PSYCHOMETRIKA—VOL. 46, NO. 2.
JUNE, 1981

A BAYESIAN APPROACH TO CONFIRMATORY FACTOR ANALYSIS

SIK-YUM LEE

THE CHINESE UNIVERSITY OF HONG KONG

Confirmatory factor analysis is considered from a Bayesian viewpoint, in which prior information on parameter is incorporated in the analysis. An iterative algorithm is developed to obtain the Bayes estimates. A numerical example based on longitudinal data is presented. A simulation study is designed to compare the Bayesian approach with the maximum likelihood method.

Key words: conjugate family, prior distributions, posterior density function, maximum likelihood estimation, longitudinal data, simulation study.

1. Introduction

One of the most widely used statistical techniques in multivariate behavioral research is the confirmatory factor analysis developed by Jöreskog [1969]. The model is of the following form

$$\Sigma = \Lambda\Phi\Lambda' + \Psi, \tag{1}$$

where Λ is a $p \times k$ matrix of factor loadings, Φ is a $k \times k$ symmetric matrix of factor covariances, and Ψ is a $p \times p$ diagonal matrix of unique variances. Let S be the sample covariance matrix which is assumed to have a Wishart distribution $W(\Sigma/(N-1), N-1)$ where N is the sample size. Traditional maximum likelihood estimates of parameters in matrices Λ, Φ and Ψ are values that minimize the function,

$$F(\Lambda, \Phi, \Psi) = \log|\Sigma| + \mathrm{tr}(S\Sigma^{-1}). \tag{2}$$

In this paper, we consider the situation in which prior knowledge about the parameters is available. From a Bayesian viewpoint, this prior information should be incorporated with the sample information in making inferences.

Basic theory of the Bayesian approach to confirmatory factor analysis is discussed in Section 2. Posterior densities associated with different prior knowledge on Λ, Φ and Ψ are presented. The modal estimate of the posterior density is defined as the Bayes estimate. Since in the present model, this estimate cannot be solved in closed form, an iterative procedure for obtaining it is described in Section 3. A numerical example is provided in Section 4. A preliminary report of a simulation study for comparison of the maximum likelihood approach with the Bayesian approach is given in Section 5. The indication is that the Bayesian procedure will give better estimates than a non-Bayesian one whenever the prior is better than a uniform distribution.

2. Bayesian Approach in Confirmatory Factor Analysis

Let θ denote the parameters of interest in the model and ξ the nuisance parameters. We consider that θ and ξ are random with joint distribution $p(\theta, \xi)$. This prior distribution

Computer facilities were provided by the Computer Services Center, The Chinese University of Hong Kong.

Requests for reprints should be sent to S. Y. Lee, Department of Mathematics, The Chinese University of Hong Kong, Shatin, N.T. HONG KONG.

0033-3123/81/0600-3009$00.75/0

is combined with the likelihood function of S, $p(S \mid \theta, \xi)$ to yield a joint posterior distribution,

$$p(\theta, \xi \mid S) \propto p(\xi)p(\theta \mid \xi)p(S \mid \theta, \xi). \tag{3}$$

One of the major problems is that of choosing the prior distribution. In most applications of the model, for example the multitrait-multimethod data and path analysis data, investigators may wish to assign fixed values on certain parameters. There are at least two reasons of doing this: (i) to achieve identificability of the model, and (ii) to obtain more meaningful interpretations. For a Bayesian, this is equivalent to assigning those particular parameters to corresponding values with probability one. The fixed values are not considered as parameters, and we are only interested in the estimate of θ, which consists of the remaining free parameters in Λ, θ and Ψ. Several types of prior information that include most of the interesting cases in real applications are presented as follows.

