iJRASET
International Journal For Research in
Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 5 **Issue:** XII **Month of publication:** December 2017

**DOI:**

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887
Volume 5 Issue XII December 2017- Available at www.ijraset.com

# A Brief Review on Inference In General Linear Models

Dr. R.V.S.S Nagabhushana Rao[1], B.Venkateswarlu[2], V.Srinivasulu[3]
[1] Assistant Professor, Department of Statistics, Vikrama Simhapuri University, Nellore - 524 320
[2] Assistant Professor, Department of Mathematics, Vikrama Simhapuri University, Nellore - 524 320
[3] Research Scholar, Department of Statistics, Sri Venkateswara University,Tirupati - 517502

***Abstract: In the Linear regression models ,under the assumptions of Spherical errors ( Homogeneous assumptions ) ,the OLS estimators of parameters have a number of optimum properties that they are BLUE's .It involves the double assumption that the error variance in constant at each observation point and that the error variances at all possible pairs of observation points are zero .It describes various inferential problems to the classical linear regression model .It specifies the linear model with the assumptions besides the OLS estimation of parameters of linear model. It also contains the procedure of testing the general linear hypothesis. Restricted least squares estimators are used for linear regression model.***
***Keywords: Linear Model, Homoscedasticity, Multicollinearity, Ordinary least Squares, Gauss –mark off Theorem, BLUE.***

## I. INTRODUCTION

It is clear from the three variable Linear Model that it would be excessively tedious and complicated to build up to the general case of K variable multiple regression model in a step wise responses.

Fortunately, the use of matrix algebra we have a compact and powerful way of treating problem and easily shown that detailed results of the two variable linear models or special cases of a few simple matrix formulae. Suppose present theory suggests has a dependent variable y is a linear with constant term or disturbances term.

Write the functional for is

$$Y=\beta_0+\beta_1 X_1+\beta_2 X_2\cdots\cdots \beta_k X_k+\in$$

Where $\in$ denotes a statistic variable or disturbances term with some specified probability and $\beta_0$ indicates the intercept made by the regression plane and the vertical y-axis and $\beta_0,\beta_1,\cdots\cdots\beta_k$ are regression coefficients .The purpose of the term $\in$ is that there may be a basic and unpredictable element of randomness in human responses .

### *A. Assumptons Of General linear Model*

The first and basic assumptions of the model is that the vector of sample observations on a dependent variable Y may be expressed as linear combination of the sample observations on independent variable X and a disturbance term , $\in$

$$i.e., Y = X_{nxk}\beta_{kx1} + \in_{nx1}$$

Where
$$Y = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_k \end{bmatrix}_{nx1} ; X = \begin{bmatrix} x_{11} & x_{12} & . & . & x_{k1} \\ x_{12} & x_{22} & . & . & x_{k2} \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{1n} & x_{2n} & . & . & x_{kn} \end{bmatrix}_{nxk} ; \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ . \\ . \\ \beta_k \end{bmatrix}_{kx1} ; \in = \begin{bmatrix} \in_1 \\ \in_2 \\ . \\ . \\ \in_k \end{bmatrix}_{nx1}$$

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887
Volume 5 Issue XII December 2017- Available at www.ijraset.com

Here all the data in terms of deviations of their model , their sample means in the case of general linear model ,the main problem is to obtain and estimates of the unknown parameter vector $\beta$ . To get estimates for $\beta$ , some further crucial assumptions as follows :

$$E(\in) = 0 \quad i.e.,\ E(y) = X\beta$$
$$\Rightarrow E(\in) = 0\ \forall\, i = 1,2,\cdots,k$$

i.e., the $\in_i^{,s}$ are variables with 0 means .This is called the assumption of unbiasedness .

$$E\left(\in\in^1\right) = \sigma^2 I$$
$$i.e.,\ \text{cov}\left(\in_i \in_j\right) = 0 ; \forall\, i \neq j = 1,2,\ldots,n \qquad 0$$
$$= \sigma^2 ; \forall\, i = j = 1,2,\ldots,n$$

1) Each $\in_i$ has the same variance.
2) All disturbances are pair wise independent and uncorrelated.
3) The assumption 1 is called the assumption of Homoscedasticity disturbance or homogeneous on is disturbance variances. The second assumption is called the assumption of uncorrelated disturbances.
4) $\ell(X) = k, k < \text{n}$ This assumption may be record as assumption of linear dependence of the explanatory variables.
5) X is a non-stochastic matrix i.e.; X is matrix of fixed known coefficients.
6) There is no error involved in the variables X and Y.
7) The disturbance vector $\in$ has a multivariate normal distribution .

The assumptions 2,3and 7 may be combined in the single statement such that $\in$ follows normal distribution with mean ‘0’ and variance $\sigma^2$ .

The general linear model with above assumptions is called a STANDARD GENERAL LINEAR MODEL.

