all independent unknown parameters in $\underset{\sim}{F}^{(g)}$, $\underset{\sim}{M}^{(g)}$ and $\underset{\sim}{E}^{(g)}$. Expressions for $\partial\underset{\sim}{\Sigma}^{(g)}/\partial\underset{\sim}{\theta}^{(g)}$ are available in Jöreskog (1971) and will not be reported here.

Artificial random samples from two populations with multivariate normal distribution $N[\underset{\sim}{0}, \underset{\sim}{\Sigma}^{(1)}]$ and $N[\underset{\sim}{0}, \underset{\sim}{\Sigma}^{(2)}]$ are generated. Here the population values of the covariance structures $\underset{\sim}{\Sigma}^{(g)}$ are given by

$$\underset{\sim}{F}^{(1)} = \begin{bmatrix} 0.7 & 0.7 & 0 & 0 \\ 0 & 0 & 0.7 & 0.7 \end{bmatrix}' , \quad \underset{\sim}{M}^{(1)} = \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix} , \quad \underset{\sim}{E}^{(1)} = 0.5\underset{\sim}{I}$$

$$\underset{\sim}{F}^{(2)} = \begin{bmatrix} 0.6 & 0.6 & 0 & 0 \\ 0 & 0 & 0.6 & 0.6 \end{bmatrix}' , \quad \underset{\sim}{M}^{(2)} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} , \quad \underset{\sim}{E}^{(2)} = 0.4\underset{\sim}{I} \tag{32}$$

where parameters with values 0 and 1 are considered as fixed parameters in the analysis. The multivariate samples $\left\{ \underset{\sim}{Y}_j^{(1)} \right\}$ and $\left\{ \underset{\sim}{Y}_j^{(2)} \right\}$ with size N_1=700 and N_2=500 were generated and then transformed to $\left\{ \underset{\sim}{Z}_j^{(1)} \right\}$ and $\left\{ \underset{\sim}{Z}_j^{(2)} \right\}$ using thresholds :

$$\begin{aligned} \underset{\sim}{\alpha}_i^{(1)} &= (-\infty, -0.5, 1.0, \infty) \\ \underset{\sim}{\alpha}_i^{(2)} &= (-\infty, -0.5, 1.0, \infty) \end{aligned} \qquad i = 1,2,3,4 \tag{33}$$

Based on the values of $\left\{ \underset{\sim}{Z}_j^{(g)} \right\}$, two independent 4-way contingency tables are constructed for analysis. Recall that for identification purposes, we need to fix $\alpha_{1,2}^{(1)}$, $\alpha_{1,3}^{(1)}$, $\alpha_{2,2}^{(1)}$, $\alpha_{3,2}^{(1)}$ and $\alpha_{4,2}^{(1)}$. These fixed values are taken to be the partition maximum likelihood (PML) estimates of the first sample $\left\{ \underset{\sim}{Z}_j^{(1)} \right\}$, see Poon & Lee (1987). In addition, the thresholds of the two groups are assumed to be equal during the analysis, i.e. a_i=1. The starting values of the various unknown parameters are arbitrary. For convenience, they are taken to be their respective population values. The full simultaneous ML estimates and their standard error estimates are presented in Table 1. The minimum function value $\hat{L}$ is equal to 4433.46 with 21 free parameters (3 thresholds and 18 free structural parameters). The minimum function value for the basic multivariate polychoric model, $\hat{L}_0$ is obtained, using a program similar to that in Poon & Lee (1987), to be 4433.03 with 23 free parameters (3 thresholds and 20 polychoric variances and covariances). According to (29), the goodness-of-fit test statistic, χ_A^2, is obtained to be 0.86 with 2 degrees of freedom. Hence, as expected, the null hypothesis that the sample data fit the proposed model is not rejected. Subsequently, we would like to test the conformity of the following exact equality constraints : (a) $\underset{\sim}{F}^{(1)} = \underset{\sim}{F}^{(2)}$, (b) $\underset{\sim}{M}^{(1)} = \underset{\sim}{M}^{(2)}$, (c) $\underset{\sim}{E}^{(1)} = \underset{\sim}{E}^{(2)}$ and (d) $\underset{\sim}{F}^{(1)} = \underset{\sim}{F}^{(2)}$, $\underset{\sim}{M}^{(1)} = \underset{\sim}{M}^{(2)}$, and $\underset{\sim}{E}^{(1)} = \underset{\sim}{E}^{(2)}$. To accomplish this task, we first consider the estimation of the general model incorporated with the corresponding constraints. The constrained estimates and their standard error estimates are reported in Table 2 to Table 5. The minimum function values obtained are respectively 4439.62, 4435.83, 4441.711 and 4453.26. Accordingly, the values of the test statistic χ_B^2 , are obtained via (30) to be 12.31, 4.74, 16.50 and 39.60 with degrees of freedom 4, 1, 4 and 9 respectively. Using 1% significance level, constraints (a) and (b) are not rejected while constraints (c) and (d) are rejected. These conclusions seem quite consistent to the pre-assigned population values.

Chapter 3

Bayesian Analysis of Stochastic Prior Information

§ 3.1 Introduction

The provision of auxiliary prior information plays an important role in the analysis of structural equation model. In this chapter, we will incorporate this kind of information in the analysis of the general model proposed in Chapter 2.

Stochastic prior information usually presents in the form of stochastic constraints defined by

$$\underset{\sim}{\mu} = \underset{\sim}{h}(\underset{\sim}{\gamma}) + \underset{\sim}{\epsilon} \tag{34}$$

where $\underset{\sim}{\mu}$ is an n by 1 observed vector, $\underset{\sim}{h}(\underset{\sim}{\gamma})$ is an n by 1 vector of differentiable functions of the parameter vector $\underset{\sim}{\gamma}$, and $\underset{\sim}{\epsilon}$ is an n by 1 random vector of error components with distribution $N[\underset{\sim}{0},\underset{\sim}{\Gamma}]$. Clearly, when $\underset{\sim}{\Gamma}$ equals to a zero matrix, $\underset{\sim}{\epsilon}$ degenerates to $\underset{\sim}{0}$ and the stochastic constraints become exact constraints. In this context, the stochastic constraints give more flexibility in studying functional relationships among the unknown parameters in $\underset{\sim}{\gamma}$. Lee (1988b) have developed an estimation procedure for the general covariance structure model in the presence of stochastic constraints by Bayesian approach. However, all the variates in the model are assumed to be continuous. Here, we will extend his work to handle polytomous variables as well. It is worth to note that we will still impose the same identification conditions specified in § 2.3 throughout the analysis.

§ 3.2 Bayesian Analysis of the Model

Now, we have the overall vector of the observed frequency counts, $\underset{\sim}{f}$ and stochastic prior information specified by (34). The manifestation of $\underset{\sim}{\mu}$ is either from previous study or from introspection. $\underset{\sim}{\Gamma}$ is considered to be an unknown nuisance parameter matrix. To simplify matters further, we assume $\underset{\sim}{f}$ to be independent of $\underset{\sim}{\epsilon}$.

To represent prior ignorance of $\underset{\sim}{\gamma}$, we assume the density function of $\underset{\sim}{\gamma}$, $p(\underset{\sim}{\gamma})$ to be a constant, see e.g. Jeffrey (1961) and Zellner (1971). Given observations on $\underset{\sim}{f}$ and $\underset{\sim}{\mu}$, the joint posterior density of $\underset{\sim}{\gamma}$ and $\underset{\sim}{\Gamma}$ is

$$p(\underset{\sim}{\gamma}, \underset{\sim}{\Gamma} \mid \underset{\sim}{f}, \underset{\sim}{\mu}) \propto p(\underset{\sim}{f} \mid \underset{\sim}{\gamma}) p(\underset{\sim}{\mu} \mid \underset{\sim}{\gamma}, \underset{\sim}{\Gamma}) p(\underset{\sim}{\Gamma}) \tag{35}$$

where $p(\underset{\sim}{f} \mid \underset{\sim}{\gamma}) \propto \prod_{g=1}^{G} \prod_{k(1)=1}^{m(1)} \cdots \prod_{k(p)=1}^{m(p)} \left[\xi_{\underset{\sim}{k}}^{(g)}\right]^{f_{\underset{\sim}{k}}^{(g)}}$,

$p(\underset{\sim}{\mu} \mid \underset{\sim}{\gamma}, \underset{\sim}{\Gamma}) \propto |\underset{\sim}{\Gamma}|^{-\frac{1}{2}} \exp\left\{- \frac{1}{2}(\underset{\sim}{\mu} - \underset{\sim}{h})' \underset{\sim}{\Gamma}^{-1} (\underset{\sim}{\mu} - \underset{\sim}{h})\right\}$, and

$p(\underset{\sim}{\Gamma})$ is the density of $\underset{\sim}{\Gamma}$.