Case 1. Let λ_i, $i = 1, \ldots, m$ denote the free parameter in Λ. Suppose that the prior distribution of λ_i satisfies the exchangeable property as discussed in Lindley & Smith [1972]. Namely, we assume independently that $\lambda_i \sim N(\eta, \sigma^2)$, $i = 1, \ldots, m$. We assign a locally uniform, vague prior for η, and hence its density is proportional to a constant [see, for example, Jeffreys, 1961, and Zellner, 1971]. To specify the prior distribution of the nuisance parameter σ^2, we use the appropriate conjugate family [see e.g. Raiffa & Schlaifer, 1961], which is here the inverse $-\chi^2$ family. Specifically for given prior constants μ and α, we assume independently that $\mu\alpha/\sigma^2 \sim \chi^2_\mu$. Thus,

$$p(\sigma^2 \mid \mu, \alpha) \propto (\sigma^2)^{-(\mu+2)/2} \exp\left\{-\frac{\mu\alpha}{2\sigma^2}\right\}.$$

According to Lindley [1971], to represent prior ignorance for σ^2, a small positive value of μ should be chosen. However, as we will see from (7), with substantial amount of data available, the prior constants scarcely affect the analysis. The conjugate distribution for Φ is to suppose Φ^{-1} has an independent Wishart distribution with matrix R and degrees of freedom ρ. To specify the prior distribution for the unique variance Ψ_j, $j = 1, \ldots, p$, we also use the conjugate inverse $-\chi^2$ family. That is, for given v_j and β_j, we suppose, independently, that $v_j\beta_j/\Psi_j \sim \chi^2_{v_j}$. Again, to represent prior ignorance for Ψ_j, a small positive value of v_j is chosen in practice.

The joint posterior density for these parameters is proportional to

$$\begin{aligned}
&|\Sigma|^{-n/2} \exp\{-\tfrac{1}{2}n \operatorname{tr} S\Sigma^{-1}\}(\sigma^2)^{-m/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{m}(\lambda_i - \eta)^2\right\} \\
&\quad \times (\sigma^2)^{-(\mu+2)/2} \exp\left\{-\frac{\mu\alpha}{2\sigma^2}\right\} |\Phi|^{-(\rho-k-1)/2} \exp\{-\tfrac{1}{2} \operatorname{tr} R\Phi^{-1}\} \\
&\quad \times \prod_{j=1}^{p} \Psi_j^{-(v_j+2)/2} \exp\left\{-\frac{1}{2}\sum_{j=1}^{p}\frac{v_j\beta_j}{\Psi_j}\right\},
\end{aligned} \tag{4}$$

where $n = N - 1$. The nuisance parameters can be eliminated by sequential integration with respect to η and σ^2. The joint posterior density for λ_i, Φ and Ψ_j is then proportional to

$$\begin{aligned}
&|\Sigma|^{-n/2} \exp\{-\tfrac{1}{2}n \operatorname{tr} S\Sigma^{-1}\} \prod_{i=1}^{m} \{(\lambda_i - \bar{\lambda})^2 + m^{-1}\mu\alpha\}^{-(m+\mu-1)/2} \\
&\quad \times |\Phi|^{-(\rho-k-1)/2} \exp\{-\tfrac{1}{2} \operatorname{tr} R\Phi^{-1}\} \prod_{j=1}^{p} \Psi_j^{-(v_j+2)/2} \exp\left\{-\frac{1}{2}\sum_{j=1}^{p}\frac{v_j\beta_j}{\Psi_j}\right\},
\end{aligned} \tag{5}$$

where $\bar{\lambda} = m^{-1} \sum_{i=1}^{m} \lambda_i$.

