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

Residual Measures for the Time Series Predictions

R.V.S.S. Nagabhushana Rao[1], J. Prabhakara Naik[2], V. Munaiah[3], K.Vasu[4], G. Mokesh Rayalu[5]
[1]*Assistant professor, Department of Statistics, Vikrama Simhapuri University, Nellore*
[2]*Lecturer in Statistics, SSBN Degree & PG College(Autonomous), Anantapur*
[3]*Lecturer, Department of Statistics, PVKN Govt College, Chittoor*
[4]*Assistant Professor, Vidya Educational Institutions, Yadamari, Chittoor*
[5]*Assistant Professor, Department of Mathematics, School of Advanced sciences ,VIT, Vellore*

Abstract: Residuals are differences between the predicted output from the model and the measured output from the validation data set or measured output from the actual data. It is the part that describes the model capability of explaining the data. Thus, residuals represent the portion of the validation data not explained by the model. They are very important measures for any statistician who builds models and relationships on the available variables. Proper measure of residual will clearly distinguish between a good model and a bad model.

I. INTRODUCTION

Any forecasting procedure is always measured through the behaviour of the residuals it generates out of prediction. It should supposed to satisfy several set of properties and assumptions in which there are some predefined consequences are there for violation of these assumptions. In popular regression forecasting, there are several diagnostic measures are associated with the residuals like DFFIT, DFBETA etc. to study the efficiency of the fitted regression model. They are well used to know how the fitted model will explain the variability in the data and also used to decide which variables are causing higher magnitude of residuals and so on.

In the literature, we can find several methods and applications based on residual in the forecasting procedures like graphical measures of errors, deviated measures of errors etc. and in the current study, we tried to cover the maximum related concepts on residual theory.

II. TYPES OF RESIDUALS

As the number of forecasting methods are increasing and developing, the types of residuals used for analysis are also increasing in the literature. In this paper we discuss about types of errors in the model forecasting.

A. Raw residuals

Raw residuals are simple difference between the predicted valued and its actual value. It is symbolically given by

$$e_i = y_i - \hat{y}$$

Raw residuals are used to study the general behaviour of error terms in the models. They are used in residual plots to verify the assumptions of the error terms in any forecasting process. They are also used to calculate the hat matrix which is further used to calculate several diagnostic measures like leverage, DFFIT, DFBETA etc.

The relationship between error term and hat matrix is given by

$$\hat{e} = (I - H)y$$

where $H=X(X^1X)^{-1}X^1$ and

$$x = \begin{pmatrix} 1 & x_{1,1} & x_{1,2} & \dots & x_{1,k} \\ 1 & x_{2,1} & x_{2,2} & \dots & x_{2,k} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & x_{n,1} & x_{n,2} & \dots & x_{n,k} \end{pmatrix}$$

B. Standardized residuals

A standardized residual is a raw residual that is divided bits standard deviation:

$$\hat{\varepsilon}_i^2 = \frac{Y_i - \hat{Y}_i}{\sqrt{Var\left[Y_i - \hat{Y}_i\right]}} = \frac{\hat{\varepsilon}_i}{\sigma^2(l - h_{ii})}$$

In practical, the value of σ^2 is unknown and hence this residual calculation is possible for the history data only. Because of these standardized residula always follows zero mean and constant variance, we can very well use them in plotting several diagnostoc plots.the standardised residuals are better suited than the raw residuals for checking the assumptions on the random errors. The standardised residuals have the same distributions as the random errors: N $(0,\sigma^2)$. However, the standardised residuals are not, in general, independent. But if the values of the diagonal elements h_{ii} of the hat-matrix h are reasonably small, the standardized residuals are `nearly' independent.

C. Studentized residuals.

In general, value of σ^2 is unknown and hence calculation of standardized residual is difficult some times, Astudentized residual is a raw residual that is divided by its estimated standard deviation.

