C
001.642
HUB

IMPROVED METHODS OF MIXED ESTIMATION IN LINEAR REGRESSION MODEL

A PROJECT REPORT PRESENTED BY

0 JUL 2007

MARIATHAS HENSLEY HUBERT

To the Board of Study in Statistics and Computer Science of the

POSTGRADUATE INSTITUTE OF SCIENCE

in partial fulfillment of the requirement

for the award of the degree of

MASTER OF PHILOSOPY

of the

UNIVERSITY OF PERADENIYA

SRI LANKA

2006

607462

IMPROVED METHODS OF MIXED ESTIMATION IN LINEAR REGRESSION MODEL

M. H. Hubert

Department of Economics and Management
Faculty of Business Studies
Vavuniya Campus of the University of Jaffna
Vavuniya
Sri Lanka

In this research work the estimation of coefficient vector and the superiority of the derived estimators in a multiple linear regression model under several circumstances are discussed. If only the sample information of the standard multiple linear regression model is available, then the estimator due to GAUSS or AITKEN is most favorable, that is, Ordinary Least Square Estimator (OLSE) is the best linear unbiased estimator of the parameter vector, when the regressor matrix is assumed to be full column rank matrix. However, in many estimation problems, often one can give auxiliary information or prior information in addition to the sample, which can be either exact or stochastic, about the unknown parameter via the routes of postulation, experimentation or past sample data which improves the linear models. Therefore, in the presence of stochastic prior information in addition to the sample, Thiel and Goldberger (1961) introduce a Mixed Regression Estimator (MRE) procedure for the parameter vector in the linear regression model. Several statisticians compared these two estimators MRE and OLSE and showed the superiority of the MRE over the OLSE under certain conditions.

Apart from this, recently, Liu Kejian (1993) introduced an alternative ridge type biased estimator to combat near multicolinearity, which is called Liu Estimator of the parameter vector, based on OLSE. Selahattin Kaciranlar, Saullah Sakallioglu, Fikiri Akdenz, George Styan and Werner (1999) introduced another new biased estimator called Restricted Liu Estimator (RLE) by augmenting the Restricted Least Squares Estimator (RLSE) and showed the superiority of the Restricted Liu Estimator (RLE) over the OLSE and Liu Estimator (LE).

In this research a new biased estimator called Stochastic Restricted Liu Estimator (SRLE) for the parameter vector was introduced and consequently its efficiency was discussed. In particular, stochastic properties of the newly introduced Stochastic Restricted Liu estimator were discussed and also the necessary and sufficient conditions for the superiority of the Stochastic restricted Liu estimator (SRLE) over the other alternatives were derived in the mean squared error matrix sense as well as scalar mean squared error sense by considering the two cases; correctly specified stochastic restrictions and incorrectly specified stochastic restrictions.

Generally, it is possible to divide every, stochastic prior information, which consists of more than one row, into two or more subsets of information. On this regard attention was also focused on the Stochastic Restricted Liu Estimators based on to competing stochastic restrictions which outperform Liu estimator with respect to the mean squared error matrix criterion. Two estimators were proposed and efficiency properties of this estimation procedure were analyzed. Conditions for superiority of one Stochastic Restricted Liu Estimator over another were derived with respect to the mean squared error matrix criterion.

Special attention was also given to the use of correct and incorrect prior information in the estimation of regression coefficients when the regression model is misspecified due to exclusion of some relevant explanatory variables. Analysis of misspecification was extended to Liu estimator, and the superiority of Liu Estimator over Ordinary Least square Estimator was discussed. The results were further extended to the Stochastic Restricted Liu Estimator based on correct and incorrect prior information, and the resulting estimator was compared with that Liu estimator. It demonstrated that the stochastic restricted Liu estimator is still superior to the Liu estimator with respect to the mean squared error matrix criterion under certain conditions.

When collecting data, one is confronted with the problem of incomplete data sets. In such cases missing data may have a strong influence on the statistical analysis of the remaining data set. Several methods have been developed for handling missing values in the regression matrix, and the efficiency of the corresponding estimators has been investigated. In this study Stochastic Restricted Liu Estimator (SLRE) was applied under the complete

case analysis for analyzing incomplete data sets, and the resulting estimators were compared.

Prediction of a study variable is an important aspect of regression analysis. As far as linear regression as the underlying model is concerned, the monograph of Bibby and Toutenburg (1978) was devoted to the problem of best predictors as well as to investigation of regions where one predictor improves another. One of the main aims of this research work was to examine the superiority of the Stochastic Restricted Liu Predictor (SRLP) over the Liu predictor (LP), and concluded that there were situations where the Stochastic Restricted Liu predictor outperforms the Liu predictor with respect to the mean squared error matrix criterion even the model was misspecified.

Finally, MATLAB software package for windows was used to analyze some theoretical results obtained in this study using the data set on Portland cement considered by Woods, Steinour and Starke (1932).