Adopting the procedure suggested by Lindley & Smith [1972], and Smith [1973], the modal estimate of the posterior density is defined as the Bayes solution. For convenience, we solve this problem by minimizing the following function

$$F_1(\Lambda, \Phi, \Psi) = F(\Lambda, \Phi, \Psi) + B_1(\Lambda, \Phi, \Psi), \tag{6}$$

where $F(\Lambda, \Phi, \Psi)$ is given in (1), and

$$B_1(\Lambda, \Phi, \Psi) = \frac{1}{n}\left[(m + \mu - 1)\log \sum_{i=1}^{m} \{(\lambda_i - \bar{\lambda})^2 + m^{-1}\mu\alpha\}\right.$$
$$\left. + (\rho - k - 1)\log |\Phi| + \operatorname{tr} R\Phi^{-1} + \sum_{j=1}^{p} \left\{(v_j + 2)\log \Psi_j + \frac{v_j\beta_j}{\Psi_j}\right\}\right]. \tag{7}$$

The solution cannot be expressed in closed form. An iterative algorithm will be described in Section 3 to obtain the estimate.

Case 2. Suppose the prior informations about Φ and Ψ_j are the same as in Case 1, but the exchangeability of the prior distribution on λ_i is doubtful. Here we assume independently that $\lambda_i \sim N(\lambda_i^*, \sigma_i^2)$, $i = 1, \ldots, m$, where λ_i^* is a prior constant, and σ_i^2 is a nuisance parameter. We again use the conjugate inverse $-\chi^2$ family for the prior distribution of σ_i^2. Namely, for given positive values μ_i and α_i, we assume independently that $\mu_i\alpha_i/\sigma_i^2 \sim \chi^2_{\mu_i}$. Under these assumptions nuisance parameters σ_i^2 can be similarly eliminated by integration. Consequently, the joint posterior density for λ_i, Φ and Ψ_j is proportional to

$$|\Sigma|^{-n/2} \exp\{-\tfrac{1}{2}n \operatorname{tr} S\Sigma^{-1}\} \prod_{i=1}^{m} \{(\lambda_i - \lambda_i^*)^2 + \mu_i\alpha_i\}^{-(\mu_i+1)/2}$$
$$\times |\Phi|^{-(\rho-k-1)/2} \exp\{-\tfrac{1}{2} \operatorname{tr} R\Phi^{-1}\} \prod_{j=1}^{p} \Psi_j^{-(v_j+2)/2} \exp\left\{-\frac{1}{2}\sum_{j=1}^{p} \frac{v_j\beta_j}{\Psi_j}\right\}. \tag{8}$$

The fit function we are going to minimize in obtaining the estimate is then equal to

$$F_2(\Lambda, \Phi, \Psi) = F(\Lambda, \Phi, \Psi) + B_2(\Lambda, \Phi, \Psi), \tag{9}$$

where

$$B_2(\Lambda, \Phi, \Psi) = \frac{1}{n}\left[\sum_{i=1}^{m} (\mu_i + 1)\log\{(\Lambda_i - \Lambda_i^*)^2 + \mu_i\alpha_i\}\right.$$
$$\left. + (\rho - k - 1)\log|\Phi| + \operatorname{tr} R\Phi^{-1} + \sum_{j=1}^{p} \left\{(v_j + 2)\log \Psi_j + \frac{v_j\beta_j}{\Psi_j}\right\}\right]. \tag{10}$$

Case 3. In this case, we wish to consider the situation in which little prior knowledge about the parameters is available. Following the invariance theory of Jeffreys [1961], to represent prior ignorance about λ_i we set each of its densities proportional to a constant. The distribution of Φ^{-1} is assumed to be Wishart with degrees of freedom ρ and unknown matrix R. Due to Jeffreys [1961] (see also Zellner [1971]), we suppose the prior density of R is proportional to $|R|^{-(k+1)/2}$. For Ψ_j we again suppose independently that $v_j\beta_j/\Psi_j \sim \chi^2_{v_j}$ with small positive value of v_j. Thus, the joint posterior density is proportional to

$$|\Sigma|^{-n/2} \exp\{-\tfrac{1}{2}n \operatorname{tr} S\Sigma^{-1}\}\, |R|^{-(k+1)/2} |R|^{\rho/2} |\Phi|^{-(\rho-k-1)/2} \exp\{-\tfrac{1}{2} \operatorname{tr} R\Phi^{-1}\}$$
$$\times \prod_{j=1}^{p} \Psi_j^{-(v_j+2)/2} \exp\left\{-\frac{1}{2}\sum_{j=1}^{p} \frac{v_j\beta_j}{\Psi_j}\right\}. \tag{11}$$