Thus, the joint posterior density $p(\underset{\sim}{\gamma}, \underset{\sim}{\Gamma} \mid \underset{\sim}{f}, \underset{\sim}{\mu})$ is proportional to

$$\prod_{g=1}^{G} \prod_{k(1)=1}^{m(1)} \cdots \prod_{k(p)=1}^{m(p)} \left[\xi_{\underset{\sim}{k}}^{(g)}\right]^{f_{\underset{\sim}{k}}^{(g)}} |\underset{\sim}{\Gamma}|^{-\frac{1}{2}} \exp\left\{- \frac{1}{2}(\underset{\sim}{\mu} - \underset{\sim}{h})' \underset{\sim}{\Gamma}^{-1} (\underset{\sim}{\mu} - \underset{\sim}{h})\right\} p(\underset{\sim}{\Gamma}) \, . \tag{36}$$

Next, we consider analysis for three types of structure for $\underset{\sim}{\Gamma}$:

Case I : $\underset{\sim}{\Gamma} = \sigma^2 \underset{\sim}{I}$. It is the simpliest case where the error components of the stochastic constraints, $\underset{\sim}{\epsilon}$, are independent and with the same variance

σ^2. Here, we have only one nuisance parameter σ^2. To specify the prior distribution of σ^2, we will use the appropriate conjugate family (Raiffa & Schlaifer (1961)), namely, the inverse χ^2 family. This conjugate family involves two parameters and is sufficiently flexible for most applications (Lindley & Smith (1972), and Lee (1981)). Therefore, for given prior constants ν and β, we assume that $\nu\beta/\sigma^2$ is distributed as χ^2_ν and so

$$p(\sigma^2 \mid \nu, \beta) \propto (\sigma^2)^{(\nu+2)/2} \exp\{-\nu\beta/2\sigma^2\} \quad . \tag{37}$$

Form (36) and (37), we find that $p(\underset{\sim}{\gamma}, \underset{\sim}{\Gamma} \mid \underset{\sim}{f}, \underset{\sim}{\mu})$ is proportional to

$$\prod_{g=1}^{G} \prod_{k(1)=1}^{m(1)} \cdots \prod_{k(p)=1}^{m(p)} \left[\xi_{\underset{\sim}{k}}^{(g)}\right]^{f_{\underset{\sim}{k}}^{(g)}} (\sigma^2)^{-(\nu+n+2)/2} \exp\left[-\frac{1}{2\sigma^2}\left\{\sum_{i=1}^{n} (\mu_i - h_i)^2 + \nu\beta\right\}\right] . \tag{38}$$

It can be shown that

$$\int_0^\infty (\sigma^2)^{-(\nu+n+2)/2} \exp\left[-\frac{1}{2\sigma^2}\left\{\sum_{i=1}^{n} (\mu_i - h_i)^2 + \nu\beta\right\}\right] d\sigma^2$$
$$\propto \left\{\sum_{i=1}^{n} (\mu_i - h_i)^2 + \nu\beta\right\}^{-(\nu+n)/2} . \tag{39}$$

Thus, the nuisance parameter σ^2 can be eliminated by integration and so the posterior density of $\underset{\sim}{\gamma}$ is given by

$$p(\underset{\sim}{\gamma} \mid \underset{\sim}{f}, \underset{\sim}{\mu}) \propto \left\{\prod_{g=1}^{G} \prod_{k(1)=1}^{m(1)} \cdots \prod_{k(p)=1}^{m(p)} \left[\xi_{\underset{\sim}{k}}^{(g)}\right]^{f_{\underset{\sim}{k}}^{(g)}}\right\} \left\{\sum_{i=1}^{n} (\mu_i - h_i)^2 + \nu\beta\right\}^{-(\nu+n)/2} . \tag{40}$$

According to the usual procedure specified by Lindley & Smith (1972) and

Smith (1973), we define our Bayes estimate $\tilde{\gamma}$ of γ as the modal estimate of this posterior density, i.e. the value of γ that maximizes (40). Since log is an increasing function, $\tilde{\gamma}$ is obtained by minimizing the function

$$L_1(\gamma) = L(\gamma) + B_1(\gamma) \tag{41}$$

where $L(\gamma)$ is the negative of the usual log-likelihood function defined by (8), and

$$B_1(\gamma) = \frac{\nu + n}{2} \log\left\{\sum_{i=1}^{n} (\mu_i - h_i)^2 + \nu\beta\right\} \quad . \tag{42}$$

It should be pointed out that $\tilde{\gamma}$ is scarcely affected by the choice of ν and β , see Lindley & Smith (1972) and Lee (1981).

Case II : Γ is a diagonal matrix with diagonal elements $\sigma_i{}^2$. This means that the error components ϵ in (34) are independent but with different variances $\sigma_i{}^2$. In this case, we have n nuisance parameters, namely, $\sigma_1{}^2$, $\sigma_2{}^2, \ldots,$ $\sigma_n{}^2$. Again we use the conjugate family to specify the prior distribution of $\sigma_i{}^2$. Thus, for given prior constants ν_i and β_i, we assume that $\nu_i\beta_i/\sigma_i{}^2$ is independently distributed as $\chi^2_{\nu_i}$. The density function $p(\sigma_i{}^2 \mid \nu_i, \beta_i)$ is similar to that in (37) except that the quantities σ^2, ν and β are now replaced by $\sigma_i{}^2$, ν_i and β_i respectively. Thus, the joint posterior density of γ and Γ is given by

$$p(\gamma, \sigma_1{}^2, \ldots, \sigma_n{}^2 \mid f, \mu) \propto \left\{\prod_{g=1}^{G} \prod_{k(1)=1}^{m(1)} \cdots \prod_{k(p)=1}^{m(p)} \left[\xi_k^{(g)}\right]^{f_k^{(g)}}\right\} \times$$
$$\prod_{i=1}^{n} (\sigma_i{}^2)^{-(\nu_i+3)/2} \exp\left[\frac{-1}{2\sigma_i{}^2}\left\{(\mu_i - h_i)^2 + \nu_i\beta_i\right\}\right] \quad . \tag{43}$$

Similarly, the nuisance parameters σ_i^2, $i = 1,\ldots,n$, can be eliminated by integration and finally the posterior density of $\underset{\sim}{\gamma}$ is given by

$$p(\underset{\sim}{\gamma} \mid \underset{\sim}{f}, \underset{\sim}{\mu}) \propto \left\{\prod_{g=1}^{G}\prod_{k(1)=1}^{m(1)}\cdots\prod_{k(p)=1}^{m(p)}\left[\xi_{\underset{\sim}{k}}^{(g)}\right]^{f_{\underset{\sim}{k}}^{(g)}}\right\}\prod_{i=1}^{n}\left\{(\mu_i - h_i)^2 + \nu_i\beta_i\right\}^{-(\nu_i+1)/2} . \tag{44}$$

As before, the Bayes estimate $\tilde{\underset{\sim}{\gamma}}$ is obtained by minimizing the function

$$L_2(\underset{\sim}{\gamma}) = L(\underset{\sim}{\gamma}) + B_2(\underset{\sim}{\gamma}) \tag{45}$$

where

$$B_2(\underset{\sim}{\gamma}) = \sum_{i=1}^{n}\frac{\nu_i + 1}{2}\log\left\{(\mu_i - h_i)^2 + \nu_i\beta_i\right\} . \tag{46}$$

<u>Case III</u> : $\underset{\sim}{\Gamma}$ is a general positive definite matrix. In this case, $\underset{\sim}{\Gamma}^{-1}$ is assumed to have an independent Wishart distribution with known positive definite $\underset{\sim}{R}$ and known degrees of freedom ρ (see, e.g. Zellner (1971), Lindley & Smith (1972) and Lee (1981)). Thus

$$p(\underset{\sim}{\Gamma}) \propto |\underset{\sim}{\Gamma}|^{-(\rho+n+1)/2}\exp\left\{-\frac{1}{2}\mathrm{tr}\,\underset{\sim}{R}\,\underset{\sim}{\Gamma}^{-1}\right\} \tag{47}$$

and hence $p(\underset{\sim}{\gamma}, \underset{\sim}{\Gamma} \mid \underset{\sim}{f}, \underset{\sim}{\mu})$ is proportional to

$$\left\{\prod_{g=1}^{G}\prod_{k(1)=1}^{m(1)}\cdots\prod_{k(p)=1}^{m(p)}\left[\xi_{\underset{\sim}{k}}^{(g)}\right]^{f_{\underset{\sim}{k}}^{(g)}}\right\}|\underset{\sim}{\Gamma}|^{-(\rho+n+2)/2}\exp\left\{-\frac{1}{2}\mathrm{tr}(\underset{\sim}{A} + \underset{\sim}{R})\underset{\sim}{\Gamma}^{-1}\right\} \tag{48}$$

where $\underset{\sim}{A} = (\underset{\sim}{\mu} - \underset{\sim}{h})(\underset{\sim}{\mu} - \underset{\sim}{h})'$. It can be shown that

$$\int |\underset{\sim}{\Gamma}|^{-(\rho+n+2)/2} \exp\left\{- \tfrac{1}{2}\mathrm{tr}(\underset{\sim}{A} + \underset{\sim}{R})\underset{\sim}{\Gamma}^{-1}\right\} d\underset{\sim}{\Gamma} \;\alpha\; |\underset{\sim}{A} + \underset{\sim}{R}|^{-(\rho+1)/2} \quad . \tag{49}$$

Therefore, the nuisance parameter $\underset{\sim}{\Gamma}$ can also be eliminated by integration and

$$p(\underset{\sim}{\gamma} \mid \underset{\sim}{f}, \underset{\sim}{\mu}) \;\alpha\; \left\{\prod_{g=1}^{G} \prod_{k(1)=1}^{m(1)} \cdots \prod_{k(p)=1}^{m(p)} \left[\xi_{\underset{\sim}{k}}^{(g)}\right]^{f_{\underset{\sim}{k}}^{(g)}}\right\} |\underset{\sim}{A} + \underset{\sim}{R}|^{-(\rho+1)/2} \quad . \tag{50}$$

Similarly, the Bayes estimate $\tilde{\underset{\sim}{\gamma}}$ is obtained by minimizing the function

$$L_3(\underset{\sim}{\gamma}) = L(\underset{\sim}{\gamma}) + B_3(\underset{\sim}{\gamma}) \tag{51}$$

where

$$B_3(\underset{\sim}{\gamma}) = \frac{\rho + 1}{2} \log|\underset{\sim}{A} + \underset{\sim}{R}| \quad . \tag{52}$$

Similar to Case I, the prior constants $\underset{\sim}{R}$ and ρ scarcely affect the analysis.

