International Journal of Statistics and Applied Mathematics 2018; 3(2): 588-597

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2018; 3(2): 588-597

www.mathsjournal.com
Received: 05-01-2018
Accepted: 06-02-2018

SS Rajkumar
Block Health Statistician, Govt primary health centre, Kolathur, Salem, Tamil nadu, India

C Mani
Head, Department of Statistics, S.V. Arts College, Tirupati, Andhra Pradesh, India

GSGN Anjaneyulu
Professor, Department of Mathematics, VIT, Vellore, Tami lnadu, India

A Chinna kesavulu
Research Scholar, Department of Statistics, S.V. University, Tirupati, Andhra Pradesh, India

P Srivyshnavi
Senior Assistant Professor, Department of CSE, SPMVV Engineering College, Tirupati, Andhra Pradesh, India

J Prabhakara Naik
Assistant Professor, SSBN Degree & PG College Autonomous, Anantapur, Andhra Pradesh, India

P Balasiddamuni
Rtd. Professor, Department of Statistics, S.V. University, Tirupati, Andhra Pradesh, India

Correspondence
SS Rajkumar
Block Health Statistician, Govt primary health centre, Kolathur, Salem, Tamil nadu, India

Tests for parameter constancy and predictive accuracy in linear model by using student zed residuals

SS Rajkumar, C Mani, GSGN Anjaneyulu, A Chinna kesavulu, P Srivyshnavi, J Prabhakara Naik and P Balasiddamuni

Abstract
Statistical models are fitted for a variety of reasons. One important reason is that of trying to observe relationships between the variables. A second reason for fitting model is to assist in carrying out a prediction; it is used in some selection processes. Thus, after having fitted a regression model from the sample of observations, one may often centers on some specific value of independent variable, and one may required to predict the value of the dependent variable likely to be associated with the specific value of the independent variable. This specific value of independent variable may lie within the range of sample independent values or more frequently, one may concerned with predicting dependent variable for a value of independent variable outside the sample observations. One of the independent criteria for an estimated depression equation is that it should have relevance for data outside the sample data used in the estimation. This criterion is embodied in the notion of parameter constancy that is the parametric vector should apply both outside and within the sample data. Parameter constancy can be examined by using a test of predictive accuracy.
This research paper proposes tests for parameter constancy and predictive accuracy with different parameter vectors in the forecast period by using internally studentized residuals.

Keywords: parameter constancy, predictive accuracy in linear model, student zed residuals

1. Introduction
Modelling is the Heart of almost all fields of Science. Modelling is part of the Problem of solving process, that problems are often solved by seeking alternative ways of looking at the problem. Generally, Modelling can avoid or reduce the need for costly, undesirable, or impossible experiments with the real world.
Statistical models are fitted for a variety of reasons. One important reason is that of trying to uncover causes by studying relationships between the variables. A second reason for fitting models, over and above prediction and explanation is to examine and test scientific hypothesis. Statistical Modelling is a tool used essentially as part of problems solving. The purpose of having a Statistical model is almost always to assist in some problem solving activity. The model assists in carrying out a prediction; it is used in some selection processes.
A Statistical model is basically a Mathematical model concerns with the functional relationships between the variables in terms of mathematical equations such as either in the form of a Set of linear equations (Linear Regression Model) or in the form of a Set of Non-linear equations (Non-linear Regression model). A large number of problems in Applied Regression Analysis are concerned with the inferential aspects including estimating the parameters and testing the hypothesis about the parameters of both the linear and non-linear regression models. Regression techniques designed to take into account random disturbances or error. They can be estimated by using various types of Residuals. The regression model and its various types of residuals is perhaps the most widely used technique in applied statistical research.

2. Some Important Types of Residuals
The Residuals are consistent estimates of the true errors or disturbances in the Regression models. Large values of Residuals indicate the bad predictions in the sample. A Residual may

be large, where the Statistician entered the corresponding observation wrongly. Alternatively, it may be an influential observation or an outlier, which differs from the data points in the sample. In other words, it may be further away from the estimated Regression equation than the other data points. Removing this influential observation or outlier from the sample may change the estimates and the Regression significantly. Thus, one should always plot the Residuals to check the data, identify the influential observations or outliers and check the violations of assumptions about the errors underlying the Regression model.

