

行政院國家科學委員會專題研究計畫成果報告

成長曲線模型的估計與預測

On Estimation and Prediction of Growth Curves Models

計畫編號：NSC 87-2119-M-009-001

執行期限：86年8月1日至87年7月31日

主持人：李昭勝 國立交通大學統計學研究所

E-mail : jclee@stat.nctu.edu.tw

本報告含三篇完成之研究成果，分別與三位共同主持人：洪慧念教授、謝國文教授、陳婉淑教授合作。

一、Analysis of Drug Dissolution Data

此乃與洪慧念教授及 D. T. Chen, James Chen 合作之文章。本文主要利用成長曲線模型來分析藥物溶解資料，並以預測的準確度來做模型選擇的根據。這些均是本計畫之研究範圍。本文將刊登於 *Statistics in Medicine*。其中英文之摘要如下。

(一)、中文摘要

人體吸收藥物之程度端賴藥物之溶解度。適當之溶解特性有助於達成藥物之療效。我們利用成長曲線模型來判斷藥物溶解度的相似性。Box-Cox 轉換及 log 轉換應用在溶解度的觀察值。利用樣本重複使用預測法來選取最佳預測模型。此外，也從假設檢定及貝氏方法討論溶解度的相似性。

(二)、英文摘要

Drug absorption in the human body depends on the dissolution rate of the drug. Suitable dissolution characteristics are important to ensure that the drug will achieve the desired therapeutic effects. To assess the similarity of dissolution rates of several drug

lots, we apply a general growth curve model with different covariance structures. The Box-Cox power transformation and the naïve log transformation are applied to a function of the dissolution rate. The predictive sample-reuse, or cross-validation, method is employed in selecting an appropriate model with best predictive accuracy. A testing procedure for examining the similarity among the drug lots is also conducted. A partially Bayesian approach is used for the assessment of dissolution equivalence.

(三) 本研究成果之參考文獻如下

- Anderson T. W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd edition. New York: Wiley.
- Chen, J. and Tsong, Y. (1997). Multiple time-point dissolution specification for a sampling acceptance plan, *J. Bioph. Stat*, 259-270.
- Geisser, S. C. (1975). The predictive sample reuse method with application. *JASA*, 70, 320-328.
- Keramidas, E. M. and Lee, J. C. (1990). Forecasting technological substitutions with concurrent short time series, *JASA*, 85, 625-632
- Keramidas, E. M. and Lee, J. C. (1995). Selection of a covariance structure for growth curves, *Biom. J.*, 37, 783-797.
- Khatri, C. G. (1966). A note on MANOVA

model applied to problems in growth curve. *Annals of the Institute of Statistical Mathematics*, 18, 75-86.

Krishnaiah, P. R. and Lee, J. C. (1980). Likelihood ratio tests for mean vectors and covariance matrices. *Handbook of Statistics*, Vol. I, 513-570.

Lee, J. C. and Geisser, S. (1972). Growth curve prediction. *Sankhya, Series A*, 34, 393-412.

Lee, J. C. (1988). Prediction and estimation of growth curves with special covariance structures. *JASA*, 83, 432-440

Lee, J. C. (1991). Tests and model selection for the general growth curve model. *Biometrics*, 47, 147-159.

Potthoof, R. F. and Roy, S. N. (1964). A generalized multivariate analysis of variance model useful especially for a growth curve problems. *Biometrika*, 51, 313-326.

二、Bayesian Prediction Analysis for Growth Curve Model Using Noninformative Priors

此乃與謝國文教授合作之成果。本文利用無信息的先驗密度函數來探討非均衡成長曲線模型的貝氏分析。主要利用MCMC法及逼近法來研究此模型的預測理論。此乃本計劃的研究範圍。以後將送到國際期刊發表。其中英文摘要如下：

(一)、中文摘要

我們利用無信息的先驗密度函數來探討非均衡成長模型的貝氏分析。由於預測密度函數非常複雜，我們提出逼近法及MCMC法來做預測分析。結果用實際資料與來當範例。簡單的逼近法比Fearn 1975年的結果好。它與MCMC法中的Rao-Blackwellization法相近。

(二)、英文摘要

We apply a Bayesian approach to the problem of prediction in an unbalanced growth curve model using noninformative priors. Due to the complexity of the model, no analytic forms of the predictive densities are available. We propose both approximations and a prediction-oriented Metropolis-Hastings sampling algorithm for two types of prediction, namely the prediction of future observations of a new subject and the prediction of future values of a partially observed subject. They are illustrated and compared through real data and simulation studies. The approximation compares favorably with the approximation in Fearn(1975) and is very comparable to the more accurate Rao-Blackwellization from Metropolis-Hastings sampling algorithm.