The expression for studentized residual is then

$$\frac{\hat{\varepsilon}_i}{\hat{\sigma}\sqrt{1 - h_{ii}}}$$

Where $\hat{\sigma}$ is an appropriate estimate of σ

If the estimate of the standard deviation is based on the same data that were used in fitting the model, the residual is also called an internally student zed residual. Then the usual estimate of σ^2 is

$$\hat{\sigma}^2 = \frac{1}{n-m}\sum_{j=1}^{n}\hat{\varepsilon}_j^2$$

Where m is the number of parameters in the model.

If the estimate of the residual's variance does not involve the i^{th} observation, it is called an externally student zed residual. In this case we have to use the estimate as

$$\sigma_{(i)}^2 = \frac{2}{n-m-1}\sum_{\substack{j=1 \\ j \neq i}}^{n}\hat{\varepsilon}_j^2$$

Studentized residuals are playing key role in model validation activities. The plots of studentised residual against predicted values or explanatory variables are very much useful in deciding about model sufficiency.

D. PRESS Residuals

The prediction error residuals (PRESS residuals) area defined by

$$e_{(i)} = y_i - \hat{y}_{(i)}$$

where $\hat{y}_{(i)}$ is the fitted value of the i^{th} response based on all observations except the i^{th} one.

Here, the variance of the i^{th} PRESS residual is

$$Var(e_{(i)}) = Var\left(\frac{e_i}{1-h_{ii}}\right) = \frac{\sigma^2}{1-h_{ii}}$$

To calculate standardized PRESS residual we need to use the expression

$$\frac{e_{(i)}}{\sqrt{Var(e_{(i)})}} = \frac{e_i}{\sqrt{\sigma^2(1-h_{ii})}}$$

PRESS residuals are applied when we want to know the impact of one particular observation on the model and to decide about inclusion of a variable in the model.

E. Scaled residuals

A scaled residual is simply a raw residual divided by a scalar quantity that is not an estimate of the variance of the residual. For example, residuals divided by the standard deviation of the response variable are scaled and referred to as Pearson or Pearson-type residuals:

$$\hat{\varepsilon}_{ic} = \frac{Y_i - \hat{Y}_i}{\sqrt{\hat{V}ar[Y_i]}}$$

In generalized linear models, where the variance of an observation is a function of the mean μ and possibly of an extra scale parameter, Var[Y] = a(μ)Φ, the Pearson residual is

$$\hat{\varepsilon}_{ic} = \frac{Y_i - \hat{\mu}_i}{\sqrt{a[\hat{\mu}_i]}}$$

Because the sum of the squared Pearson residuals equals the Pearson X^2statistic:

$$X^2 = \sum_{i=1}^{n} \hat{\varepsilon}_{iP}^2$$

When the scale parameter Φparticipates in the scaling, the residual is also referred to as a Pearson-type residual:

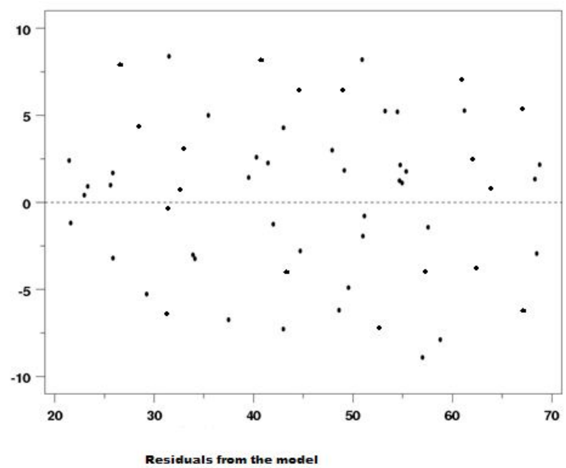
$$\hat{\varepsilon}_{iP} = \frac{Y_i - \hat{\mu}_i}{\sqrt{a(\hat{\mu})\phi}}$$

III. PLOTS FOR RESIDUALS

The plotting of residual values against various observations gives an idea of different patterns and relationships among the analysis variables. Each of the plot has its own significant result and they are explained briefly in the following sections.