Residuals can be examined, especially in the cross section data for omission of relevant variables or incorrect functional form under model specification. Residuals have vital role in dealing with the various inferential problems in the Statistical Modelling. Particularly, the various Diagnostic methods to find certain inadequacies in the Statistical Modelling are based on Residuals.

3. The OLS, Studentized and Predicted Residuals

Consider the Classical Linear Regression model as

$$y_{n\times 1} = X_{n\times k}\beta_{k\times 1} + \varepsilon_{n\times 1} \tag{2.1}$$

Such that $\mathrm{E}[\varepsilon]=0$ and $\mathrm{E}\left[\varepsilon\varepsilon^{I}\right]=\sigma^{2}I_{n}$. Here, the error vector ε consists of unknown errors which are assumed to be uncorrelated. Generally, ε can be estimated by the Ordinary Least Squares Residual vector e, which is given by

$$e = y - X\hat{\beta} \tag{2.2}$$

Where $\hat{\beta}=\left(X^{I}X\right)^{-1}X^{I}y$ is the Best Linear Unbiased Estimator (BLUE) of β One may have,

$$\hat{y} = X\hat{\beta} = X\left[\left(X^{I}X\right)^{-1}X^{I}y\right] \tag{2.3}$$

And $e=\left[y-\hat{y}\right]=\left[y-X\left(X^{I}X\right)^{-1}X^{I}y\right]$

$$e=\left[I-X\left(X^{I}X\right)^{-1}X^{I}\right]y \tag{2.4}$$

$$e=\left[I-V\right]y$$

Where $V = X\left(X^{I}X\right)^{-1}X^{I}$ (2.5)

Generally, $V = X\left(X^{I}X\right)^{-1}X^{I}$ is an important matrix and is known as HAT matrix. One may write,

$y=\left[V+I-V\right]y=Vy+\left(I-V\right)y$ Or $y=X\left(X^{I}X\right)^{-1}X^{I}y+e$

$$\Rightarrow y = X\hat{\beta}+e \tag{2.6}$$

Mean and Variance of e are given by

$$\mathrm{E}\left[e-\varepsilon\right]=\mathrm{E}\left[\left(I-V\right)y-\varepsilon\right]$$

$$=\mathrm{E}\left[\left(I-V\right)\left(X\beta+\varepsilon\right)-\varepsilon\right]$$

$$=\mathrm{E}\left[\left(I-V\right)\varepsilon-\varepsilon\right],\left[\because\left(I-V\right)X=0\right]$$

$$=\mathrm{E}\left[M-I\right]\varepsilon,$$

Where $M=\left[I-X\left(X^{I}X\right)^{-1}X^{I}\right]$ is a Symmetric Idempotent matrix,

$$\Rightarrow \mathrm{E}[e-\varepsilon]=(M-I)\mathrm{E}[\varepsilon]=0 \quad (2.7)$$

or $\mathrm{E}[e]=0 \quad [\because \mathrm{E}(\varepsilon)=0]$

Thus, the Ordinary Least Squares Residual vector e is a Linear Unbiased Estimator of ε . Also,

$$\mathrm{var}(e)=\mathrm{E}\left[ee^{I}\right]$$

$$=\mathrm{E}\left[M\varepsilon\varepsilon^{I}M^{I}\right]$$

$$=M\mathrm{E}\left[\varepsilon\varepsilon^{1}\right]M^{I}$$

$$=\sigma^{2}MM^{I}$$

Or $\mathrm{var}(e)=\sigma^{2}M$ (2.8)

Hence, each of the Residual vector has zero mean and they all have different variances and are all correlated with each other.

Since, $e=My,$ which is a Linear combination of elements of y , if the ε follows normal distribution then e will also have normal distribution. If the Linear Regression model has intercept term then for the Least Squares fitting, the sum of the Residuals be zero. Write the HAT matrix as

$$V=\left(\left(v_{ij}\right)\right)$$

Where, $v_{ij}=x_{i}^{I}\left(X^{I}X\right)^{-1}x_{j}$ (2.9)

And $v_{ii}=x_{j}^{I}\left(X^{I}X\right)^{-1}x_{i}$ (2.10)

Here, x_{i}^{I} and x_{j}^{I} are the i[th] row and j[th] row of the data matrix X respectively.