### *B. Ordinary Least Squares Estimation*

The most frequently used estimating technique for the general linear model is called Ordinary Least Squares estimation.
Consider general linear model as

$$Y = X\beta + \in \text{(2.1)}$$

Where $\hat{\beta}$ is a column vector of the estimates of $\beta$ , then define a vector of disturbances as

$$e = Y - X\hat{\beta}$$

Therefore $e = Y - \hat{Y}$

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

We have $e^1 e = \left(Y - X\hat{\beta}\right)^1 \left(Y - X\hat{\beta}\right) \cdots\cdots (\ 2.2\ )$

By the least squares principle for choosing $\hat{\beta}$ the residual sum of squares is to be minimized .

Thus , $\dfrac{\partial\left(e^1 e\right)}{\partial\hat{\beta}} = 0 \Rightarrow -2X^1 Y + 2X^1 Y\hat{\beta} = 0$

$$\Rightarrow \quad X^1 X\beta = X^1 Y \text{ (2.3)}$$

From the assumption (4) , $E\left(\hat{\beta}\right) = \beta$

Thus , $\hat{\beta}$ is a linear unbiased estimator of $\beta$ .

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887*
*Volume 5 Issue XII December 2017- Available at www.ijraset.com*

According to Gauss – Mark off the orem , $\hat{\beta}$ has a smaller variance than any other linear unbiased estimator of $\beta$ . $\rho(X) = k_a$

we have $\hat{\beta} = \left(X^1 X\right)^{-1}\left(X^1 Y\right)$ ( 2.4 )

The equations (2.3) are called the ordinary least squares normal equations.

The estimate $\hat{\beta} = \left(X^1 X\right)^{-1}\left(X^1 Y\right)$ is the OLS estimator for $\beta$ .

The mean and variance of $\hat{\beta}$ are : $E\left(\hat{\beta}\right) = \beta$ ( 2.5)

Thus, the OLS estimators are unbiased. Now consider

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

And $V\left(\hat{\beta}\right) = \sigma^1\left(X^1 X\right)^{-1}$ ( 2.6 )

Here $\sigma^2$ is unknown parameter . The variance of $\hat{\beta}$ may be obtained by taking the $i^{th}$ term from the principle diagonal of $\left(X^1 X\right)^1$ an multiplied by $\sigma^2$ .

The half diagonal elements of $\left(X^1 X\right)^{-1}$ multiplied by $\sigma^2$ gives sampling variances .

### C. *Ordinary Least Squares Estimators Are Blue's*

The most important result in least squares theory is that no other linear unbiased estimators can have smaller sampling variance than those of the OLS estimators.

$\hat{\beta} = \left(X^1 X\right)^{-1}\left(X^1 Y\right)$ Which is a linear function of $Y$ .

### D. *Testing General Linear hypothesis*

Linear hypothesis is a statement about the parameters of a linear model which is in the form of linear function of parameters. General linear hypothesis consists of a set of linear hypothesis about the parameters of a linear model.

Consider a classical linear regression model as

$$Y_{nx1} = X_{nxk}\, \beta_{kx1} + \in_{nx1}$$

A general linear hypothesis consists of a set of $q(\leq k)$ linear restrictions about the elements of $\beta$ can be expressed in the matrix notation as (2.7)

Where R: (qxk) m

$H_0 : R_{qxk}\beta_{kx1} = r_{qx1}$ atrix of known constants with full row rank;

To test $H_0$ , first of all , the unknown $\beta$ in ( 2.7) may be replaced by the OLS estimator $\hat{\beta}$ and obtaining $R\hat{\beta}$ .

The sampling distribution of $R\hat{\beta}$ is derived as follows :

$$E\left[R\hat{\beta}\right] = R\beta$$
$$V\left[R\hat{\beta}\right] = E\left[R\left(\hat{\beta}-\beta\right)\right]\left[R\left(\hat{\beta}-\beta\right)\right]^1 \quad (2.8)$$

Since , $\hat{\beta}$ follows multivariate normal distribution ,

$$R\hat{\beta} \approx N\left[R\beta, \sigma^2 R\left(X^1 X\right)^{-1} R^1\right]$$
$$or\ R\left[\hat{\beta}-\beta\right] \approx N\left[0, \sigma^2 R\left(X^1 X\right)^{-1} R^1\right] \quad (2.9)$$

If the hypothesis ( 2.7 ) is true , one can replace $R\beta$ in ( 2.9) by $r$ ,obtaining

$$R\left[\hat{\beta}-r\right]\approx N\left[0,\sigma^2 R\left(X^1X\right)^{-1}R^1\right]$$
$$\Rightarrow \left(R\hat{\beta}-r\right)^1\left[\sigma^2 R\left(X^1X\right)^{-1}R^1\right]^{-1}\left(R\hat{\beta}-r\right)\approx \chi_q^2$$

Here $\left[R\left(X^1X\right)^{-1}R^1\right]$ is positive definite matrix ,

Since n independently with $\hat{\beta}$ and hence independently with $R\hat{\beta}$ .