§ 3.3 Computational Procedure

As before, the modal estimates of $L_k(\underset{\sim}{\gamma})$, k=1,2,3; cannot be obtained in closed form. Therefore, we still employ the scoring algorithm in computing the solutions, i.e.

$$\underset{\sim}{\gamma}^{j+1} = \underset{\sim}{\gamma}^{j} - \delta\, \underset{\sim}{I}_k(\underset{\sim}{\gamma}^{j})^{-1}\, \dot{\underset{\sim}{L}}_k(\underset{\sim}{\gamma}^{j}) \qquad k = 1,2,3 \tag{53}$$

where

$$\dot{\underset{\sim}{L}}_k(\underset{\sim}{\gamma}) = \dot{\underset{\sim}{L}}(\underset{\sim}{\gamma}) + \dot{\underset{\sim}{B}}_k(\underset{\sim}{\gamma}) \quad , \qquad \underset{\sim}{I}_k(\underset{\sim}{\gamma}) = \underset{\sim}{I}(\underset{\sim}{\gamma}) + \underset{\sim}{B}_k(\underset{\sim}{\gamma}) \tag{54}$$

are the gradient vector and the information matrix of $L_k(\underset{\sim}{\gamma})$ with $\dot{\underset{\sim}{L}}(\underset{\sim}{\gamma})$ and $\underset{\sim}{I}(\underset{\sim}{\gamma})$ defined in (9) an (10); and

$$\dot{\mathbf{B}}_k(\boldsymbol{\gamma}) = \partial B_k(\boldsymbol{\gamma})/\partial\boldsymbol{\gamma} \quad , \quad \mathbf{B}_k(\boldsymbol{\gamma}) = \partial^2 B_k(\boldsymbol{\gamma})/\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma} \qquad (55)$$

are the first and second derivatives of $B_k(\boldsymbol{\gamma})$.

For completeness, expressions for $\dot{\mathbf{B}}_k(\boldsymbol{\gamma})$ and $\mathbf{B}_k(\boldsymbol{\gamma})$ are presented below (see Lee, 1988b) :

$$\dot{\mathbf{B}}_1(\boldsymbol{\gamma})_c = -(\nu + n)\left\{\sum_{i=1}^{n}(\mu_i - h_i)^2 + \nu\beta\right\}^{-1}\left\{\sum_{i=1}^{n}(\mu_i - h_i)\frac{\partial h_i}{\partial\gamma_c}\right\} \qquad (56)$$

$$\mathbf{B}_1(\boldsymbol{\gamma})_{cd} = (\nu + n)\left[\left\{\sum_{i=1}^{n}(\mu_i - h_i)^2 + \nu\beta\right\}^{-1} \text{x} \sum_{i=1}^{n}\left\{\frac{\partial h_i}{\partial\gamma_c}\frac{\partial h_i}{\partial\gamma_d} - (\mu_i - h_i)\frac{\partial^2 h_i}{\partial\gamma_c\partial\gamma_d}\right\}\right.$$
$$\left. - 2\left\{\sum_{i=1}^{n}(\mu_i - h_i)^2 + \nu\beta\right\}^{-2}\left\{\sum_{i=1}^{n}(\mu_i - h_i)\frac{\partial h_i}{\partial\gamma_c}\right\}\left\{\sum_{i=1}^{n}(\mu_i - h_i)\frac{\partial h_i}{\partial\gamma_d}\right\}\right] \qquad (57)$$

$$\dot{\mathbf{B}}_2(\boldsymbol{\gamma})_c = -\sum_{i=1}^{n}(\nu_i + 1)(\mu_i - h_i)\left\{(\mu_i - h_i)^2 + \nu_i\beta_i\right\}^{-1}\frac{\partial h_i}{\partial\gamma_c} \qquad (58)$$

$$\mathbf{B}_2(\boldsymbol{\gamma})_{cd} = -\sum_{i=1}^{n}(\nu_i + 1)\left[\left\{(\mu_i - h_i)^2 + \nu_i\beta_i\right\}^{-2}\left\{(\mu_i - h_i)^2 - \nu_i\beta_i\right\}\frac{\partial h_i}{\partial\gamma_c}\frac{\partial h_i}{\partial\gamma_d}\right.$$
$$\left. + \left\{(\mu_i - h_i)^2 + \nu_i\beta_i\right\}^{-1}(\mu_i - h_i)\frac{\partial^2 h_i}{\partial\gamma_c\partial\gamma_d}\right] \qquad (59)$$

$$\dot{\mathbf{B}}_3(\boldsymbol{\gamma})_c = -(\rho + 1)\frac{\partial\mathbf{h}}{\partial\gamma_c}\mathbf{A}^*(\boldsymbol{\mu} - \mathbf{h}) \qquad (60)$$

$$B_3(\gamma)_{cd} = -(\rho + 1)\left[\frac{\partial^2 h}{\partial\gamma_c \partial\gamma_d} A^*(\mu - h) + \right.$$

$$\left. \frac{\partial h}{\partial\gamma_d} \left\{ A^* \otimes (\mu - h)' A^*(\mu - h) + (\mu - h)' A^* \otimes A^*(\mu - h) - A^* \right\} \frac{\partial h}{\partial\gamma_c} \right] \qquad (61)$$

where $A^* = (A + R)^{-1}$.

§ 3.4 Test the Compatibility of the Prior Information

In this section, we will develop theory to test the compatibility of the prior information to the sample information. The null hypothesis is given by

$$H_0 : \mu = h(\gamma) + \epsilon \qquad (62)$$

where $\epsilon \sim N[0, \Gamma]$ for some Γ.

Suppose γ is identified in the general hypothesis parameter space and $\hat{\gamma}$ is the corresponding unconstrained maximum likelihood estimate. Under mild regularity conditions, we have

$$(\hat{\gamma} - \gamma) \xrightarrow{L} N\left[0, I(\gamma)^{-1} \right] \qquad (63)$$

where N is the overall sample size, i.e. $N = N_1 + \dots + N_G$; '$\xrightarrow{L}$' denotes convergent in distribution and $I(\gamma)$ is the usual information matrix defined in (10). Then, by Delta's theorem, we have

$$\left[h(\hat{\gamma}) - h(\gamma) \right] \xrightarrow{L} N[\, 0, \Omega \,] \qquad (64)$$

where $\Omega = (\partial h/\partial\gamma)\ I(\gamma)^{-1}(\partial h/\partial\gamma)'$. Hence, $\left[h(\hat{\gamma}) - \mu \right]$ will converge in distribution to $N[\, 0, \Gamma + \Omega \,]$, and the Wald's type test statistics

$$\chi_C^2 = \left[h(\hat{\gamma}) - \mu \right]' (\Gamma + \Omega)^{-1} \left[h(\hat{\gamma}) - \mu \right] \tag{65}$$

will converge in distribution to a chi-square distribution with n degrees of freedom, see Lee (1988a). To apply this test in practice, Ω is replaced by $\hat{\Omega} = \Omega(\hat{\gamma})$ and also we have to know Γ. Naturally, Γ can be specified by the null hypothesis. This means that the null hypothesis not only specifies the stochastic functional relationships among the various parameters but also their precision by giving known values for Γ. Otherwise, a consistent estimate of Γ is required. It is worth to note that if we specify $\Gamma = 0$, then the stochastic constraints in (34) will become exact equality constraints and the χ_C^2 reduces back to the test statistic given by Lee (1985).