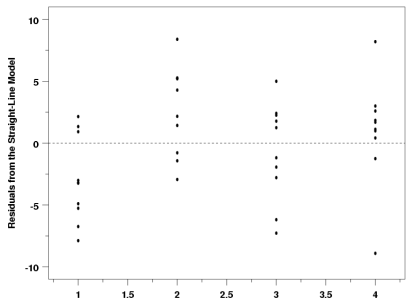
A. Residuals versus the predictor variable

The scatter plots of residuals against predictor values or residuals against omitted varaibles in the model will give good idea on the model sufficiency. In other words, these scatter plots are useful to know whether the currently included variables are sufficient to explain the predicted variables or not. It also gives an idea to add or remove the omitted variables in the models. The general decision of model sufficiency is exact randomly distribution of points in the graph. Plots of the residuals versus other predictor variables, or potential predictors that exhibit systematic structure indicate that the form of the function can be improved in some way.

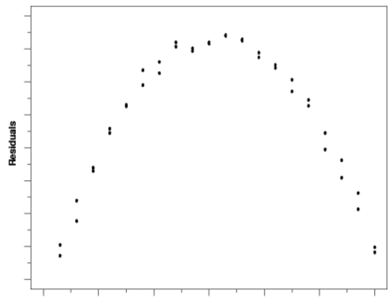


Plot of residuals Vs predictor values

The residuals from the above plot shows clearly that there are no systematic pattern in the distribution of points and the model is sufficient and it is a good model. In the other hand the following are some of the examples of patterns that can indicate that model fitted is not good. Either we need to modify the functional form of variables or need to decide about the variables included in the model.



Plot of observed effects

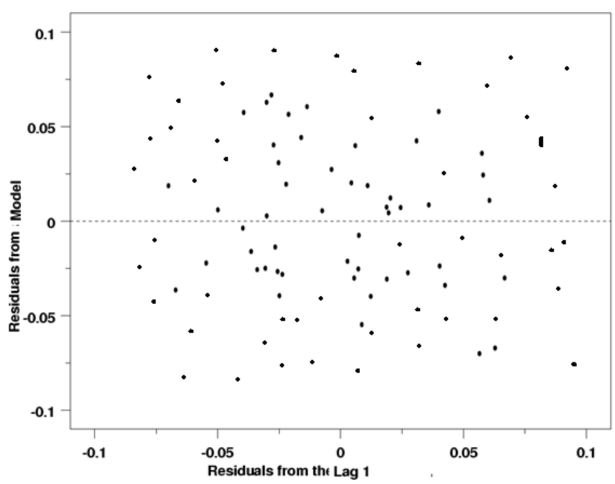


Plot of heteroscadasticity

B. *Residual lag plot*

The lag plot of the residuals is another tool to decide about the independence of error terms. In this plot, the normal residuals are plotted against the lagged values of the residual and order of lag is determined based on the given scenario. If the errors are not independent, then the estimate of the error standard deviation will be biased and it leads to improper inferences about the process. The lag plot works by plotting each residual value versus the value of the successive residual (in chronological order of observation). The first residual is plotted versus the second, the second versus the third, etc. Because of the way the residuals are paired, there will be one less point on this plot than on most other types of residual plots

If the errors are independent, there should be no pattern or structure in the lag plot. In this case the points will appear to be randomly scattered across the plot in a scatter shot fashion. If there is significant dependence between errors, some sort of deterministic pattern will likely be evident. The following is plot residuals where they are seem to independent.