The nuisance matrix R can be removed by integration. The joint posterior density for the remaining parameters is then proportional to

$$|\Sigma|^{-n/2} \exp\{-\tfrac{1}{2}n \operatorname{tr} S\Sigma^{-1}\} \times |\Phi|^{(k+1)/2} \times \prod_{j=1}^{p} \Psi_j^{-(v_j+2)/2} \exp\left\{-\frac{1}{2}\sum_{j=1}^{p}\frac{v_j\beta_j}{\Psi_j}\right\}. \tag{12}$$

To obtain the Bayes estimate, we consider

$$F_3(\Lambda, \Phi, \Psi) = F(\Lambda, \Phi, \Psi) + B_3(\Lambda, \Phi, \Psi), \tag{13}$$

where

$$B_3(\Lambda, \Phi, \Psi) = \frac{1}{n}\left[-(k+1)\log|\Phi| + \sum_{j=1}^{p}\left\{(v_j+2)\log\Psi_j + \frac{v_j\beta_j}{\Psi_j}\right\}\right]. \tag{14}$$

Case 4. In the context of exploratory factor analysis Martin and McDonald [1975] considered a Bayesian approach to handle the problem of Heywood cases (nonpositive estimates of unique variances). In the exploratory model, the loading matrix Λ is free to rotate; hence only prior information on Ψ_j is considered. They assumed independently that the prior density of Ψ_j is proportional to $\exp\{-\tfrac{1}{2}\beta_j/\Psi_j\}$ for some prior constant β_j. Such prior knowledge can be easily incorporated into our previous analysis. It can be shown, by similar arguments, that the functions $F_k(\Lambda, \Phi, \Psi)$, $k = 4, 5, 6$ for obtaining the corresponding Bayesian estimates are in the same form as in (6), (9) and (13), with

$$B_4(\Lambda, \Phi, \Psi) = \frac{1}{n}\left[(m+\mu-1)\log\sum_{i=1}^{m}\{(\Lambda_i - \bar{\Lambda})^2 + m^{-1}\mu\alpha\}\right.$$
$$\left. + (\rho - k - 1)\log|\Phi| + \operatorname{tr} R\Phi^{-1} + \sum_{j=1}^{p}\frac{\beta_j}{\Psi_j}\right], \tag{15}$$

$$B_5(\Lambda, \Phi, \Psi) = \frac{1}{n}\left[\sum_{i=1}^{m}(\mu_i + 1)\log\{(\Lambda_i - \Lambda_i^*)^2 + \mu_i\alpha_i\}\right.$$
$$\left. + (\rho - k - 1)\log|\Phi| + \operatorname{tr} R\Phi^{-1} + \sum_{j=1}^{p}\frac{\beta_j}{\Psi_j}\right], \tag{16}$$

$$B_6(\Lambda, \Phi, \Psi) = \frac{1}{n}\left[-(k+1)\log|\Phi| + \sum_{j=1}^{p}\frac{\beta_j}{\Psi_j}\right]. \tag{17}$$

It should be noted that since the confirmatory model subsumes the exploratory model as special case, our study includes those cases given by Martin & McDonald [1975].

It should be pointed out that other prior knowledge, for example with λ_i distributed as in Case 2, and with Φ and Ψ distributed as in Case 3, can be similarly considered in the analysis. Moreover, we admit that certain types of prior distributions of interest are left out, such as subsets of parameters having exchangeable prior distributions different from other subsets and subsets of parameters which are equal but not constant.