§ 3.5 Example

Consider again the simultaneous confirmatory factor analysis model in § 2.7 with population values given by (32). Using the same contingency tables simulated in § 2.7, we analyze the model separately with the following sets of stochastic constraints :

(a) $F^{(2)} = F^{(1)} + \epsilon$;

(b) $M^{(2)} = M^{(1)} + \epsilon$;

(c) $E^{(2)} = E^{(1)} + \epsilon$; and

(d) $F^{(2)} = F^{(1)} + \epsilon_1$, $M^{(2)} = M^{(1)} + \epsilon_2$ and $E^{(2)} = E^{(1)} + \epsilon_3$ with $\epsilon = (\epsilon_1' , \epsilon_2' , \epsilon_3')'$;

In all cases, ϵ is assumed to be multivariately normally distributed with mean $\underset{\sim}{0}$ and covariance matrix $\underset{\sim}{\Gamma}$.

A modified computer program has been written to find the Bayes estimates $\tilde{\underset{\sim}{\gamma}}$. To implement the procedure, $\underset{\sim}{\Gamma}$ is assumed to be $\sigma^2\underset{\sim}{I}$ (Case I) and the prior constants ν and β are taken to be 1.0 and 0.1 respectively. As before, for identification purposes, we fix $\alpha_{12}^{(1)}$, $\alpha_{13}^{(1)}$, $\alpha_{22}^{(1)}$, $\alpha_{32}^{(1)}$ and $\alpha_{42}^{(1)}$ to their partition maximum likelihood (PML) estimates and assume the thresholds of both groups to be equal. The program converges nicely in a few iterations with the starting values of the parameters taken to be their corresponding population values. The Bayes estimates and their various standard error estimates are presented in Table 6 to Table 9. From the tables, we observe that the various Bayes estimates do not differ too much from the unconstrained maximum likelihood estimates. On the other hand, they are quite different from the corresponding constrained maximum likelihood estimates presented in Table 2 to Table 5. Lastly, we would like to test the compatibility of the stochastic prior information to the sample data using the Wald's type test statistic, χ_C^2, given in (65). The results of the tests with 5% significance level are reported in Table 10. From the table, we observe that only constraint (a) are compatible to the sample data regardless to the values of σ^2 chosen. Meanwhile, the conclusions of the tests concerning other constraints depend solely on the values of σ^2 chosen. As a rule of thumb, the smaller the value of σ^2, the more likely that the tests will be rejected.

Chapter 4

Simulation Study

§ 4.1 Introduction

In this chapter, a number of simulations have been implemented to illustrate the behavior of Bayes estimates derived in Chapter 3 and compare them with the full maximum likelihood estimates discussed in Chapter 2. Since in practice it is more realistic to assume the error components of the stochastic constraints are independent, we only consider Case I and II in our study.

§ 4.2 Simulation 1

The first part of the study is to investigate the usefulness of stochastic prior information in covariance structure analysis with polytomous variables in several groups. The simulation is based on the simultaneous confirmatory factor analysis model defined in (31) with the following population values for $\underset{\sim}{F}^{(g)}$, $\underset{\sim}{M}^{(g)}$, and $\underset{\sim}{E}^{(g)}$:

$$\begin{aligned} \underset{\sim}{F}^{(1)} &= \begin{bmatrix} 1.0 & 0 & 0.5 \\ 0 & 1.0 & 0.5 \end{bmatrix}', \quad \underset{\sim}{M}^{(1)} = \begin{bmatrix} 1.0 & 0.3 \\ 0.3 & 1.0 \end{bmatrix}, \quad \underset{\sim}{E}^{(1)} = 0.5\underset{\sim}{I} \\ \underset{\sim}{F}^{(2)} &= \begin{bmatrix} 1.0 & 0 & 0.5 \\ 0 & 1.0 & 0.5 \end{bmatrix}', \quad \underset{\sim}{M}^{(2)} = \begin{bmatrix} 1.0 & 0.4 \\ 0.4 & 1.0 \end{bmatrix}, \quad \underset{\sim}{E}^{(2)} = 0.6\underset{\sim}{I} \end{aligned} \tag{66}$$

where parameters with values 0 or 1 are considered as fixed parameters throughout the analysis. The population covariance matrices $\underset{\sim}{\Sigma}^{(g)}$ were computed from $\underset{\sim}{F}^{(g)}$, $\underset{\sim}{M}^{(g)}$, and $\underset{\sim}{E}^{(g)}$ according to (31). Multivariate random samples $\left\{\underset{\sim}{Y}_j^{(g)}\right\}$ of size N_g were generated and then transformed to $\left\{\underset{\sim}{Z}_j^{(g)}\right\}$

using the following thresholds :

$$\underset{\sim}{\alpha}_i^{(1)} = (-\infty, -0.5, 1.0, \infty)$$

$$i = 1, 2, 3$$

$$\underset{\sim}{\alpha}_i^{(2)} = (-\infty, -0.5, 1.0, \infty) \ .$$

Following the similar procedure described in § 2.7, two independent 3-way contingency tables are constructed for latter analysis. To simplify matter further, we deliberately fix the thresholds to their population values during estimation. Afterwards, we analyze the data by the following methods :

(i) ML1 : Full maximum likelihood approach;

(ii) ML2 : Full maximum likelihood approach with additional exact-equality constraints,

$$F_{31}^{(1)} = F_{31}^{(2)} \quad \text{and} \quad F_{32}^{(1)} = F_{32}^{(2)} \tag{67}$$

(iii) BAY1 : Bayesian approach with stochastic constraints

$$F_{31}^{(2)} = F_{31}^{(1)} + \epsilon_1 \quad \text{and} \quad F_{32}^{(2)} = F_{32}^{(1)} + \epsilon_2 \tag{68}$$

where $\underset{\sim}{\epsilon} = (\epsilon_1, \epsilon_2)'$ is distributed as $N[\underset{\sim}{0}, \underset{\sim}{\Gamma}]$ with $\underset{\sim}{\Gamma}$ taken to be $\sigma^2 \underset{\sim}{I}$ and prior constants $\nu = 5.0$, $\beta = 0.1$;

(iv) BAY2 : same as (iii), but with $\nu = 1.0$;

(v) BAY3 : same as (iii), but the covariance matrix of $\underset{\sim}{\epsilon}$ is taken to be diagonal and $\nu_1 = \nu_2 = 5.0$, $\beta_1 = \beta_2 = 0.1$; and

(vi) BAY4 : same as (v), but with $\nu_1 = \nu_2 = 1.0$.

Apart from the different estimation methods, we are also interested to examine the effect of the sample sizes on the performance of the various estimates. Therefore, two sets of sample sizes are chosen, they are respectively N_1=100, N_2=150 and N_1=200, N_2=300. Throughout the process of study, for every chosen sample sizes and estimation methods, we encountered some generated data that gave rise to improper solutions for the unique variance estimates. In literature, these solutions are called Heywood cases. For comparison sake, these improper cases were deleted and the simulation continued until we have completed 50 cases. The root mean squares errors between the various estimates and their corresponding population values are given by

$$RMS_i = \left\{ \sum_{j=1}^{50} (\hat{\gamma}_{ij} - \gamma_i)^2 / 50 \right\}^{\frac{1}{2}} \qquad i = 1,\ldots,q \tag{69}$$

where q is the total number of free parameters in the model; and $\hat{\gamma}_{ij}$ is the estimate of γ_i for the j^{th} case. In addition to the root mean squares errors, the sample means and the sample standard deviations of the estimates are also very useful measures of the performance of the various estimation methods. They are defined by

$$\bar{\gamma}_i = \sum_{j=1}^{50} \hat{\gamma}_{ij} / 50 \tag{70}$$

$$S_i = \left\{ \sum_{j=1}^{50} (\hat{\gamma}_{ij} - \bar{\gamma}_i)^2 / 49 \right\}^{\frac{1}{2}} . \tag{71}$$

The results of the simulation are respectively reported in Table 11 to 16. After examining the tables carefully, we have the following observations :

1. For both sample sizes, the various Bayes estimates are comparably better than the ML1 estimates, especially for estimates of $F_{31}^{(1)}$, $F_{32}^{(1)}$, $F_{31}^{(2)}$ and $F_{32}^{(2)}$ and when the sample sizes are small.

2. As expected, for both sample sizes, the ML2 estimates of $F_{31}^{(1)}$, $F_{32}^{(1)}$, $F_{31}^{(2)}$ and $F_{32}^{(2)}$ are the best among the other estimation methods. However, it is interesting to note that the majority of other ML2 estimates are slightly worse than that of the various Bayes estimates, though the differences are quite minor.

3. For N_1=200 and N_2=300, the performance of the various Bayes estimates are quite similar to each other. On the other hand, for smaller sample sizes N_1=100 and N_2=150, the Bayes estimates for Case I (BAY1 and BAY2) are superior than that for Case II (BAY3 and BAY4).

4. For both sample sizes, it is worth to note that the choice of prior constants ν and β for Case I or ν_i and β_i for Case II scarcely affect the overall performance of the Bayes estimates. In fact, I have checked some individual estimates and found that they are very close to each other.