Residual lag plot

C. *Partial regression plot*

Partial regression plots are very much useful to study the effect of adding an independent variable in the existing model. (Given that one or more independent variables are already in the model). Partial regression plots are constructed from the following steps

1) Compute the residuals of regressing the response variable against the independent variables by omitting X_i

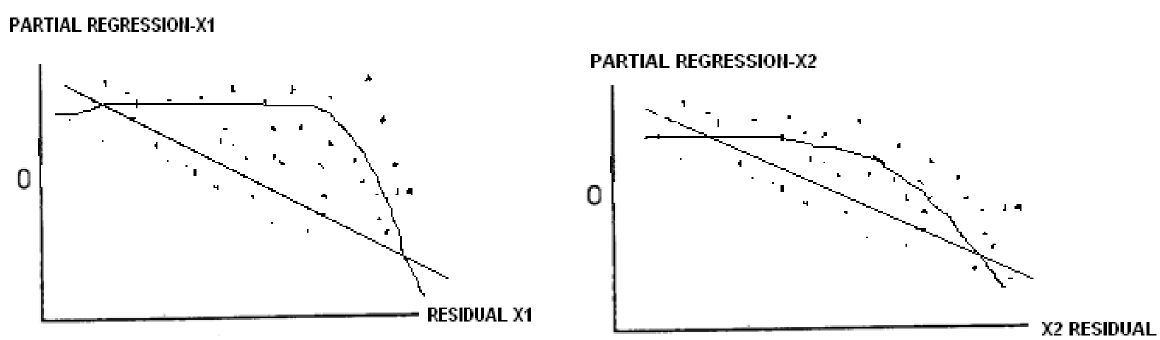
2) Compute the residuals from regressing X_i against the remaining independent variables.
3) Plot the residuals from (1) against the residuals from (2).

Mathematically, we can express this plot as $Y_{.[i]}$ versus $X_{i.[i]}$ where
$Y_{.[i]}$ = residuals from regressing Y (the response variable) against all the independent variables except X_i
$_{i.[i]}$ = residuals from regressing X_i against the remaining independent variables.

D. Properties

1) The least squares linear fit to this plot has the slope β_i and intercept zero.
2) The residuals from the least squares linear fit to this plot are identical to the residuals from the least squares fit of the original model (Y against all the independent variables including X_i).
3) The influences of individual data values on the estimation of a coefficient are easy to see in this plot.
4) It is easy to see many kinds of failures of the model or violations of the underlying assumptions (nonlinearity, heteroscedasticity, unusual patterns).

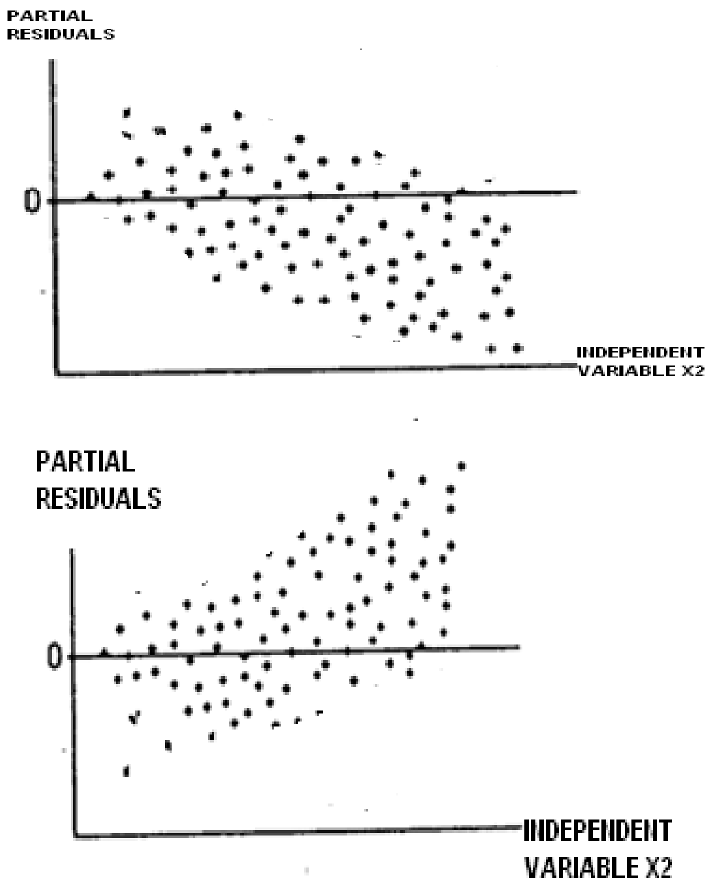
Partial regression plots are most commonly used to identify leverage points and influential data points that might not be leverage points. Partial residual plots are most commonly used to identify the nature of the relationship between Y and X_i (given the effect of the other independent variables in the model). Note that since the simple correlation between the two sets of residuals plotted is equal to the partial correlation between the response variable and X_i partial regression plots will show the correct strength of the linear relationship between the response variable and X_iThis is not true for partial residual plots. On the other hand, for the partial regression plot, the x axis is not X_i.
In the following partial regression plots we can conclude that influence of outliers is more in the case of X_1 rather than X_2.