Now, the variance, covariance and correlation coefficients of the Residuals are given by,

$$\mathrm{var}(e_{i})=\sigma^{2}(1-v_{ii}) \quad (2.11)$$

$$\mathrm{cov}(e_{i}e_{j})=-\sigma^{2}v_{ij} \quad (2.12)$$

And $r_{e_{i}e_{j}}=\dfrac{-v_{ij}}{\sqrt{(1-v_{ii})(1-v_{jj})}}$ (2.13)

Also, V is (nxn) symmetric Idempotent matrix and Rank (V) = Rank (X) = k, k<n, where k is the number of parameters including intercept parameter and

$$Trace(V)=\sum_{i=1}^{n}v_{ii}=Rank(X)=k \quad (2.14)$$

Further, one may have, $\sum_{j}v_{ij}^{2}=v_{ii}$ (2.15)

And $\sum_{i}v_{ij}=\sum_{j}v_{ij}=1$ (2.16)

Each v_{ii} must fall in the interval, $0\le v_{ii}\le 1$

The notion of v_{ii} measures the distance from the point x_i to the center of the data, and cases with unusual values for the independent variables will tend to have large values of v_{ii}

Generally, $\operatorname{var}(e_i)$ will be small whenever v_{ii} is large, so cases with X_i near $\bar{X}$ will be fit poorly and cases with x_i far from $\bar{X}$ will be fit well, which is undesirable

The improved set of Residuals known as Studentized Residuals, can be obtained by scaling, so that the cases will larger v_{ii} get larger scaled Residual, and cases with smaller v_{ii} get smaller scaled Residuals. By Scaling, one may divide each of the Residuals by an estimate of its Standard deviation. In Other words, these Residuals are the Standerdized version of the Residuals that does not depend upon σ^2 and, the v_{ii} 's scale Quantities Margolin (1977) used the term Studentized Residuals instead of Standardized Residuals. David (1981) distinguished two types of Studentized Residuals namely (i) Internally Studentized Residuals and (ii) Externally Studentized Residuals.

The Internally Studentized Residuals are given by,

$$e_i^* = \frac{e_i}{\hat{\sigma}\sqrt{1-v_{ii}}}, i = 1, 2\cdots, n \tag{2.17}$$

Under the assumption of the normality of error vector ε in the Linear Regression model $y = X\beta + \varepsilon$, say $\varepsilon \sim N\left(0, \sigma^2 I_n\right)$, the Internally Studentized Residuals $\left[\frac{e_i^{*2}}{n-k}\right]$ follows Beta distributions with parameters $\left[\frac{1}{2}, \frac{n-k-1}{2}\right]$. One may obtain, $\mathrm{E}\left[e_i^*\right] = 0$ and $\operatorname{var}\left[e_i^*\right] = 1$ and

$$\operatorname{cov}\left[e_i^*, e_j^*\right] = \frac{-v_{ij}}{\sqrt{(1-v_{ii})(1-v_{jj})}}, i \neq j = 1, 2, \cdots, n \tag{2.18}$$

The Externally Studentized Residuals are given by

$$e_i^{**} = \frac{e_i}{\hat{\sigma}_{(i)}\sqrt{1-v_{ii}}}, i = 1, 2\cdots, n \tag{2.19}$$

Where, $$\hat{\sigma}_{(i)}^2 = \left[\left(n-k^I\right)\hat{\sigma}^2 - \frac{e_i^2}{1-v_{ii}}\right] \Big/ n-k^I-1 \tag{2.20}$$

Or $$\hat{\sigma}_{(i)}^2 = \hat{\sigma}^2\left[\frac{n-k^I-e_i^{*2}}{n-k^I-1}\right] \tag{2.21}$$

And $$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-k} \tag{2.22}$$

Here, Rank (V) = k^1

$\hat{\sigma}_{(i)}^2$ is an estimate of error variance σ^2 with i^{th} observation has been omitted from the estimation. Also, e_i^{**} follows t-distribution with ($n-k^1-1$) degrees of freedom. It can be shown that e_i^{**} is a monotonic transformation of e_i^* and a relationship between External and Internal Studentized residuals is given by,

$$e_i^{**} = e_i^* \left[\frac{n-k^I-1}{n-k^I-e_i^{*2}} \right]^{\frac{1}{2}} \tag{2.23}$$

In various diagnostic methods in Statistical Modelling, the predicted Residuals $e_{(i)}$ which is based on a fit to the data with the ith observation excluded, has been frequently used. Suppose that $\hat{\beta}_{(i)}$ be the Ordinary Least Square (OLS) estimator of β with ith observation has been excluded. Then the ith predicted Residual is defined by,

$$e_{(i)} = y_i - X_i^I \beta_{(i)}, i = 1, 2, \cdots, n \tag{2.24}$$

Sometimes, $e_{(i)}$ may be known as prediction error.