§ 4.3 Simulation 2

The second part of the study is also based on the confirmatory factor analysis model with population values given by (66), except now $F_{31}^{(2)}$ and $F_{32}^{(2)}$ are taken to be 0.3 and 0.6 respectively. Therefore, in ML2, the exact equality constraints in (67) are incorrectly specified. Furthermore, for the various Bayes estimates, the stochastic constraints in (68) are not exactly in accordance with the population values. During the simulation, we still encountered Heywood cases in obtaining the various estimates for both sets of sample sizes. Similarly, they were deleted and the simulation continued until 50 cases have been completed. The results obtained are summarized in Table 17 to Table 22. From these tables, we have the following observations :

1. As expected, with both sample sizes, the ML1 estimates are comparably better than the ML2 estimates which fixed some parameters at incorrect values. On the contrary, it is surprising to find that the various Bayes estimates are better than the ML1 estimates.

2. Similar to simulation 1, the BAY1 and BAY2 estimates are constantly superior to the BAY3 and BAY4 estimates, especially for the estimates of $F_{31}^{(1)}$, $F_{32}^{(1)}$, $F_{31}^{(2)}$ and $F_{32}^{(2)}$ and when small sample sizes are used.

3. As before, the choice of prior constants ν and β or ν_i and β_i has negligible effect on the various Bayes estimates.

§ 4.4 Summary and Discussion

Since both estimation procedures demand heavy computational efforts, we only consider the popular simultaneous confirmatory factor analysis model in our simulation study. For the same reason, we let the dimension of the polytomous vector to be 3. Looking at the population values of the chosen model, it seems quite restrictive and artificial. Models of higher dimensions are more interesting as well as more useful in practice. Nevertheless, the basic conclusions are unaffected and deserve discussion here.

First of all, the various Bayes estimates seem superior to the unconstrained maximum likelihood estimates (ML1). The superiority of the former to the latter holds even when the stochastic constraints are not in accordance with the population values. Besides, as compared to the constrained maximum likelihood estimates, the ML2 performs better only when the exact constraints are correctly specified. On the contrary, when the exact constraints are incorrectly specified, the Bayesian approach produces estimates that are far better than that produced by ML2. Therefore, unless you have strong confidence on the plausibility of the exact equality constraints, it is wise to impose the less restrictive stochastic counterparts.

Finally, it is interesting to note that in obtaining the ML1 and ML2 estimates, we frequently encountered improper Heywood cases, especially when the sample sizes are small. On the other hand, the phenomenon is quite rare in deriving the various Bayes estimates. Thus, special attention should be paid in applying the maximum likelihood method.

Chapter 5

Concluding Remarks

In this thesis, stochastic prior information in the form of stochastic constraints on the parameters of the model has been introduced in the analysis of structural equation model with polytomous data in several groups. A method based on the Bayesian approach (see Lee, 1988b) has been developed to obtain the various Bayes estimates. Based on the results of the simulation study, it has been shown that the provision of stochastic prior information not only provides us more freedom in studying the functional relationship among the parameters in the model but also gives more accurate and reliable estimates generally.

Nevertheless, the method suffer a major drawback of computational inefficiency, especially when the dimension of the observed polytomous vector, p is large. It is because the procedure requires the evaluation of multiple integrals with complexity increases dramatically with p. As a result, the technique is practically infeasible for higher dimension of polytomous vector, say p=6 or more.

To remedy this deficiency, one obvious direction of future development is to apply the concept to a more efficient estimation procedure in the analysis of structural equation model for polytomous variables. As a typical example, the computationally efficient multi-stage estimation procedure described in Poon, Lee, Bentler and Afifi (1989) is clearly a possible candidate to entertain.

At last, it is worth to note that all the results developed here are based on the normality assumption of the latent random vector. Hence, if the underlying distribution is unknown or other than normal, the applicability of the procedures is suspected. Therefore, the problem of robustness of the various Bayes estimates may be an interesting research topic in the future.

Table 1 : Full Simultaneous Maximum Likelihood Estimates

Parameters	Group 1	Group 2
α_{12}	-0.512^{*}	
α_{13}	0.983^{*}	
α_{22}	-0.464^{*}	
α_{23}	0.839(0.070)	
α_{32}	-0.491^{*}	
α_{33}	1.080(0.103)	
α_{42}	-0.519^{*}	
α_{43}	1.102(0.106)	
F_{11}	0.882(0.125)	0.755(0.072)
F_{21}	0.552(0.083)	0.453(0.049)
F_{32}	0.761(0.102)	0.668(0.066)
F_{42}	0.814(0.106)	0.658(0.069)
M_{21}	0.302(0.064)	0.540(0.073)
E_{11}	0.225(0.234)	0.207(0.128)
E_{22}	0.572(0.152)	0.348(0.089)
E_{33}	0.693(0.255)	0.346(0.154)
E_{44}	0.528(0.254)	0.624(0.194)

Notes : 1. In all tables, asterisks denote parameter values fixed at those values.

2. Standard error estimates are in parentheses.

Table 2 : Full Simultaneous Maximum Likelihood Estimates with constraint (a) $\underset{\sim}{F}^{(1)} = \underset{\sim}{F}^{(2)}$

Parameters	Group 1		Group 2
α_{12}		-0.512^*	
α_{13}		0.983^*	
α_{22}		-0.464^*	
α_{23}		0.838(0.070)	
α_{32}		-0.491^*	
α_{33}		1.076(0.102)	
α_{42}		-0.519^*	
α_{43}		1.097(0.105)	
F_{11}		0.810(0.056)	
F_{21}		0.500(0.038)	
F_{32}		0.713(0.050)	
F_{42}		0.730(0.053)	
M_{21}	0.321(0.067)		0.540(0.061)
E_{11}	0.317(0.119)		0.159(0.116)
E_{22}	0.602(0.126)		0.324(0.090)
E_{33}	0.724(0.211)		0.311(0.151)
E_{44}	0.618(0.199)		0.559(0.196)

Table 3 : Full Simultaneous Maximum Likelihood Estimates with constraint (b) $\underset{\sim}{\mathbf{M}}^{(1)} = \underset{\sim}{\mathbf{M}}^{(2)}$

Parameters	Group 1		Group 2
a_{12}		-0.512^{*}	
a_{13}		0.983^{*}	
a_{22}		-0.464^{*}	
a_{23}		0.843(0.071)	
a_{32}		-0.491^{*}	
a_{33}		1.071(0.102)	
a_{42}		-0.519^{*}	
a_{43}		1.092(0.105)	
F_{11}	0.829(0.088)		0.836(0.102)
F_{21}	0.587(0.067)		0.408(0.058)
F_{32}	0.762(0.083)		0.700(0.083)
F_{42}	0.802(0.085)		0.623(0.081)
M_{21}		0.391(0.049)	
E_{11}	0.322(0.165)		0.064(0.182)
E_{22}	0.541(0.146)		0.388(0.091)
E_{33}	0.682(0.237)		0.282(0.168)
E_{44}	0.540(0.230)		0.648(0.196)

Table 4 : Full Simultaneous Maximum Likelihood Estimates with constraint (c) $\underset{\sim}{E}^{(1)} = \underset{\sim}{E}^{(2)}$

Parameters	Group 1		Group 2
a_{12}		-0.512^{*}	
a_{13}		0.983^{*}	
a_{22}		-0.464^{*}	
a_{23}		0.823(0.068)	
a_{32}		-0.491^{*}	
a_{33}		1.025(0.094)	
a_{42}		-0.519^{*}	
a_{43}		1.081(0.103)	
F_{11}	0.809(0.070)		0.762(0.067)
F_{21}	0.549(0.051)		0.467(0.049)
F_{32}	0.837(0.068)		0.637(0.060)
F_{42}	0.675(0.062)		0.673(0.064)
M_{21}	0.282(0.059)		0.577(0.073)
E_{11}		0.273(0.109)	
E_{22}		0.428(0.096)	
E_{33}		0.390(0.146)	
E_{44}		0.649(0.179)	

Table 5 : Full Simultaneous Maximum Likelihood Estimates with constraints (d) $\underset{\sim}{\mathbf{F}}^{(1)} = \underset{\sim}{\mathbf{F}}^{(2)}$, $\underset{\sim}{\mathbf{M}}^{(1)} = \underset{\sim}{\mathbf{M}}^{(2)}$ and $\underset{\sim}{\mathbf{E}}^{(1)} = \underset{\sim}{\mathbf{E}}^{(2)}$

<table>
<tr><th>Parameters</th><th>Group 1</th><th>Group 2</th></tr>
<tr><td>a_{12}</td><td colspan="2">-0.512*</td></tr>
<tr><td>a_{13}</td><td colspan="2">0.983*</td></tr>
<tr><td>a_{22}</td><td colspan="2">-0.464*</td></tr>
<tr><td>a_{23}</td><td colspan="2">0.849(0.072)</td></tr>
<tr><td>a_{32}</td><td colspan="2">-0.491*</td></tr>
<tr><td>a_{33}</td><td colspan="2">1.035(0.099)</td></tr>
<tr><td>a_{42}</td><td colspan="2">-0.519*</td></tr>
<tr><td>a_{43}</td><td colspan="2">1.080(0.104)</td></tr>
<tr><td>F_{11}</td><td colspan="2">0.820(0.071)</td></tr>
<tr><td>F_{21}</td><td colspan="2">0.512(0.048)</td></tr>
<tr><td>F_{32}</td><td colspan="2">0.702(0.059)</td></tr>
<tr><td>F_{42}</td><td colspan="2">0.730(0.061)</td></tr>
<tr><td>M_{21}</td><td colspan="2">0.385(0.048)</td></tr>
<tr><td>E_{11}</td><td colspan="2">0.226(0.127)</td></tr>
<tr><td>E_{22}</td><td colspan="2">0.470(0.106)</td></tr>
<tr><td>E_{33}</td><td colspan="2">0.492(0.167)</td></tr>
<tr><td>E_{44}</td><td colspan="2">0.569(0.186)</td></tr>
</table>