(三) 本研究成果之參考文獻如下

Carter, R. L., Resnick, M. B., Ariet, M., Shieh, G. and Voncsh, E. F. (1992), A random coefficient growth curve analysis of mental development in low-birth-weight infants, *Statistics in Medicine*, 11, 243-256.

Chub, S. and Greenberg, E. (1995), Understanding the Metropolis-Hastings algorithm, *The American Statistician*, 49, 327-335.

Casella, G. and George, E. (1992), Explaining the Gibbs sampler, *The American Statistician*, 46, 167-174.

Elston, R. C. and Grizzle, J. E. (1962), Estimation of time response curves and their confidence band *Biometrics*, 18, 148-159.

Fearn, T. (1975), A Bayesian approach to growth curves, *Biometrika*, 62, 89-100.

Geisser, S. (1970), Bayesian analysis of growth curves, *Sankhya, Ser. A*, 32, 53-64.

Geisser, S. (1981), Sample reuse procedures for prediction of the unobserved portion of a partially observed vector, *Biometrika*, 68,

243-250.

Gelfand, A. E., Hills, S. E., Racine-Poon, A. and Smith, A. F. M. (1990), Illustration of Bayesian inference in normal data models using Gibbs sampling, *Journal of the American Statistical Association*, 85, 972-985.

Gelman, A. and Rubin, D. B. (1992), Inference from iterative simulation using multiple sequences, *Statistical Science*, 7, 457-472.

Gilks, W. R., Richardson, S. and Spiegelhalter, D. J. eds. (1996), *Markov Chain Monte Carlo in Practice*, Chapman and Hall, London.

Gilks, W. R., Wang, C. C., Yvonnet, B. and Cuursaget, P. (1993), Random-effects models for longitudinal data using Gibbs sampling, *Biometrics*, 49, 441-453.

Grizzle, J. E. and Allen, D. M. (1969), Analysis of growth and response curves, *Biometrics*, 25, 357-381.

Khatri, C. G. (1966), A note on MANOVA model applied to problems in growth curve, *Annals of the Institute of Statistical Mathematics*, 18, 75-86.

Kleinbaum, D. G. (1973), A generalization of the growth curve model which allows missing data, *Journal of Multivariate Analysis*, 3, 117-124.

Lee, J. C. (1988), Prediction and estimation of growth curves with special covariance structures, *Journal of the American Statistical Association*, 83, 432-440.

Lee, J. C. and Geisser, S. (1972), Growth curve prediction, *Sankhya, Ser. A*, 34, 393-412.

Lee, J. C. and Geisser, S. (1975), Applications of growth curve prediction, *Sankhya, Ser. A*, 37, 239-256.

Potthoff, R. F. and Roy, S. N. (1964), A generalized multivariate analysis of variance model useful especially for growth curve problems, *Biometrika*, 51, 313-326.

Rao, C. R. (1965), The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves, *Biometrika*, 52, 447-458.

Rao, C. R. (1966), Covariance adjustment and related problems in multivariate analysis, in *Multivariate Analysis (Vol. I)*, ed. P. R. Krishnaiah, New York: Academic Press, 87-103.

Rao, C. R. (1987), Prediction of future observations in growth curve models, *Statistical Science*, 2, 434-471.

Reinsel, G. C. (1985), Mean squared error properties of empirical Bayes estimators in a multivariate random effects general linear model, *Journal of the American Statistical Association*, 80, 642-650.

Von Rosen, D. (1991), The growth curve model: a review, *Communications in Statistics, Theory and Methods*, 20, 2791-2822.

Yang, R. and Chen, M. H. (1995), Bayesian analysis for random coefficient regression models using noninformative priors, *Journal of Multivariate Analysis*, 55, 283-311.