3. An Algorithm for Bayesian Estimation

The classical iterative procedure for searching a local minimum of a function is the Newton-Raphson algorithm. A basic step of the algorithm in applying to $F_k(\theta)$, $k = 1, \ldots, 6$ is defined by

$$\Delta\theta = -\gamma\ddot{F}_k(\theta)^{-1}\dot{F}_k(\theta),$$

where γ is a step-halving parameter which takes the first value in the sequence $1, \frac{1}{2}, \frac{1}{4}, \ldots$ that reduces $F_k(\theta)$,

$$\dot{F}_k(\theta) = \dot{F}(\theta) + \dot{B}_k(\theta),$$

and

$$\ddot{F}_k(\theta) = \ddot{F}(\theta) + \ddot{B}_k(\theta),$$

with $\dot{F}(\theta) = (\partial F/\partial \theta_i)$, $\ddot{F}(\theta) = (\partial^2 F/\partial \theta_i \partial \theta_j)$, $\dot{B}_k(\theta) = (\partial B_k/\partial \theta_i)$, and $\ddot{B}_k(\theta) = (\partial^2 B_k/\partial \theta_i \partial \theta_j)$ as the gradient vectors and the Hessian matrices of $F(\theta)$ and $B_k(\theta)$, respectively. Since in confirmatory factor analysis, $\ddot{F}(\theta)$ is very complicated to program and requires long computer time to evaluate, it is advantageous to approximate it by the information matrix, $I(\theta)$ [see, for example, Lee & Jennrich, 1979]. Thus, the algorithm we are going to use is defined by

$$\Delta\theta = -\gamma U_k(\theta)^{-1}\dot{F}_k(\theta), \tag{18}$$

where $U_k(\theta) = I(\theta) + \ddot{B}_k(\theta)$. At the end of each step, θ is replaced by $\theta + \Delta\theta$, and the process is continued until the root mean squares of $\Delta\theta$ and $\dot{F}(\theta)$ are sufficiently small.

Expressions for $\dot{F}(\theta)$ and $I(\theta)$ have been reported in Jöreskog [1969], so to conserve space they are not presented here. Derivatives $\dot{B}_k(\theta)$ and $\ddot{B}_k(\theta)$ can be obtained easily by using the following results:

$$\frac{\partial}{\partial \Phi_{ab}} \log|\Phi| = (2 - \delta_{ab})\Phi_{ab}^{-1}, \qquad \frac{\partial}{\partial \Phi_{ab}} \operatorname{tr} R\Phi^{-1} = -(2 - \delta_{ab})(\Phi^{-1}R\Phi^{-1})_{ab},$$

$$\frac{\partial^2 \log|\Phi|}{\partial \Phi_{ab}\,\partial \Phi_{cd}} = -\tfrac{1}{2}(2 - \delta_{ab})(2 - \delta_{cd})(\Phi_{ac}^{-1}\Phi_{bd}^{-1} + \Phi_{ad}^{-1}\Phi_{bc}^{-1}),$$

$$\frac{\partial^2 \operatorname{tr} R\Phi^{-1}}{\partial \Phi_{ab}\partial \Phi_{cd}} = \tfrac{1}{2}(2 - \delta_{ab})(2 - \delta_{cd})\{\Phi_{ac}^{-1}(\Phi^{-1}R\Phi^{-1})_{bd} + \Phi_{ad}^{-1}(\Phi^{-1}R\Phi^{-1})_{bc}\},$$

where δ denotes the Kronecker delta.