Table 6 : Bayes Estimates corresponding to stochastic constraint (a) $\underset{\sim}{F}^{(2)} = \underset{\sim}{F}^{(1)} + \underset{\sim}{\epsilon}$

Parameters	Group 1		Group 2
α_{12}		-0.512^{*}	
α_{13}		0.983^{*}	
α_{22}		-0.464^{*}	
α_{23}		0.839(0.056)	
α_{32}		-0.491^{*}	
α_{33}		1.079(0.067)	
α_{42}		-0.519^{*}	
α_{43}		1.101(0.067)	
F_{11}	0.862(0.071)		0.767(0.058)
F_{21}	0.557(0.053)		0.456(0.043)
F_{32}	0.764(0.070)		0.673(0.056)
F_{42}	0.794(0.071)		0.672(0.060)
M_{21}	0.307(0.057)		0.540(0.063)
E_{11}	0.257(0.122)		0.193(0.098)
E_{22}	0.565(0.102)		0.348(0.072)
E_{33}	0.683(0.130)		0.342(0.101)
E_{44}	0.554(0.128)		0.610(0.116)

Table 7 : Bayes Estimates corresponding to stochastic constraint (b) $\underset{\sim}{\mathbf{M}}^{(2)} = \underset{\sim}{\mathbf{M}}^{(1)} + \underset{\sim}{\epsilon}$

Parameters	Group 1	Group 2
a_{12}	-0.512*	
a_{13}	0.983*	
a_{22}	-0.464*	
a_{23}	0.839(0.063)	
a_{32}	-0.491*	
a_{33}	1.080(0.080)	
a_{42}	-0.519*	
a_{43}	1.103(0.080)	
F_{11}	0.871(0.087)	0.762(0.067)
F_{21}	0.559(0.062)	0.450(0.047)
F_{32}	0.763(0.078)	0.671(0.061)
F_{42}	0.813(0.081)	0.656(0.064)
M_{21}	0.315(0.061)	0.524(0.070)
E_{11}	0.246(0.157)	0.195(0.115)
E_{22}	0.565(0.122)	0.351(0.081)
E_{33}	0.691(0.173)	0.340(0.124)
E_{44}	0.533(0.170)	0.627(0.146)

Table 8 : Bayes Estimates corresponding to stochastic constraint (c) $\underset{\sim}{E}^{(2)} = \underset{\sim}{E}^{(1)} + \underset{\sim}{\epsilon}$

Parameters	Group 1	Group 2
α_{12}	-0.512^{*}	
α_{13}	0.983^{*}	
α_{22}	-0.464^{*}	
α_{23}	0.827(0.064)	
α_{32}	-0.491^{*}	
α_{33}	1.047(0.084)	
α_{42}	-0.519^{*}	
α_{43}	1.091(0.089)	
F_{11}	0.862(0.076)	0.758(0.062)
F_{21}	0.552(0.055)	0.451(0.044)
F_{32}	0.805(0.072)	0.643(0.057)
F_{42}	0.737(0.068)	0.669(0.061)
M_{21}	0.299(0.057)	0.547(0.066)
E_{11}	0.250(0.136)	0.211(0.109)
E_{22}	0.528(0.119)	0.353(0.084)
E_{33}	0.523(0.179)	0.371(0.133)
E_{44}	0.622(0.178)	0.604(0.167)

Table 9 : Bayes Estimates corresponding to stochastic constraints

(d) $\underset{\sim}{F}^{(2)} = \underset{\sim}{F}^{(1)} + \underset{\sim}{\epsilon}_1$, $\underset{\sim}{M}^{(2)} = \underset{\sim}{M}^{(1)} + \underset{\sim}{\epsilon}_2$ and $\underset{\sim}{E}^{(2)} = \underset{\sim}{E}^{(1)} + \underset{\sim}{\epsilon}_3$

Parameters	Group 1	Group 2
a_{12}	-0.512*	
a_{13}	0.983*	
a_{22}	-0.464*	
a_{23}	0.825(0.062)	
a_{32}	-0.491*	
a_{33}	1.032(0.079)	
a_{42}	-0.519*	
a_{43}	1.082(0.084)	
F_{11}	0.850(0.078)	0.778(0.070)
F_{21}	0.546(0.056)	0.451(0.047)
F_{32}	0.774(0.070)	0.653(0.060)
F_{42}	0.734(0.069)	0.675(0.065)
M_{21}	0.327(0.059)	0.505(0.067)
E_{11}	0.258(0.138)	0.190(0.122)
E_{22}	0.522(0.113)	0.357(0.085)
E_{33}	0.525(0.166)	0.358(0.133)
$E_{\$\$}$	0.609(0.171)	0.591(0.164)

Table 10 : Summary of the tests of the compatibility of the various stochastic constraints

σ^2	χ^2_H	d.f.	Conclusion
		(a) $\underset{\sim}{F}^{(2)} = \underset{\sim}{F}^{(1)} + \underset{\sim}{\epsilon}$	
0.01	5.86	4	Do not reject
0.05	1.17	4	Do not reject
		(b) $\underset{\sim}{M}^{(2)} = \underset{\sim}{M}^{(1)} + \underset{\sim}{\epsilon}$	
0.01	5.64	1	Reject
0.05	1.13	1	Do not reject
		(c) $\underset{\sim}{E}^{(2)} = \underset{\sim}{E}^{(1)} + \underset{\sim}{\epsilon}$	
0.01	17.89	4	Reject
0.05	3.60	4	Do not reject
(d) $\underset{\sim}{F}^{(2)} = \underset{\sim}{F}^{(1)} + \underset{\sim}{\epsilon}_1$, $\underset{\sim}{M}^{(2)} = \underset{\sim}{M}^{(1)} + \underset{\sim}{\epsilon}_2$ and $\underset{\sim}{E}^{(2)} = \underset{\sim}{E}^{(1)} + \underset{\sim}{\epsilon}_3$			
0.01	29.37	9	Reject
0.05	5.90	9	Do not reject

Table 11 : Root Mean Squares errors between the various estimates and the population values for simulation 1

	Parameters	RMS					
		ML1	ML2	BAY1	BAY2	BAY3	BAY4
	$F_{31}^{(1)}$	0.154	0.116	0.124	0.119	0.135	0.133
	$F_{32}^{(1)}$	0.162	0.128	0.137	0.134	0.132	0.132
	$M_{21}^{(1)}$	0.168	0.171	0.167	0.167	0.169	0.168
	$E_{11}^{(1)}$	0.193	0.190	0.190	0.189	0.193	0.193
	$E_{22}^{(1)}$	0.221	0.221	0.217	0.216	0.220	0.220
$N_1=100$	$E_{33}^{(1)}$	0.188	0.171	0.183	0.180	0.189	0.188
$N_2=150$	$F_{31}^{(2)}$	0.158	0.116	0.127	0.123	0.146	0.143
	$F_{32}^{(2)}$	0.159	0.128	0.131	0.129	0.131	0.130
	$M_{21}^{(2)}$	0.181	0.186	0.181	0.182	0.185	0.185
	$E_{11}^{(2)}$	0.289	0.274	0.276	0.273	0.285	0.284
	$E_{22}^{(2)}$	0.236	0.237	0.235	0.234	0.236	0.236
	$E_{33}^{(2)}$	0.233	0.229	0.226	0.224	0.233	0.232

Table 12 : Root Mean Squares errors between the various estimates and the population values for simulation 1

	Parameters	RMS					
		ML1	ML2	BAY1	BAY2	BAY3	BAY4
	$F_{31}^{(1)}$	0.100	0.085	0.090	0.087	0.099	0.097
	$F_{32}^{(1)}$	0.103	0.083	0.093	0.089	0.084	0.084
	$M_{21}^{(1)}$	0.130	0.123	0.129	0.128	0.130	0.130
	$E_{11}^{(1)}$	0.213	0.213	0.210	0.208	0.213	0.213
	$E_{22}^{(1)}$	0.176	0.177	0.173	0.172	0.175	0.176
$N_1=200$	$E_{33}^{(1)}$	0.139	0.124	0.135	0.133	0.136	0.136
$N_2=300$	$F_{31}^{(2)}$	0.139	0.085	0.107	0.102	0.117	0.114
	$F_{32}^{(2)}$	0.124	0.083	0.097	0.093	0.098	0.095
	$M_{21}^{(2)}$	0.136	0.138	0.132	0.132	0.133	0.133
	$E_{11}^{(2)}$	0.224	0.228	0.223	0.222	0.223	0.224
	$E_{22}^{(2)}$	0.177	0.174	0.174	0.173	0.175	0.175
	$E_{33}^{(2)}$	0.147	0.144	0.141	0.141	0.143	0.142