The relationships among Ordinary least squares, Studentized and predicted Residuals are given by,

(i) $e_{(i)} = \dfrac{e_i}{1-v_{ii}}, i = 1, 2, \cdots, n$ (2.25)

(ii) $e_i^* = \dfrac{e_{(i)}}{\hat{\sigma} / \sqrt{(1-v_{ii})}}$ (2.26)

(iii) $e_i^{**} = \dfrac{e_{(i)}}{\hat{\sigma}_{(i)} / \sqrt{(1-v_{ii})}}$ (2.27)

A widely used Criterion for Model selection known as "Predicted Residual Sum of Squares (PRESS)" based on $e_{(i)}$ is given by

$$PRESS = \sum e_{(i)}^2 \tag{2.28}$$

Alternatively, interms of Studentized Residuals, one way express PRESS as

$$PRESS = \sum e_i^{*2} \tag{2.29}$$

Or $PRESS = \sum e_i^{**^2}$ (2.30)

4. A Test for Parameter Constancy Using Studentized Residuals

After having estimated a Regression equation from the sample of observations, our interest often centers on some specific value of the independent variable, and one may required to predicting the value of dependent variable likely to be associated with the specific value of the regressor variable. This specific value of regressor may lie within the range of sample regressor values or more frequently, one may concerned with predicting dependent variable for a value of regressor outside the sample observations. One of the most important criteria for an estimated Regression equation is that it should have relevance for data outside the sample data used in the estimations. This criterion is embodied in the notion of parameter constancy, that is, that the parametric vector should apply both outside and within the sample data. Parameter constancy can be examined by using a test of predictive accuracy.

Under test of predictive accuracy one may divide the sample data of n observations into n_1 observations to be used for estimation and $n_2 = (n-n_1)$ observations to be used for testing the hypothesis of parameter constancy. In the case of time series data, one may use the first n_1 observations for estimation and the last $n_2 = (n-n_1)$ for testing procedure. In the case of cross-section data, it could be partitioned by the values of a size variable. There are no hard and fast rules for determining the relative sizes of n_1 and n_2. Generally 5 to 10 percent of the observations may be used for testing procedure.

Consider the Classical Regression model as

$$y_{n\times1} = X_{n\times k}\beta_{k\times1} + \varepsilon_{n\times1} \tag{3.1}$$

Such that $\varepsilon \sim N\left(0,\sigma^2 I_n\right)$

Partitioning the data of n observations into n_1 and n_2 observations, and express X, y and ε as

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad and \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

Where X_1 is $n_1$xk; X_2 is $n_2$xk, y_1 is $n_1$xk; y_2 is $n_2$xk; ε_1 is n_1x1 and ε_2 is n_2x1 matrices.

By using n_1 observations, the OLS estimator of β is given by

$$\hat{\beta}_1 = \left(X_1^I X_1\right)^{-1} X_1^I y_1 \tag{3.2}$$

Now, prediction of y_2 by using $\hat{\beta}_1$ is given by,

$$\hat{y}_2 = X_2 \hat{\beta}_1 \tag{3.3}$$

The prediction errors vector is defined as

$$\mathrm{U} = \left[y_2 - \hat{y}_2\right] = y_2 - X_2 \hat{\beta}_1 \tag{3.4}$$

$$\Rightarrow \mathrm{U} = X_2\beta + \varepsilon_2 - X_2\hat{\beta}_1 = \varepsilon_2 - X_2\left(\hat{\beta}_1 - \beta\right) \tag{3.5}$$

Consider $\hat{\beta}_1 = \left(X_1^I X_1\right)^{-1} X_1^I y_1$

$$= \left(X_1^I X_1\right)^{-1} X_1^I \left(X_1\beta + \varepsilon_1\right)$$
$$= \beta + \left(X_1^I X_1\right)^{-1} X_1^I \varepsilon_1$$