Partial regression plots

E. Partial residual plots

When performing a linear regression with a single independent variable, a scatter plot of the response variable against the independent variable provides a good indication of the nature of the relationship. If there is more than one independent variable, things become more complicated. Although it can still be useful to generate scatter plots of the response variable against each of the independent variables, this does not take into account the effect of the other independent variables in the model.
One of the aims of performing a multiple regression is to determine each inde-pendent variable's ability to "explain" the dependent variable in the presence of the other independent variables. Often, the original data plots (y vs Xi) tell us little or nothing about the relationship between X, and (Y_1, X_1, X_2, * * *, X_{i-1} , X_{i+1}, .. , X_p). The partial residual plot explicitly displays this latter relation-ship while displaying deviations from the standard assumptions.
Partial residual plots attempt to show the relationship between a given independent variable and the response variable given that other independent variables are also in the model.
Partial residual plots are formed as:
Res + β_i* X_i versus X_i Where Res = residuals from the full model β_i = regression coefficient from the i^{th} independent variable in the full model, X_i = the i^{th} independent variable Consider the following two partial residual plots as



Partial residual plots

Both of the above type of graphs will indicate that variance is hugely changing with variable X_2. This means the variable X_2 has larger effect on the model.

A partial residual plot of following type will indicate that none of the least squares assumptions are violated with respect to particular independent variable.

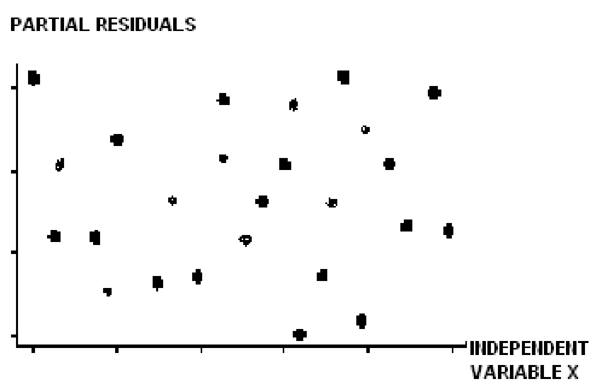


Fig. 5.10 Normal plot of partial residuals

IV. NUMERICAL MEASURES OF ERRORS

Numerical measures of errors play an important role in diagnostic study of time series analysis. The forecast accuracy can be identified and compared by using these numerical measures of errors only. The measures are most commonly dependent on the residual which is defined as the difference between actual and predicted value. The following are some of the important measures of the numerical measures of errors.

A. Mean Squared Error (MSE)

The mean squared error is useful to understand how close the predicted value is from its original value. It relay the concepts of bias, precision, and accuracy in statistical forecasting. The average of the square of the difference between the desired response and the actual system output is known as Mean Squared Error. The smaller the Mean Squared Error, the closer the forecast to the data. It is obtained from

$$MSE = \frac{\sum_{t=1}^{N} E_t^2}{N}$$

Where E= Residual measure or difference between predicted and original values
N= Total number of observations in the data.