Table 13 : Sample means of the various estimates for simulation 1

	Parameters	Sample Mean					
		ML1	ML2	BAY1	BAY2	BAY3	BAY4
	$F_{31}^{(1)}$	0.434	0.461	0.448	0.451	0.437	0.439
	$F_{32}^{(1)}$	0.486	0.492	0.481	0.482	0.488	0.488
	$M_{21}^{(1)}$	0.285	0.274	0.281	0.280	0.281	0.281
	$E_{11}^{(1)}$	0.409	0.417	0.410	0.411	0.408	0.408
	$E_{22}^{(1)}$	0.472	0.478	0.470	0.470	0.473	0.473
$N_1=100$	$E_{33}^{(1)}$	0.537	0.520	0.537	0.534	0.540	0.539
$N_2=150$	$F_{31}^{(2)}$	0.504	0.461	0.483	0.478	0.495	0.493
	$F_{32}^{(2)}$	0.483	0.492	0.496	0.497	0.491	0.491
	$M_{21}^{(2)}$	0.377	0.392	0.378	0.380	0.376	0.377
	$E_{11}^{(2)}$	0.663	0.654	0.655	0.654	0.663	0.662
	$E_{22}^{(2)}$	0.550	0.550	0.550	0.549	0.550	0.550
	$E_{33}^{(2)}$	0.605	0.652	0.623	0.628	0.613	0.616

Table 14 : Sample means of the various estimates for simulation 1

	Parameters	Sample Mean					
		ML1	ML2	BAY1	BAY2	BAY3	BAY4
	$F_{31}^{(1)}$	0.462	0.469	0.464	0.465	0.458	0.459
	$F_{32}^{(1)}$	0.481	0.488	0.481	0.481	0.485	0.485
	$M_{21}^{(1)}$	0.316	0.313	0.315	0.314	0.316	0.316
	$E_{11}^{(1)}$	0.497	0.501	0.496	0.496	0.497	0.497
	$E_{22}^{(1)}$	0.501	0.505	0.500	0.500	0.502	0.502
N_1=200	$E_{33}^{(1)}$	0.526	0.516	0.527	0.526	0.527	0.527
N_2=300	$F_{31}^{(2)}$	0.473	0.469	0.474	0.473	0.474	0.475
	$F_{32}^{(2)}$	0.498	0.488	0.495	0.496	0.496	0.495
	$M_{21}^{(2)}$	0.415	0.421	0.414	0.414	0.414	0.414
	$E_{11}^{(2)}$	0.538	0.536	0.536	0.536	0.537	0.537
	$E_{22}^{(2)}$	0.624	0.622	0.623	0.622	0.623	0.623
	$E_{33}^{(2)}$	0.614	0.635	0.620	0.622	0.619	0.619

Table 15 : Standard deviation of the various estimates for simulation 1

	Parameters	Standard deviation					
		ML1	ML2	BAY1	BAY2	BAY3	BAY4
	$F_{31}^{(1)}$	0.141	0.111	0.114	0.110	0.121	0.119
	$F_{32}^{(1)}$	0.163	0.129	0.137	0.134	0.133	0.133
	$M_{21}^{(1)}$	0.170	0.171	0.168	0.168	0.169	0.169
	$E_{11}^{(1)}$	0.172	0.172	0.169	0.169	0.171	0.171
	$E_{22}^{(1)}$	0.221	0.223	0.217	0.216	0.221	0.221
$N_1=100$	$E_{33}^{(1)}$	0.186	0.172	0.182	0.179	0.187	0.186
$N_2=150$	$F_{31}^{(2)}$	0.159	0.111	0.127	0.122	0.148	0.145
	$F_{32}^{(2)}$	0.160	0.129	0.133	0.130	0.132	0.131
	$M_{21}^{(2)}$	0.181	0.188	0.181	0.183	0.185	0.185
	$E_{11}^{(2)}$	0.285	0.271	0.273	0.270	0.281	0.280
	$E_{22}^{(2)}$	0.233	0.234	0.232	0.231	0.233	0.233
	$E_{33}^{(2)}$	0.236	0.226	0.227	0.224	0.235	0.234

Table 16 : Standard deviation of the various estimates for simulation 1

	Parameters	Standard deviation					
		ML1	ML2	BAY1	BAY2	BAY3	BAY4
	$F_{31}^{(1)}$	0.094	0.080	0.084	0.081	0.090	0.089
	$F_{32}^{(1)}$	0.102	0.083	0.092	0.088	0.084	0.084
	$M_{21}^{(1)}$	0.130	0.123	0.129	0.128	0.130	0.130
	$E_{11}^{(1)}$	0.215	0.216	0.212	0.210	0.215	0.215
	$E_{22}^{(1)}$	0.177	0.178	0.175	0.173	0.177	0.177
N_1=200	$E_{33}^{(1)}$	0.138	0.124	0.134	0.132	0.135	0.135
N_2=300	$F_{31}^{(2)}$	0.137	0.080	0.105	0.099	0.116	0.112
	$F_{32}^{(2)}$	0.125	0.083	0.097	0.094	0.099	0.096
	$M_{21}^{(2)}$	0.137	0.138	0.133	0.133	0.134	0.134
	$E_{11}^{(2)}$	0.217	0.221	0.216	0.215	0.217	0.217
	$E_{22}^{(2)}$	0.177	0.174	0.175	0.173	0.175	0.175
	$E_{33}^{(2)}$	0.148	0.141	0.141	0.140	0.143	0.142

Table 17 : Root Mean Squares errors between the various estimates and the population values for simulation 2

	Parameters	RMS					
		ML1	ML2	BAY1	BAY2	BAY3	BAY4
	$F_{31}^{(1)}$	0.166	0.193	0.153	0.152	0.167	0.165
	$F_{32}^{(1)}$	0.155	0.196	0.128	0.125	0.132	0.131
	$M_{21}^{(1)}$	0.157	0.166	0.158	0.158	0.164	0.163
	$E_{11}^{(1)}$	0.215	0.215	0.214	0.213	0.216	0.216
	$E_{22}^{(1)}$	0.202	0.210	0.195	0.193	0.197	0.197
$N_1=100$	$E_{33}^{(1)}$	0.212	0.222	0.204	0.201	0.202	0.202
$N_2=150$	$F_{31}^{(2)}$	0.152	0.180	0.123	0.123	0.139	0.136
	$F_{32}^{(2)}$	0.188	0.204	0.148	0.145	0.152	0.151
	$M_{21}^{(2)}$	0.183	0.191	0.182	0.184	0.181	0.181
	$E_{11}^{(2)}$	0.255	0.260	0.252	0.252	0.257	0.257
	$E_{22}^{(2)}$	0.261	0.280	0.249	0.247	0.251	0.250
	$E_{33}^{(2)}$	0.195	0.210	0.183	0.182	0.184	0.184

Table 18 : Root Mean Squares errors between the various estimates and the population values for simulation 2

	Parameters	RMS					
		ML1	ML2	BAY1	BAY2	BAY3	BAY4
	$F_{31}^{(1)}$	0.125	0.176	0.127	0.130	0.134	0.134
	$F_{32}^{(1)}$	0.116	0.148	0.108	0.107	0.110	0.110
	$M_{21}^{(1)}$	0.106	0.115	0.107	0.108	0.109	0.109
	$E_{11}^{(1)}$	0.247	0.260	0.241	0.240	0.246	0.246
	$E_{22}^{(1)}$	0.218	0.223	0.217	0.217	0.222	0.222
N_1=200	$E_{33}^{(1)}$	0.139	0.145	0.136	0.135	0.137	0.137
N_2=300	$F_{31}^{(2)}$	0.149	0.165	0.128	0.126	0.135	0.134
	$F_{32}^{(2)}$	0.143	0.157	0.114	0.109	0.113	0.112
	$M_{21}^{(2)}$	0.142	0.150	0.135	0.134	0.135	0.135
	$E_{11}^{(2)}$	0.245	0.262	0.243	0.242	0.246	0.246
	$E_{22}^{(2)}$	0.231	0.243	0.224	0.222	0.225	0.224
	$E_{33}^{(2)}$	0.136	0.155	0.129	0.128	0.130	0.130

Table 19 : Sample means of the various estimates for simulation 2

	Parameters	Sample Mean					
		ML1	ML2	BAY1	BAY2	BAY3	BAY4
	$F_{31}^{(1)}$	0.437	0.388	0.411	0.404	0.412	0.411
	$F_{32}^{(1)}$	0.462	0.535	0.494	0.503	0.513	0.514
	$M_{21}^{(1)}$	0.260	0.256	0.257	0.258	0.250	0.251
	$E_{11}^{(1)}$	0.414	0.404	0.406	0.405	0.409	0.408
	$E_{22}^{(1)}$	0.447	0.474	0.455	0.458	0.463	0.463
$N_1=100$	$E_{33}^{(1)}$	0.557	0.547	0.559	0.558	0.541	0.542
$N_2=150$	$F_{31}^{(2)}$	0.301	0.388	0.350	0.360	0.351	0.353
	$F_{32}^{(2)}$	0.639	0.535	0.589	0.577	0.573	0.571
	$M_{21}^{(2)}$	0.374	0.375	0.365	0.366	0.373	0.372
	$E_{11}^{(2)}$	0.635	0.648	0.637	0.638	0.640	0.640
	$E_{22}^{(2)}$	0.579	0.556	0.565	0.562	0.562	0.562
	$E_{33}^{(2)}$	0.584	0.641	0.611	0.617	0.623	0.623