三、Bayesian Estimation of Time Series Regressions with Applications

此乃與陳婉淑教授合作的成果。本文考慮在線性迴歸模型中，其誤差項之共變矩陣具有 P 階自我相關性質，且其限制在平穩的條件下對參數做估計問題。主要利用 Wise 在 1951 的結果導出正確 (exact) 的最大概似函數。然後用 MCMC 法對本數作貝氏估計。這是本計劃的研究範圍。擬送到國際期刊發表。其中英文摘要如下：

(一)、中文摘要

本文考慮在線性迴歸模式中，其誤差項之共變矩陣具有 P 階自我相關性質，且其限制在平穩的條件下對參數作估計問題。我們應用 Wise (1951) 之結果獲得此模式

之 exact 最大概似函數，進而對參數作貝氏推論。Chib (1993) 的論文亦對具有自我相關誤差項之迴歸模型推論，但該論文貝氏推論是在條件最大概似函數 (conditional likelihood) 下對參數估計，也就是假設起始值是已知的。我們應用 MCMC 法 (馬可夫鍊蒙地卡羅) 對參數作貝氏估計，採用 Tanner (1991) 的 griddy Gibbs 抽樣方法。文中以模擬研究與實例分析來闡明本法，而由模擬研究之結果得知我們的結果較 Chib (1993) 和 Chib and Greenberg (1994) 之結果精確，且由實例分析得知，許多文獻上的資料可用此模式描述得極佳。

(二)、英文摘要

We consider the parameter estimation of regression where error terms follow a stationary AR (P) process. Utilizing Wise (1951), we obtained the exact likelihood and the subsequent posterior distribution of the parameters. Posterior distributions are obtained using MCMC. The result obtained compared favorably with Chib (1993) and Chib and Greenberg (1994).

(三) 本研究成果之參考文獻如下

Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis Forecasting and Control*, 2nd edition, San Francisco: Holden-Day.
Carlin, B. P., Gelfand, A. E., and Smith, A. F. M. (1992). Hierarchical Bayesian of Changepoint problems, *Appl. Statist.*, 41, 2459-2482.
Carlin, B. P. and Chib, S. (1995). Bayesian Model Choice via Markov Chain Monte Carlo Methods, *Journal of the Royal Statistical Society*, B57, 473-484.
Chen, C. W. S. and Lee, J. C. (1995). Bayesian Inference of Threshold Autogressive Models, *Journal of Time Series Analysis*, 16, 483-492.
Chen, C. W. S., McCulloch, R. E. and Tsay,

R. S. (1997). A Unified Approach to Estimating and Modeling Linear and Nonlinear Time Series, *Statistica Sinica*, 7, 451-472.

Chib, S. (1993). Bayes regression with autogressive errors: A Gibbs sampling approach, *Journal of Econometrics*, 58, 275-294.

Chib, S., and Greenberg, E. (1994). Bayes Inference in Regression Model With ARMA(p,q) Errors, *Journal of Econometrics*, 64, 183-206.

Fogiel, M. (1980). *Handbook of Mathematical Formulas, Tables, Graphs, Functions, Transforms For Mathematicians, Scientists, Engineers.*

Gelman, A., and Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7, 457-472.

Geman, S., and Geman, D. (1984). Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images, *IEEE Transactions in Pattern Analysis and Machine Intelligence*, 6, 721-741.

George, E. I. And McCulloch, R. E. (1993). Variable Selection via Gibbs Sampling, *J. Am. Statist. Assoc.*, 88, 881-889

Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-Based Approaches to Calculating Marginal Densities, *Journal of the American Statistical Association*, 85, 398-409.

Harvey, A. (1981), *Time Series Models* (Philip Allan, London).

Hastings, W. K. (1970). Monte Carlo Sampling Methods Using Markov Chains and Their Application, *Biometrika*, 57, 97-100.

Lung, G. M. and Box, G. E. P. (1978). On a Measure of Lack of Fit in Time Series Models, *Biometrika*, 65, 297-303.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. (1953). Equations of State Calculations by Fast

Computing Machines, Journal of Chemical Physics, 21, 1987-1091.

McCulloch, R. E. and Tsay, R. S. (1994b). Statistical Inference of Economic Time Series via the Switching Models, J. Time Ser. Anal, 15, 523-539.

Tanner, M. A. (1991). Tools for Statistical Inference. Springer-verlag, New York.

Wise, J. (1951). The Autocorrelation Function And Spectral Density Function, (London School of Economics).

Zeger, S. L. and Karim, M. R. (1991). Generalized Linear Models With Random Effects: A Gibbs Sampling Approach, J. Am. Statist. Assoc., 86, 79-86.