Or $\left(\hat{\beta}_1 - \beta\right) = \left(X_1^I X_1\right)^{-1} X_1^I \varepsilon_1$

$$\therefore \mathrm{U} = \varepsilon_2 - X_2\left[\left(X_1^I X_1\right)^{-1} X_1^I \varepsilon_1\right] \tag{3.6}$$

$$\Rightarrow \quad \text{(i)} \quad \mathrm{E}\left[\mathrm{U}\right] = 0$$

$$\text{(ii)} \quad \mathrm{var}\left(\mathrm{U}\right) = \mathrm{E}\left[\mathrm{U\text{-}E}\left(\mathrm{U}\right)\right]\left[\mathrm{U\text{-}E}\left(\mathrm{U}\right)\right]^I$$

$$= \mathrm{E}\left[\varepsilon_2 - X_2\left(X_1^I X_1\right)^{-1} X_1^I \varepsilon_1\right]\left[\varepsilon_2 - X_2\left(X_1^I X_1\right)^{-1} X_1^I \varepsilon_1\right]^I$$
$$= \mathrm{E}\left[\varepsilon_2\varepsilon_2^I\right] + X_2\left(X_1^I X_1\right)^{-1} X_1^I \mathrm{E}\left[\varepsilon_1\varepsilon_1^I\right] X_1\left(X_1^I X_1\right)^{-1} X_2^I$$

$$\mathrm{var}\left(\mathrm{U}\right) = \sigma^2 I_{n_2} + X_2\left\{\left(\sigma^2 I_{n_2}\right)\left(X_1^I X_1\right)^{-1}\right\} X_2^I \tag{3.7}$$

$$\begin{bmatrix} \because \mathrm{E}\left[\varepsilon_1\varepsilon_2^I\right] = 0 \quad and \\ \mathrm{E}\left[\varepsilon_i\varepsilon_i^I\right] = \sigma^2 I_{n_i}, i = 1,2 \end{bmatrix}$$

$$\text{var}(\text{U}) = \sigma^2 \left[I_{n_2} + X_2 \left(X_1^I X_1 \right)^{-1} X_2^I \right] \tag{3.8}$$

Since, $\varepsilon \sim N(0, \sigma^2)$, the prediction error vector U follows normal distribution $N(0, \text{var}(\text{U}))$ and

Hence,

$$\text{U}^I \left[\text{var}(\text{U}) \right]^{-1} \text{U} \sim \chi^2_{n_2} \tag{3.9}$$

Where $e_1^{*I} e_1^*$ is the Internally Studentized residual sum of squares from the estimated regression equation. Further, the two χ^2 Statistics $\text{U}^I \left[\text{var}(\text{U}) \right]^{-1} \text{U}$ and $\frac{\left[e_1^{*I} e_1^* \right]}{\sigma^2}$ are distributed independently.

The test statistic for testing the hypothesis of parameter constancy is given by

$$\text{F} = \frac{\text{U}^I \left[\text{var}(\text{U}) \right]^{-1} \text{U} / n_2}{\left[\left[e_1^{*I} e_1^* \right] / \sigma^2 \right] / (n_1 - k)} \sim \text{F}_{(n_2, n_1-k)} \tag{3.11}$$

$$\text{F} = \frac{\text{U}^I \left[I_{n_2} + X_2 \left(X_1^I X_1 \right)^{-1} X_2^I \right]^{-1} \text{U} / n_2}{e_1^{*I} e_1^* / (n_1 - k)} \sim \text{F}_{(n_2, n_1-k)} \tag{3.12}$$

For larger values of the F test Statistic, the hypothesis of parameter constancy (i.e. the same parametric vector β applies within and outside the estimation data) may be rejected.

Remark: In this test, it has been assumed that σ^2 is the same in each subset of data. Otherwise, the errors are Heteroscedastic errors and F test Statistic may be redefined.