Table 20 : Sample means of the various estimates for simulation 2

	Parameters	Sample Mean					
		ML1	ML2	BAY1	BAY2	BAY3	BAY4
	$F_{31}^{(1)}$	0.432	0.391	0.416	0.410	0.413	0.412
	$F_{32}^{(1)}$	0.478	0.536	0.496	0.502	0.509	0.510
	$M_{21}^{(1)}$	0.317	0.317	0.316	0.317	0.314	0.314
	$E_{11}^{(1)}$	0.458	0.450	0.452	0.451	0.455	0.454
	$E_{22}^{(1)}$	0.506	0.526	0.510	0.511	0.516	0.516
N_1=200	$E_{33}^{(1)}$	0.538	0.524	0.538	0.537	0.526	0.527
N_2=300	$F_{31}^{(2)}$	0.335	0.391	0.361	0.368	0.363	0.365
	$F_{32}^{(2)}$	0.609	0.536	0.583	0.575	0.573	0.572
	$M_{21}^{(2)}$	0.424	0.424	0.418	0.418	0.422	0.421
	$E_{11}^{(2)}$	0.640	0.648	0.641	0.641	0.642	0.642
	$E_{22}^{(2)}$	0.616	0.562	0.609	0.607	0.606	0.606
	$E_{33}^{(2)}$	0.564	0.632	0.576	0.580	0.584	0.584

Table 21 : Standard deviation of the various estimates for simulation 2

	Parameters	Standard deviation					
		ML1	ML2	BAY1	BAY2	BAY3	BAY4
	$F_{31}^{(1)}$	0.155	0.159	0.126	0.119	0.144	0.140
	$F_{32}^{(1)}$	0.152	0.195	0.129	0.126	0.133	0.132
	$M_{21}^{(1)}$	0.154	0.162	0.153	0.154	0.158	0.157
	$E_{11}^{(1)}$	0.199	0.194	0.194	0.192	0.198	0.197
	$E_{22}^{(1)}$	0.196	0.210	0.191	0.190	0.195	0.195
N_1=100	$E_{33}^{(1)}$	0.207	0.219	0.198	0.194	0.200	0.200
N_2=150	$F_{31}^{(2)}$	0.153	0.159	0.114	0.108	0.130	0.126
	$F_{32}^{(2)}$	0.186	0.195	0.149	0.145	0.152	0.149
	$M_{21}^{(2)}$	0.183	0.191	0.181	0.182	0.180	0.180
	$E_{11}^{(2)}$	0.255	0.258	0.252	0.251	0.256	0.256
	$E_{22}^{(2)}$	0.262	0.279	0.249	0.247	0.250	0.250
	$E_{33}^{(2)}$	0.196	0.208	0.184	0.183	0.184	0.184

Table 22 : Standard deviation of the various estimates for simulation 2

	Parameters	Standard deviation					
		ML1	ML2	BAY1	BAY2	BAY3	BAY4
	$F_{31}^{(1)}$	0.106	0.139	0.096	0.094	0.103	0.102
	$F_{32}^{(1)}$	0.115	0.145	0.109	0.108	0.111	0.111
	$M_{21}^{(1)}$	0.105	0.115	0.107	0.107	0.109	0.109
	$E_{11}^{(1)}$	0.245	0.258	0.239	0.237	0.244	0.244
	$E_{22}^{(1)}$	0.220	0.223	0.219	0.218	0.224	0.224
N_1=200	$E_{33}^{(1)}$	0.135	0.144	0.132	0.131	0.136	0.135
N_2=300	$F_{31}^{(2)}$	0.146	0.139	0.114	0.108	0.121	0.118
	$F_{32}^{(2)}$	0.144	0.145	0.113	0.107	0.111	0.109
	$M_{21}^{(2)}$	0.141	0.148	0.135	0.134	0.135	0.135
	$E_{11}^{(2)}$	0.244	0.260	0.242	0.241	0.245	0.245
	$E_{22}^{(2)}$	0.233	0.242	0.226	0.224	0.227	0.226
	$E_{33}^{(2)}$	0.132	0.153	0.128	0.128	0.130	0.130

REFERENCE

Bentler, P. M. (1989). *EQS : Structural Equation Program manual.* Los Angeles : BMDP Statistical Software.

Bock, R. D. & Lieberman, M. (1970). *Fitting a response model for n dichotomously scored items.* Psychometrika, 35, 179-197.

Bock, R. D. (1975). *Multivariate Statistical Methods in Behavioral Research.* McGraw-Hill.

Christoffersson, A. (1975). *Factor analysis of dichotomized variables.* Psychometrika, 40, 5-32.

Jeffrey, H. (1961). *Theory of Probability.* 3rd edition, Oxford.

Johnson, N. L. & Kotz, S. (1972). *Distributions in Statistics : Continuous Multivariate Distributions.* New York : Wiley.

Jöreskog, K. G. (1971). *Simultaneous factor analysis in several populations.* Psychometrika, 36, 409-426.

Jöreskog, K. G. & Sörbom, D. (1988). *LISREL VII : A Guide to the program and applications.* Chicago, IL : SPSS, Inc.

Lee, S. Y. (1981). *A Bayesian approach to confirmatory factor analysis.* Psychometrika, 46, 153-160.

Lee, S. Y. (1985). *On testing functional constraints in structural equation models.* Biometrika, 72, 125-131.

Lee, S. Y. (1988a). *Estimation of structural equation models with exact and stochastic prior information.* Communication in Statistics : Theory and Methods, 17, 1988, 1565-1590.

Lee, S. Y. (1988b). *Bayesian analysis of stochastic constraints in structural equation models.* Technical report, Dept. of Statistics, The Chinese University of Hong Kong.

Lee, S. Y. & Bentler, P. M. (1980). *Some asymptotic properties of constrained generalized least squares estimation in covariance structure models.* South African Statistical Journal, 14, 121-136.

Lee, S. Y., & Jennrich, R. I. (1979). *A study of algorithms for covariance structure analysis with specific comparisions using factor analysis.* Psychometrika, 43, 99-113.

Lee, S. Y. & Poon, W. Y. (1985). *Further developments on constrained estimation in analysis of covariance structures.* The Statistician, 34, 305-316.

Lee, S. Y. & Tsui, K. L. (1982). *Covariance structure analysis in several populations.* Psychometrika, 47, 297-308.

Lee, S. Y., Poon, W. Y. & Bentler, P. M. (1989a). *Full Maximum likelihood analysis of structural equation models with polytomous variables.* Statistics and Probability Letters, 9, 91-97.

Lee, S. Y., Poon, W. Y. & Bentler, P. M. (1989b). *Simultaneous analysis of multivariate polytomous variates in several groups.* Psychometrika, 54, 63-73.

Lindley, D. V. & Smith, A. F. M. (1972). *Bayes esitmates for the linear model.* Journal of the Royal Statistical Society, Series B, 34, 1-42.

Muthen, B. (1978). *Contribution to factor analysis of dichotomous variables.* Psychometrika, 43, 551-560.

Muthen, B. & Christoffersson, A. (1981). Simultaneous factor analysis of dichotomous variables in several groups. Psychometrika, 46, 407-420.

Newcomb, M. D. & Bentler, P. M. (1988). *Consequences of adolescent drug use : Impact on the lives of young adults.* Newbury Park, CA : Sage.

Poon, W. Y. & Lee, S. Y. (1987). *Maximum likelihood estimation of multivariate polyserial and polychoric correlation coefficients.*

Psychometrika, 52, 409-430.

Poon, W. Y., Lee, S. Y., Afifi, A. A. & Bentler, P. M. (1989). *Analysis of multivariate polytomous variates in several groups via partition maximum likelihood approach.* Communicational Statistics & Data Analysis, in press.

Poon, W. Y., Lee, S. Y., Bentler, P. M. & Afifi, A. A. (1989). *Covariance structural analysis with polytomoous variables in several populations.* (Submitted)

Raiffa, H. & Schlaifer, H. (1961). *Applied Statistical Decision Theory.* 2nd ed., Boston : Graduate School of Business Adminstration, Harvard.

Schervish, M. J. (1984). *Multivariate normal probabilities with error bound.* Applied Statistics, 33, 81-94.

Smith, A. F. M. (1973). *Bayes estimates in one-way and two ways models.* Biometrika, 60, 319-329.

Sörbom, D. (1974). *A general method for studying differences in factor means and factor structure between groups.* British Journal of Mathematical and Statistical Psychology, 27, 229-239.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics* New York : Wiley.