Thus ,to test the general linear hypothesis , $H_0 : R\beta = r$ , the F-statistic is given by

$$F = \frac{\left(R\hat{\beta}-r\right)^1\left[R\left(X^1X\right)^{-1}R^1\right]^{-1}\left(R\hat{\beta}-r\right)/q}{\frac{e^1e}{n-k}} - F_{q,(n-k)}$$

*E. Restricted Least Squares Estimator For $\beta$*

Consider the general linear hypothesis $H_0 : R\beta = r$

If $H_0$ is not rejected ,then one may re-estimate the model , in cooperating the linear restrictions in the estimations process .This estimation process involves the efficiency of estimators .The restricted least squares ( RLS) estimator $\beta^*$ satisfying the set of $q(\le k)$ restrictions in $R\beta = r$ can be obtained by minimizing the residual sum of squares with respect to $\beta^*$ subject to the constraints $R\beta^* = r$

It is given by $\beta^* = \hat{\beta} + \left(X^1X\right)^{-1}R^1\left[R\left(X^1X\right)^{-1}R^1\right]^{-1}\left(r - R\beta\right)$ (2.10 )

The mean vector and covariance matrix of $\beta^*$ is given by

$$E\left(\beta^*\right) = \beta + \left(X^1X\right)^{-1}R^1\left[R\left(X^1X\right)^{-1}R^1\right]\left(r - R\beta\right)$$
$$and\ V\left(\beta^*\right) = \sigma^2\left[\left(X^1X\right)^{-1} - \left(X^1X\right)^{-1}R^1\left\{R\left(X^1X\right)^{-1}R^1\right\}^{-1}R\left(X^1X\right)^{-1}\right] \quad (2.11)$$

Some alternative expressions for F-statistic for testing the general linear hypothesis $H_0 : R\beta = r$ using RLS estimator is given by

$$(i)\, F = \frac{\left(\beta^* - \beta\right)^1\left(X^1X\right)\left(\beta^* - \beta\right)/q}{\frac{e^1e}{n-k}} \sim F_{q,n-k}$$

$$(ii)\, F = \frac{e^{*1}e^* - e^1e}{\frac{e^1e}{n-k}} \sim F_{q,n-k}$$

$$(iii)\, F = \frac{\left[R_{OLS}^2 - R_{RLS}^2\right]/q}{\left[1 - R_{OLS}^2\right]/n-k} \sim F_{q,n-k}$$

Where $e^{*1}e^*$ = Restricted least squares residual sum of squares

$e^1e$ = Unrestricted least squares residual sum of squares

$R_{RLS}^2$ and $R_{OLS}^2$ are respectively the $R^2$ - values obtained from the OLS and RLS regressions .

It should be noted that $(i)\, F = R_{OLS}^2 \ge R_{RLS}^2$ *and* $(ii)\, F = e^{*1}e^* \ge e^1e$

 

## II. CONCLUSION

In the present paper, linear regression model under the assumption of spherical errors, the parameters of ordinary least squares estimates are BLUE. various inferential problems to the classical linear regression model are discussed. Under the procedure of testing the general linear hypothesis, Restricted least squares estimators are used.

## REFERENCES

[1] Firth, D (1991),"Generalized Linear Models", in statistical theory and modeling, by D.V. Hinkley,N. Reid and E.J. Snell,London: Chapman and Hall,PP: 55-82
[2] Johnston, J., (1984),"EconometricMethods", Mc Graw Hill International Editions,Singapore.
[3] Judge, G.G., Griffiths, W.E., Hill, R.C., andLee, T.C. (1980), "The theory and practice of Econometrics", John Wiley, New York.
[4] Kmenta, J., and Gilbert, R.F., (1968), "Small sample properties of Alternative Seemingly Unrelated Regression Equations", JASA ,63, 1180-1200.
[5] Rao C.R. (1973),"Linear Statistical Inference and its applications", John Wiley, NewYork.
[6] Srivastava, V.K., andGiles,D.E.A. (1987), "Seemingly Unrelated Regression Equations models: Estimation and Inference", marcel DekkerInc. NewYork.
[7] Theil, H. (1971),"Principles of Econometrics", John Wiley,NewYork.
[8] Tsimikas, J.V., (2012),"Inference in generalized Linear regression models", Computational Statistics a Data analysis, Volume 56, Issue 6, PP: 1854-1868.
[9] Zellner, A (1963),"Estimators for Seemingly Unrelated Regression Equations", Journal of the American Statistical Association,58, 348-368.
[10] Zellner, A and Haung D.S(1962),"Further properties of Efficient estimators for Seemingly Unrelated Regression Equations", International Economic Review, 3, 300-313.