International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-6, August, 2019

On OLS Estimation of Stochastic Linear Regression Model

B. Mahaboob, B.Venkateswarlu, K.A. Azmath, C. Narayana, J. Peter Praveen

Abstract: ***This paper mainly discusses the formulation of stochastic linear statistical model and its assumptions and finally explores an important aspect namely the Ordinary Least Squares (OLS) estimation of stochastic linear regression model. In addition to these inference in stochastic linear regression model is also presented here. Nimitozbay et.al [1], in their paper proposed the weighted mixed regression estimation of the coefficient vector in a linear regression model with stochastic linear restrictions binding the regression coefficients. In 1980, P.A.V.B. Swamy et.al proposed a linear regression model where the coefficient vector is a weekly stationary multivariate stochastic process and that model provides a convenient representation of a general class of non-stationary processes. They proposed prediction and estimation methods which are linear and easy to compute. Daojiang et.al [2] in 2014, in their paper depicted an innovative estimation technique to the multicollinearity in statistical model which is linear in the case of existence of stochastic linear constraints on the parameters and a very different estimation technique was presented by mixing the OME and PCR estimator also known as SRPC regression estimator. In 2014, Shuling Wang et.al [3] in their paper proposed some diagnostic methods in restricted stochastic statistical models which are linear. Gil Gonjalez et.al [4], in 2007, in their paper, derived the LSEs for the simple linear statistical model and examined them from a theoretical perspective.***

Index Terms: ***Regression coefficient, Data matrix, error vector, Homoscedasticity, OLS estimator, t-test statistic. OME(Ordinary Mixed Estimator),PCR(Principal Components Regressor), SRPC (Stochastic Restricted Principal Components),LSE(Least Square Estimator).***

I. INTRODUCTION

The words "Stochastic", 'Probability', 'Uncertainty', 'Random', 'Likelihood', Significance and Chance are synonym words. The words are "Deterministic', 'Sure', certain and Non-random are antonym words. A deterministic model predicts a single outcome from a given set of situations. A stochastic model predicts a set of possible outcomes weighted by their probabilities or likelihoods. Stochastic model building can be used for several reasons. The main objectives for stochastic model building are: (i) Developing scientific understanding (ii) Test the effect of changes is a system; and (iii) aid decision making including, (a) Tactical decisions by managers and (b) Strategic decisions by planners. The various tools for stochastic modelling have been studied by several American and British mathematicians and statisticians. In India, a few mathematicians and statisticians have made contributions to the problems of stochastic models. In stochastic model building, a great deal of research has been directed to establishing the functional relationships among different variables such as dependent independent and error variables. Stochastic model building in new and very fascinating fertile filed of research in statistics.

Revised Manuscript Received on October 30, 2019.
* Correspondence Author
Dr B. Mahaboob*, Department of Mathematics, Koneru Lakshmaiah Educational Foundation, Vaddeshwaram, Guntur, A.P., INDIA. Email.id: mahaboob@kluniversity.in
Dr B.Venkateswarlu, School of Advanced Science (SAS), Vellore Institute of Technology, Vellore, Tamilnadu Corresponding Author Email id: venkatesh.reddy@vit.ac.in
Dr K.A. Azmath, Department of Mathematics, Sree Vidyanikethan Engineering College, Tirupati, Andhra Pradesh. Email id: kaajmath@gmail.com
Dr C. Narayana, Mathematics Department, Sri Harsha institute of P.G Studies, Nellore Email id:nareva.nlr@gmail.com
J.Peter Praveen, Department of Mathematics, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India, 522502 Email id: jppraveen17kluniversity.in

II. FORMULATION OF STOCHASTIC LINEAR REGRESSION MODEL

Suppose Y depends on free variables $X_2, X_3, \ldots, X_k$ and an error variable (or disturbance term) $\in$ linearly. For a sample of 'n' observations on each of these variables, one may write a k-variable stochastic linear regression model as