4. *Numerical Example*

Our example is part of the longitudinal data from a growth study conducted by Education Testing Service [see, Hilton, Note 1]. The data set consists of 383 school girls. Their scores were recorded in Grades 7 and 9 on six tests: Scholastic Aptitude Tests in verbal (SCATV) and quantitative (SCATQ) part, Achievement Tests in mathematics (MATH), science (SCI), social studies (SS) and reading (READ). For completeness, the sample covariance matrices [see, Sörbom & Jöreskog, 1976] are reported in Table 1. The covariance matrix from Grade 7 students was analyzed using the maximum likelihood method, with two correlated factors, namely, the quantitative ability factor (Q) and verbal ability factor (V). The pattern of the loading matrix is given by

$$\begin{array}{cccccc} \text{MATH} & \text{SCI} & \text{SS} & \text{READ} & \text{SCATV} & \text{SCATQ} \end{array}$$
$$\Lambda' = \begin{pmatrix} * & * & * & 0 & 0 & 1 \\ 0 & * & * & * & 1 & 0 \end{pmatrix},$$

where * denotes a free parameter to be estimated, and both 0 and 1 are fixed loadings of zero and one, respectively. The unit loading is used to fix the scales of measurement in factors V and Q. The covariance matrix of Grade 9 was then analyzed by the Bayesian

Table 1

Covariance Matrices of the Data from Grade 7 and Grade 9 Students.

a) Grade 7

	MATH 1	SCI 2	SS 3	READ 4	SCATV 5	SCATQ 6
1	129.978					
2	68.313	85.032		symmetric		
3	101.273	88.562	170.195			
4	116.164	94.846	144.585	214.351		
5	91.015	76.307	118.944	141.015	135.095	
6	99.872	68.384	104.240	110.713	87.998	154.267

b) Grade 9

	1	2	3	4	5	6
1	135.974					
2	89.207	130.938		symmetric		
3	109.619	112.694	177.551			
4	94.151	96.424	112.176	147.687		
5	94.353	98.539	115.459	112.119	137.098	
6	127.463	96.029	117.465	104.273	100.720	211.259

approach with prior distributions as described in Case 2, using maximum likelihood estimates from Grade 7 students as prior distributions.

According to Lindley [1971], to represent prior ignorance for σ_i^2 and Ψ_j, small positive values of μ_i and v_j should be taken. It can be seen from (10) that with a sufficiently large sample size, prior constants μ_i, α_i, ρ, v_j and β_j in $B_2(\Lambda, \Phi, \Psi)$ scarcely affect the analysis. In this example, for convenience we chose appropriate values so that $B_2(\Lambda, \Phi, \Psi)$ in approximately equal to

$$\frac{1}{n}\left[\sum_{i=1}^{m} \log\{(\lambda_i - \lambda_i^*)^2 + 0.1\} + \log|\Phi| + \operatorname{tr} R\Phi^{-1} + \sum_{j=1}^{p}\left\{2 \log \Psi_j + \frac{1}{\Psi_j}\right\}\right].$$

Based on the algorithm described in Section 3, a computer program has been implemented; the listings of this program can be obtained upon request from the author. In this example the program converged in seven iterations, where both the root mean squares of $\dot{F}_k(\theta)$ and $\Delta\theta$ are less than 0.0005. Bayes estimates together with the maximum likelihood estimates are reported in Table 2.

The data have been reanalyzed with the prior distribution of Ψ_j replaced by that given in Case 4 of Section 2. We found that the solution is insensitive to the small positive values β_j; and it is very close to the Bayes estimates given in Table 2.

Table 2

Bayes Estimates and Maximum Likelihood Estimates (in parenthesis).

$$\Lambda_{ij} = \begin{bmatrix} .931(.926) & 0\ (0) \\ .138(.123) & .766(.783) \\ .269(.263) & .800(.806) \\ 0\ (0) & .989(.988) \\ 0\ (0) & 1\ (1) \\ 1\ (1) & 0\ (0) \end{bmatrix} \qquad \Phi_{ij} = \begin{bmatrix} 136.472(137.482) & \\ 101.509(102.125) & 111.802(111.489) \end{bmatrix} \qquad \Psi_j = \begin{bmatrix} 17.389(18.072) \\ 40.922(40.917) \\ 52.104(52.260) \\ 38.521(38.949) \\ 25.142(26.609) \\ 73.752(73.777) \end{bmatrix}$$

The method developed here has been applied to several Heywood case examples. The results indicate that with moderate sample sizes, Bayes solutions obtained from different prior distributions on Ψ_i are different. Moreover, these solutions are quite different from the maximum likelihood estimates that are subject to inequality constraints $\Psi_j \geq 0$ [see, Lee, 1980]. A more detailed discussion on this topic will be reported later.