B. Root Mean Square Error (RMSE)

The RMSE measures variance of the error. In other words the difference between forecast and corresponding observed values are squared and then averaged over the sample. Finally, the square root of the average is taken. Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors. This means the RMSE is most useful when we need to identify the large residuals in the data.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{N} E_t^2}{N}}$$

C. Mean Absolute Percentage Error (MAPE):

MAPE is considered as a measure of accuracy in the time series data and it is always expressed in percentage.` The absolute difference between actual and predicted value is summed and averaged to minimize the effect of bulk data. Then it is converted to percentage in its final form. The expression for MAPE is given by

$$MAPE = \sqrt{\frac{\sum_{t=1}^{N} \left| \frac{E_t}{Y_t} \right|}{N}}$$

V. CONCLUSIONS

Time series analysis provides tools for selecting the best suited model that can be used to forecast of future events. Modeling the time series is a statistical problem that involves some statistical tools like estimation technique and testing of hypothesis. Forecasts are used in computational procedures to estimate the parameters of a model being used to allocate limited resources or to describe random processes such as those mentioned above. Time series models assume that observations vary according to some probability distribution about an underlying function of time. In this paper we discuss about types of errors in the model forecasting.

REFERENCES

[1] Andrews, D.F (1976), "Significance Tests Based of Residuals", Biometrica, 58, 139-148.
[2] Anscombe, F.J. (May 1960), "Rejection of Outliers," Technometries, 123-46.
[3] Anscombe, F.J., and J.W. Tukey (1963), "The Examination and Analysis of Residuals", Technometrics, 5,141-60.
[4] Belsley, D.A., Kuh, E., and Welsh, R.E. (1972), "Regression Diagnostics", Wiley, New York.
[5] Carter, A.H. (1949), "The Estimation and Comparision of Residual Regressions where there are two or more related sets of observations, Biometrika, 36, 27-47.
[6] Carter, A.H. (1949), "The estimation and comparision of residual regressions where there are two or more related sets of observations, Biometrika, 36, 36-46, 109-11.

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

[7] Chatterjee, S, and Hadi, A.S, (1986), "Influential Observations, High Leverage points, and Outliers in linear Regression", Statistical Science, 179-416.
[8] Cook, R.D. and Chih-Ling Tsai (1985), "Residuals in Nonlinear Regression", Biometrika, 72, 1, 23-29.
[9] Cook, R.D. and Weisberg, S. (1982), "Residuals and Influence in Regression", Chapman and Hall, New York.
[10] Daniel, C. (1978), "Patterns in Residuals in the two-way layout", Technometrics, 20, 385-395.
[11] Dennis E. Jennings (1986), "Outliers and Residual Distributions in Logistic Regression", JASA, 81, 396, 987-990.
[12] Donald, A. Pierce and Doniel W. Schafer (1986), "Residuals in Generalised Linear Models", JASA, 81, 977-986.
[13] Draper, N.R. and Smith, H. (1988), "Applied Regression Analysis", 3rd Edition, Wiley, New York.
[14] Farebrother, R.W. (1976a), "BLUS Residuals, Algorithm", AS104, Applied Statistics 25, 317-319.
[15] Freds, Wood (1973), "The use of individual effects and residuals in fitting equations to data", Technometrics, 15, 677-695.
[16] Freund J.R and F.V.W. Richard and C.W. Clunies Ross (1961), Residual Analysis, JASA, 56, 98-110.
[17] Hoanglin, D.C., and Welson, R. (1978), "The hat matrix in regression and ANOVA", American Statistician, 32, 17-22.
[18] Wood, F.S. (1963), "The use of Individual Effects and residual in fitting equations to data, Biometrika, 55, 1-17.
[19] Wooding, W.M. (1969), "The Computation and use of residuals in the analysis of Experimental Data", Journal of Quality Technology, 1, 175-188.
[20] Zyskind G. (1963), "A note on Residuals Analysis", JASA, 58, 1128-1132.