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots\ldots + \beta_k X_{ki} + \in_i ,$$

values of i are from 1to n ………………(2.1)

where, Y_i : i[th] entry on a dependent variable Y, X_{ji} is the i[th]entry on j[th]free variable, j=2,3,…….k, β_j is the parameter of X_j , and $\in_i$ is i[th] observation on the error variable $\in$.In the matrix notation one way express the mode (1.1) as

$$Y_{nx1}=X_{nxk}\,\beta_{kx1}+\in_{nx1}. \ldots\ldots\ldots\ldots\ldots (2.2)$$

where

$$Y=\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, X=\begin{bmatrix} 1 & X_{21} & X_{31} & \ldots\ldots & X_{K1} \\ 1 & X_{22} & X_{32} & \ldots. & X_{K2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{2n} & X_{3n} & \ldots.. & X_{Kn} \end{bmatrix}$$

$$\beta=\begin{bmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_k\end{bmatrix} \text{ and } \in=\begin{bmatrix}\in_1\\ \in_2\\ \vdots\\ \in_n\end{bmatrix}$$

Y is observation vector of order nx1; X is fixed known coefficient matrix (Data matrix) of order (nxk), β is parametric vector of order (kx1) and $\in$ is error vector of order (nx1).

III. ASSUMPTIONS OF STOCHASTIC LINEAR STATISTICAL MODEL

The various crucial assumptions of stochastic linear model are given by

(i) Assumption of Linearity: The vector of sample observations may be written as a linear function of sample observations on (k-1) independent variables x's and an error variable $\in$.

i.e., $Y = X\beta + \in$

(ii) Assumption of Unbiasedness:

$$E[\in]=0 \Rightarrow \begin{bmatrix}E(\in_1)\\ E(\in_2)\\ \vdots\\ E(\in_n)\end{bmatrix}=\begin{bmatrix}0\\ 0\\ \vdots\\ 0\end{bmatrix}$$

(or) $E(Y) = X\beta$

(iii) Assumption of Homoscedasticity and Non Autocorrelation:

$$E(\in\in') = \sigma^2 I_n$$

$$\Rightarrow \begin{bmatrix}Var(\in_1) & Cov(\in_1,\in_2) & \cdots & Cov(\in_1,\in_n)\\ & Var(\in_2) & \ldots & Cov(\in_2,\in_n)\\ & \vdots & & \vdots\\ & \vdots & \ldots\ldots & \vdots\\ & & & Var(\in_n)\end{bmatrix}=\begin{bmatrix}\sigma^2 & 0 & \ldots\ldots & 0\\ 0 & & & \\ \vdots & & & \vdots\\ 0 & & & \sigma^2\end{bmatrix}$$

(or) $Cov(\in_i \in_j) = \sigma^2, \forall i = j$ are from 1 to n

$= 0$ For unequal i and j

i.e., the errors are unobservable and they are pair wise uncorrelated and have the same unknown variance σ^2.

iv. Assumption of No Multi-collinearly: Rank (X) = K, K<n. There is no linear dependence among the independent variables.

v. Assumption of Non Stochastic Data Matrix: X is a non-stochastic
matrix of fixed known coefficients.

vi. Assumption of Normality:

$$\in \sim N(0, \sigma^2 I_n)$$

The statistical model which is linear is $Y = X\beta + \in$ along with the above assumptions (i) to (vi) is known as the standard or Gauss-Markoff or Classical Stochastic Linear Regression model.