5. Simulation Study

The following simulation study is designed to compare the Bayesian approach with the maximum likelihood approach. Since this type of study is expensive, we only present here some admittedly limited results. In particular, we consider the Bayes estimates with prior information where the fit function of interest is given by $F_5(\Lambda, \Phi, \Psi)$. The maximum likelihood estimates reported in Table 2 are taken as the population values for Λ, Φ and Ψ, and the population covariance matrix Σ was computed according to (1). The main steps of the simulation are:

(i) Based on a sample of size $N = 50$, a pseudo-random Wishart matrix S^* with covariance matrix Σ/n and degree of freedom n was generated according to the method described by Kshirsagar [1959].

(ii) Based on S^*, maximum likelihood estimates of Λ and Φ were obtained. These estimates are used as prior knowledge for λ_i^* and R.

(iii) Another 50 pseudo-random Wishart matrices with same covariance matrix and degree of freedom as in (i) were simulated. These matrices were analyzed by both the maximum likelihood procedure (MLP) and the Bayesian procedure (BP) described as before. The corresponding root mean squared errors (RMS) of the estimates were computed respectively.

(iv) Steps (i) to (iii) were repeated with $N = 100$, 150 and 200.

The results are summarized in Table 3. We observe that the Bayes estimates of the factor loadings are substantially better, while for the estimates of Φ and Ψ_j, Bayes estimates are slightly better. Hence, if appropriate prior information is available, for example in longitudinal studies or multiple group factor analysis, the Bayesian approach represents an attractive procedure for analyzing the confirmatory factor analytic model.

Table 3

Summary of Result from Simulation Study.

	N = 50		N = 100		N = 150		N = 200	
	MLP	BP	MLP	BP	MLP	BP	MLP	BP
RMS(Λ)	0.252	0.151	0.174	0.131	0.134	0.106	0.105	0.088
RMS(Φ)	36.832	33.817	24.329	23.287	19.985	19.398	17.600	17.373
RMS(Ψ)	12.880	12.373	9.397	9.223	7.300	7.199	5.930	5.878

REFERENCE NOTE

1. Hilton, T. L. *Growth study annotated bibliography* (Prog. Rep. 69–71). Princeton, N.J.: Educational Testing Service, 1969.

REFERENCES

Jeffrey, H. *Theory of Probability*, 3rd edition. Oxford: Clarendon, 1961.

Jöreskog, K. G. A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 1969, *34*, 183–202.

Kshirsagar, A. M. Bartlett decomposition and Wishart distribution. *Annals of Mathematical Statistics*, 1959, *30*, 230–241.

Lee, S. Y. Estimation of covariance structure models with parameters subject to functional restraints. *Psychometrika*, 1980, *45*, 309–324.

Lee, S. Y. & Jennrich, R. I. A study of algorithms for covariance structure analysis with specific comparisons using factor analysis. *Psychometrika*, 1979, *44*, 99–113.

Lindley, D. V. The estimation of many parameters. In V. P. Godamke & D. A. Sprott (eds.) *Foundations of Statistical Inference*. Toronto: Holt, Rinehart and Winston, 1971, pp. 435–455.

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

Martin, J. K. & McDonald, R. P. Bayesian estimation in unrestricted factor analysis: A treatment for Heywood cases. *Psychometrika*, 1975, *40*, 505–517.

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

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

Sörbom, D. & Jöreskog, K. G. *Confirmatory Factor Analysis with Model Modification, User's Guide*. Chicago: National Educational Resources, 1976.

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

Manuscript received 8/11/80
Final version received 1/20/81