5. A Test for Predictive Accuracy with a Different Parametric Vector in the Forecast Period Using Studentized Residuals

Consider the Classical Linear Regression model as

$$y_{n\times1} = X_{n\times k}\beta_{k\times1} + \varepsilon_{n\times1} \tag{4.1}$$

such that $\varepsilon \sim N(0, \sigma^2 I_n)$, Partitioning the data of n observations into n_1 and n_2 observations and allowing for the possibility of different Regression coefficient vectors for two subsets of observations gives

$$y_{1_{n_1\times1}} = X_{1_{n_1\times k}} \delta_{k\times1} + \varepsilon_{1_{n_1\times k}} \tag{4.2}$$

$$y_{2_{n_2\times1}} = X_{2_{n_2\times k}} \text{r}_{k\times1} + \varepsilon_{2_{n_2\times k}} \tag{4.3}$$

Where $\delta = \beta$, write the model (4.3) as

$$y_2 = X_2\text{r} + \varepsilon_2 = X_2\delta + X_2(\text{r} - \delta) + \varepsilon_2 \tag{4.4}$$

Or $y_2 = X_2\delta + \eta + \varepsilon_2$ (4.5)

Where $\eta = X_2(\text{r} - \delta)$

If $\eta = 0$ then $r = \delta$ and the Regression coefficient vector is constant over the estimation and forecast periods. Now, the Regression model can be compactly expressed as

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ X_2 & I_{n_2} \end{bmatrix} \begin{bmatrix} \delta \\ \eta \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \quad (4.6)$$

From (4.1) and (4.6), are may write the augmented matrix X and (X^1 X) as

$$X = \begin{bmatrix} X_1 & 0 \\ X_2 & I \end{bmatrix} \text{ And } X^I X = \begin{bmatrix} X_1^I X_1 + X_2^I X_2 & X_2^I \\ X_2 & I \end{bmatrix}$$

$$\Rightarrow \left(X^I X\right)^{-1} = \begin{bmatrix} \left(X_1^I X_1\right)^{-1} & -\left(X_1^I X_1\right)^{-1} X_2^I \\ -X_2^I \left(X_1^I X_1\right)^{-1} & I + X_2 \left(X_1^I X_1\right)^{-1} X_2^I \end{bmatrix} \quad (4.7)$$

The OLS estimators of δ and η are given by

$$\begin{bmatrix} \hat{\delta} \\ \hat{\eta} \end{bmatrix} = \begin{bmatrix} \left(X_1^I X_1\right)^{-1} & -\left(X_1^I X_1\right)^{-1} X_2^I \\ -X_2^I \left(X_1^I X_1\right)^{-1} & I + X_2 \left(X_1^I X_1\right)^{-1} X_2^I \end{bmatrix} \begin{bmatrix} X_1^I y_1 + X_2^I y_2 \\ y_2 \end{bmatrix} \quad (4.8)$$

$$\text{Or } \begin{bmatrix} \hat{\delta} \\ \hat{\eta} \end{bmatrix} = \begin{bmatrix} \left(X_1^I X_1\right)^{-1} X_1^I y_1 \\ y_2 - X_2 \left(X_1^I X_1\right)^{-1} X_1^I y_1 \end{bmatrix} \quad (4.9)$$

$$\Rightarrow \hat{\eta} = y_2 - X_2 \hat{\delta} \quad (4.10)$$

The hypothesis of constant β or δ is equivalent to the null hypothesis of $\eta = 0$

The null hypothesis H: $\eta = 0$ can be tested by examining the joint significance of the last $n_2 = (n-n_1)$ variables in the augmented Regression equation:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ X_2 & I \end{bmatrix} \begin{bmatrix} \hat{\delta} \\ \hat{\eta} \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \quad (4.11)$$

Where e_1 and e_2 are the OLS residual vectors.

Since, $\hat{\eta} = y_2 - X_2 \hat{\delta}$ gives the prediction errors vector, and $y_2 = X_2 \hat{\delta} + \hat{\eta} + e_2$, one may obtain $e_2 = 0$. Now, the residual sum of squares from the fitting equation (4.6) is $e_1^I e_1$. By defining Internally Studentized residual vector e_1^* by using the OLS residual vector e_1 the F-test Statistic for testing the hypothesis of predictive accuracy is given by

$$F = \frac{\hat{\eta}^I \left[\operatorname{var}(\hat{\eta})\right]^{-1} \hat{\eta} \big/ n_2}{\left[\left[e_1^{*I} e_1^*\right] \big/ \sigma^2\right] \big/ (n_1 - k)} \sim F_{(n_2, n_1 - k)} \quad (4.12)$$

or $$F = \frac{\hat{\eta}^I \left[I_{n_2} + X_2 \left(X_1^I X_1 \right)^{-1} X_2^I \right]^{-1} \hat{\eta} \Big/ n_2}{e_1^{*^I} e_1^* \Big/ (n_1 - k)} \sim F_{(n_2, n_1-k)} \qquad (4.13)$$