IV. OLS ESTIMATION PROCEDURE OF STOCHASTIC LINEAR REGRESSION MODEL

Standard stochastic linear regression model in matrix notation as

$$Y_{nx1} = X_{nxk}\beta_{kx1} + \in_{nx1} \quad \text{.... (4.1)}$$

Let $\hat{\beta}$ be a column vector with k rows of β and the residual vector may be defined as

$$e = \left[Y - X\hat{\beta}\right] \quad \text{.... (4.2).}$$

By the principle of least squares estimation, one may minimize the residual sum of squares e^1e with respect to $\hat{\beta}$. This

$$e^1e = \left(Y - X\hat{\beta}\right)^1\left(Y - X\hat{\beta}\right)$$

gives

$$= Y^1Y - 2\hat{\beta}X^1Y + \hat{\beta}^1X^1X\hat{\beta}$$

and

$$\frac{\partial(e^1e)}{\partial\hat{\beta}} = 0 \Rightarrow -2X^1Y + 2X^1X\hat{\beta} = 0$$

or

$$X^1X\hat{\beta} = X^1Y \quad \text{........... (4.3)}$$

The equation in (4.6.3) is known as a group of OLS normal equations which ensure the consistency from the assumption Rank (X) = K, one may obtain the ordinary least square estimation to β as

$$\hat{\beta} = (X^1X)^{-1}X^1Y \quad \text{.... (4.4)}$$

(i)

$$E(\hat{\beta}) = (X^1X)^{-1}X^1\, E(Y) = (X^1X)^{-1}X^1X\,\beta = \beta \quad \text{.... (4.5)}$$

i.e., $\hat{\beta}$ is linear unbiased estimator of β

(ii)

$$V(\hat{\beta}) = E\left[\hat{\beta} - E(\beta)\right]\left[\hat{\beta} - E(\hat{\beta})\right]^1$$

$$= E\left[(X^1X)^{-1}X^1\in\right]\left[(X^1X)^{-1}X^1\in\right]^1$$

$$= (X^1X)^{-1}X^1E(\in\in^1)X(X^1X)^{-1}$$

$$\Rightarrow V(\hat{\beta}) = \sigma^2(X^1X)^{-1} \quad \text{.... (4.6)}$$

i.e., The OLS estimator $\hat{\beta}$ has mean vector β and covariance matrix $\sigma^2(X^1X)^{-1}$

Retrieval Number F7930088619/2019©BEIESP
DOI: 10.35940/ijeat.F7930.088619
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Remarks:

1. From the Gauss –Markoff Theorem, it can be shown that the OLS estimator $\hat{\beta}$ has smaller sampling covariance matrix than any other linear unbiased estimator of β. Thus $\hat{\beta}$ is the Best Linear Unbiased Estimator (BLUE) of β.
2. One can obtain an unbiased estimator of unknown variance σ^2 as $\hat{\sigma}^2 = \left[\frac{e^1 e}{n-k}\right]$, where $e = \left[Y - X\hat{\beta}\right]$ is called ordinary least squares vector residual

V. INFERENCE IN LINEAR STOCHASTIC STATISTICAL MODEL

Consider the linear standard stochastic statistical model as

$$Y_{nx1} = X_{nxk}\beta_{kx1} + \in_{nx1} \quad \text{…. (5.1)}$$

$$\in \sim N\left(O, \sigma^2 I_n\right) \quad \text{….. (5.2)}$$

Since, $\in$ follows multivariate normal distribution, one can see that

$$\hat{\beta} \sim N\left(\beta, \sigma^2 \left(X^1 X\right)^{-1}\right) \quad \text{…. (5.3)}$$

$$\Rightarrow \hat{\beta}_j \sim N\left[\beta_j, \sigma^2 C_{jj}\right], j = 1, 2, .., k \quad \text{……….. (5.4)}$$

Where C_{jj} is the j^{th} element in the principal diagonal matrix $\left(X^1 X\right)^{-1}$.

It can be shown that $\left[\frac{(n-k)\hat{\sigma}^2}{\sigma^2}\right] \sim \chi^2_{(n-k)}$ …. (5.5)

Also, $\hat{\beta}$ and $\frac{(n-k)\hat{\sigma}^2}{\sigma^2}$ are independently distributed each other. To test $H_0 : \beta_j = 0$, one may use the t-test statistic as

$$t = \frac{\hat{\beta}_j}{\hat{\sigma}\sqrt{C_{jj}}} \sim t_{(n-k)}, \quad j = 1, 2, …,k \quad \text{…. (5.6)}$$

VI. CONCLUSION

In the above research study an attempt has been made by describing some important aspects namely formulation of linear statistical stochastic model, assumptions of stochastic linear regression model, Ordinary Least Squares (OLS) estimation of stochastic linear regression model and inference in stochastic linear regression model. In the context of future research one can extend this study to Generalized LSE of Linear Stochastic Statistical Model with non-spherical errors.

REFERENCES

1. Ninetozbay, SelahattinKacirnlar (2017), “Estimation in a linear regression model with stochastic linear restrictions: A new two parameter weighted mixed estimator”, Journal of statistical computation and simulation, vol. (88), issue (9).
2. P.A.V.B Swamy, P.A Trinsley (1980), “Linear prediction and estimation methods for regression models with stationary stochastic coefficients”, Journal of Economics, Vol. (12), issue (2), Pp: 103-142.
3. Pinsky, M.A. and Samuel Kanlin, (2011), “An Introduction to Stochastic Modeling”, Academic Press, 978-0-12-381416-6.
4. Berry, J.S. and Houston, S.K., (1995), “Mathematical Modelling”, Jordan Hill, 0-340-61404-8.
5. Byron J.T. Morgan (2008), “Applied stochastic modelling”, CRC press, 978-1-58488-666-2
6. Taylor, H.M. and Samuel karlin, (1998), “An Introduction to Stochastic Modeling”, Academic Press, London, 978-0-12-684887-8.
7. Venkateswarlu .B et.al (2011), “On problems of statistical modelling”, International journal Engineering science and Technology, Vol. (3), issue (6), Pp: 4914-4920
8. ShulingWang, Man Liu, et.al (2014), “On diagnostics in stochastic restricted linear regression models”, Open journal of statistics, vol. (4), Pp; 757-764.
9. Gil Gonzalez et.al (2007), “Least squares estimation of linear regression models for convex compact random sets”, Advances in Data Analysis and classification, Vol.(1), issue(1), Pp:67-81

AUTHORS PROFILE

Dr.B.Mahaboob has been working as Associate professor in the Department of Mathematics, Koneru Lakshmaiah Education Foundation Deemed to be University, Vaddeswaram, Guntur dist, A.P for 2 years. He completed his Ph.D. from Sri Venkateswara University in 2016, Tirupati, A.P, India. He stood as university **Gold Medalist** in his post-graduation (M.Sc. Mathematics) during the academic year 1996-1998.He has published 25 research articles in reputed international journals. His areas of research are Nonlinear Regression Models, Stochastic models, Econometric models and Operations Research. His total experience in teaching is 21 years.

Dr.B.Venkateswarlu has been working as Senior Assistant professor in the Department of Mathematics, School of Advanced Sciences, Vellore Institute of Technology (Deemed to be University), Vellore for 5 years. He completed his Ph.D. from Sri Venkateswara University in 2008, Tirupati, A.P, India. He completed his post-graduation (M.Sc. Mathematics) during the academic year 1996-1998. He has published 30 research articles in reputed international journals. His areas of research are Operations Research, Fuzzy set theory, linear algebra and Data mining. His total experience in teaching is 21 years.

Dr.K.A.Ajmath, Professor of Mathematics in Sree Vidyanikethan Engineering College, A.Rangampet. He is a member of Andhra Pradesh Society of Mathematical Sciences, He is also a member of Cryptography Research Society in India. Dr.K.A.Ajmath is also discharging duties as Student Welfare Officer for Sree Vidyanikethan Educational Institutions. Dr.K.A.Ajmath published a number of research papers in the field of cryptography in prestigious national and international journals.