Remark: Consider the Regression equation (4.11) as unrestricted least squares Regression equation with the unrestricted OLS residual sum of squares $e_1^I e_1$. One may obtain the unrestricted internally Studentized residual sum of squares by using $e_1^I e_1$ as $\left[e_1^{*^I} e_1^* \right]_{UR}$

By setting $\eta = 0$ the restricted OLS regression equation can be written as

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \hat{\beta}_R + e_R \qquad (4.14)$$

Remark: Consider the Regression equation (4.11) as unrestricted least squares Regression equation with the unrestricted OLS residual sum of squares $e_1^I e_1$. One may obtain the unrestricted internally Studentized residual sum of squares by using $e_1^I e_1$ as $\left[e_1^{*^I} e_1^* \right]_{UR}$

By setting $\eta = 0$ the restricted OLS regression equation can be written as

$$F = \frac{\left[\left(e^{*^I} e^* \right)_R - \left(e^{*^I} e^* \right)_{UR} \right] \Big/ n_2}{\left[e_1^{*^I} e_1^* \right]_{UR} \Big/ (n_1 - k)} \sim F_{(n_2, n_1-k)} \qquad (4.15)$$

One may reject the hypothesis of parametric constancy or predictive accuracy if F value exceeds a preselected critical value.

6. Conclusions

In the Statistical Modelling, the various substantial portions of the modeling come under the headings of linear regression models, nonlinear regression models and Sets of linear regression models or Seemingly Unrelated Regression Equations (SURE) models. A large number of problems in Statistical Modelling are concerned with the inferential aspects including estimation of parameters and testing the hypotheses about the parameters of the various types of regression models.

In the present research study, a test for parameter constancy in the classical linear statistical model has been proposed by using Studentized residuals. After having estimated a regression equation from the sample of observations, our interest often centers on some specific value of the independent variable and one may require to predict the value of dependent variable likely to be associated with the specific value of the regressor variable. One of the most important criteria for an estimated regression equation is that it should have relevance for data outside the sample data used in the estimation. Parameter constancy means that the parametric vector should apply both outside and within the sample data. Parameter constancy can be examined by using a test of predictive accuracy.

Also, a test for predictive accuracy with different parameter vector in the forecast period has been developed by using Internally Studentized residuals. In testing parameter constancy, it has been assumed that error variance is the same in each subset of data. But, in the case of testing for predictive accuracy, parameter vector is different in the forecast period.

7. Reference

1. Brown RL, Durbin J, Evans JM. Techniques for testing the constancy of Regression Relationships over time", Journal of the Royal Statistical Society Series-B. 1975; 37:149-192.
2. Chambers MJ, McGarry JS. Modeling cyclical behavior with differential-difference equations in an unobserved components framework", Econometric Theory. 2002; 18(2):387-419.
3. Cox DR. Tests of Separate Families of Hypotheses, in: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, 1961.
4. Durga Prasad S, Balasiddamuni P, Ramesh Mummineni, Statistical Inference In Time Series Regression Models, LAP, Lambert Academic Publishers, Germany, 2013.
5. Efron B, Johnstone I, Hastie T, Tibshirani R. Least angle regression, Ann. Statist. 2004; 32:407-499.

6. Gallant AR. Seemingly Unrelated Nonlinear Regressions, Journal of Econometrics. 1975; 3:35-50.
7. Gallant AR. Nonlinear Statistical Models, Wiley, New York, 1987.
8. Hari Babu O, Balasiddamuni P, Ramana Murthy B. Inferential Aspects of Regression Models, LAP, Lambert Academic Publishers, Germany, 2013.
9. Kadane JB, Lazar NA. Methods and Criteria for Model Selection", Journal of the American Statistical Association. 2004; 99(465):279-290.
10. Lang S, Adebayo S, Fahremir L. Bayesian Semiparametric Seemingly Unrelated Regression, Proceedings in Computational Statistics (Ed. By W. Hardle and B. Ronz), Physika- Verlag, Heidel berg, (195-200), 2002.
11. Narayana P, Balasiddamuni P, Subbarami Reddy C. Statistical Inference in Sets of Linear Regression Models, LAP, Lambert Academic Publishers, Germany, 2013.