

INTERNALLY STUDENTIZED RESIDUALS TEST FOR MULTIPLE NON-NESTED REGRESSION MODELS

G. Mohan Naidu¹, P. Balasiddamuni², M. Venkata Ramanaiah², D. Giri³, M. Venkata Subba Reddy² and S. Venkata Rama Raju²

¹*Department of Statistics and Maths, S.V. Agricultural College, Tirupati*

²*Department of Statistics, S.V. University, Tirupati*

³*Degree College, Mahadevpur, Karimnagar District*

**Author for Correspondence*

ABSTRACT

Two models are said to be non-nested models, if one can not be derived as a special case of another. Much attention in classical statistics has been devoted to testing non-nested regression models. Within the classical framework, there are three alternative general approaches to test non-nested models namely, the use of specification error tests; the use of comprehensive model method; and the use of procedures based upon Cox (1961, 1962) and Atkinson (1970) methods. The testing of multiple non-nested models has gained widespread importance in applied statistical research. This paper proposes a simple and convenient test for multiple non-nested regression models by using Internally Studentized residuals.

Keywords: *Non- Nested Model, Studentized Residuals*

INTRODUCTION

The selection of a good model is an art. The basic idea in statistics is how to select a good model for the purpose of the study. Once a model is given, however, there are statistical criteria to judge whether the given model is bad or not. Since, many models can explain the same set of data about equally well, a given set of data can be used to screen out bad models but not to generate good models, whatever statistical techniques are used. The subject of model selection is treated in classical statistics, which deals with the two topics of estimation and testing of hypotheses. The problem of determining an appropriate model based on a subset of the original set of variables contains three basic ingredients namely, i) The computational technique used to provide the information for the analysis; ii) The criterion used to analyze the variables and select a subset, if that is an appropriate; and iii) The estimation of coefficients in the final model.

In model selection criteria, there may be two important problems those arising from nested and non-nested model structures. The nested models arise with, for instance, two models specified in such a way that one model is a special case of the other; the non-nested model arise when neither model follows as a special case of the other.

The model selection criterion is a problem of choice among competing models. The choice of a model follows some preliminary data search. In the context of the linear model, it leads to the specification of explanatory variables that appear to be the most important on prior grounds. Often, some explanatory variables appear in one model and reappear in another model gives rise to the nested models; often again neither model, in the case of two models appears to be a special case of the other model gives rise to the non-nested models.

In the process of choosing models, statisticians have developed a variety of diagnostic tests. These tests have been classified into two categories:

- (i) Tests of Nested Regression models, and
- (ii) Tests of Non-nested Regression models

If a model–I can be derived as a special case of another model–II then model–I is said to be nested model within model–II. Two models are said to be non-nested models, if one can not be derived as a special case of another.

Research Article

Much attention in classical statistics has been devoted to testing non-nested regression models. Within the classical framework, there are three alternative general approaches to test non-nested models:

- i) The use of Misspecification tests or Specification error tests
- ii) The use of comprehensive model method in which the non-nested models are embedded in a general specification.
- iii) The use of procedures based upon the Cox (1961, 1962) and Atkinson (1970) methods.

Mallows (1973) has derived a Conditional Mean Squared Error Prediction (C_p) criterion for choosing between the two nested regression models. The problem of selecting non-nested regression models was first posed by Hotelling (1940) and later generalized by Chow (1957, 1980). According to Harvey (1990) there are two approaches to test a non-nested hypothesis namely (i) Discrimination Approach and (ii) Discerning Approach. Under discrimination approach, given two or more competing models, one chooses a model based on some criteria of goodness of fit and under discerning approach, in investigating one model, one may take into account information provided by other models.

The problem of testing hypothesis regarding non-nested regression models is still an important area of research in statistics. Some of the methods for selecting non-nested regression models were suggested by Cox (1961, 1962) and Atkinson (1970), Quandt (1974), Vuong (1989), Pesaran and Deaton (1978), Davidson and MacKinnon (1981) Fisher and McAleer (1981), Deaton (1982), Berger and Pericchi (1996), Hansen and Yu (2001) and others.

In the presented study, an attempt has been made by proposing a selection criterion for multiple non-nested linear statistical models. Besides this criterion, a modified method for non-nested linear statistical model by using J- Regression model has been suggested, using internally studentized residuals.

Nested and non-nested linear regression models

Consider a linear regression model

$$Y_{nx1} = X_{nxk} \beta_{kx1} + \epsilon_{nx1} \tag{2.1}$$

Or $Y = X_1 \beta_1 + X_2 \beta_2 + \epsilon$ (2.2)

Where X_1 is $n \times k_1$; X_2 is $n \times k_2$;

And $E[\epsilon] = 0$; $E[\epsilon \epsilon'] = \sigma^2 I_n$

To select model (2.1) or the linear model including X_1 alone as the design matrix, one may test the null hypothesis $H_0: \beta_2 = 0$ by using F-test for testing general linear hypothesis. The choice between model (2.1) and a linear regression model including X_1 alone is a case of ‘Nested linear models’. Two linear models are said to be nested models if one model assumes that the parametric vector lies in a subspace of the parametric space assumed under the other model.

Two linear models are said to be Non-nested linear models if they may consist of model (2.1) and

$$Y_{nx1} = Z_{n \times p} \gamma_{px1} + \mu_{nx1} \tag{2.3}$$

Where Z includes a different set of explanatory variables from those included in the (nxk) matrix X in model (2.1). These may be some variables common to both Z and X , but neither hypothesis results from restricting the values of the parameter vector permitted by the other hypothesis.

Testing non-nested linear statistical models

Consider two non-nested linear statistical models under two hypotheses as

$H_1: Y = X_1 \beta_1 + \epsilon_1$ (3.1)

$H_2: Y = X_2 \beta_2 + \epsilon_2$ (3.2)

These two regression models are non-nested because the regressors under one model are not a subset of the other model even though X_1 and X_2 may have some common variables. In order to test H_1 against H_2 , Cox (1961) modified the likelihood ratio test to allow for the non-nested case. The idea behind Cox approach is to consider to what extent model (3.1) under H_1 , is capable of predicting the performance of model (3.2) under H_2 .

Alternatively one may nest the two models as

$H_3: Y = X_1 \beta_1 + X_2^* \beta_2 + \epsilon_3$ (3..3)

Research Article

Where X_2^* excludes from X_2 the common variables with X_1 .

A test for H_1 is simply the F-test for $H_0 : \beta_2^* = 0$. This tests H_1 versus H_3 which is a hybrid of H_1 and H_2 and not H_1 versus H_2 .

Davidson and Mac Kinnon (1981) have proposed to test $H_0 : \alpha = 0$ in the linear combination of H_1 and H_2 :

$$Y = (1-\alpha) X_1 \beta_1 + \alpha X_2 \beta_2 + \epsilon \quad (3.4)$$

Where α is an unknown scalar. Since α is not identified, one may replace β_2 by $\hat{\beta}_2 = \left(\frac{X_2^l X_2}{n} \right)^{-1}$

$\left(\frac{X_2^l Y}{n} \right)$ which can be obtained from regressing Y on X_2 under H_2 .

- i.e., (i) Regress Y on X_2 get $\hat{Y}_2 = X_2 \hat{\beta}_2$
 (ii) Regress Y on X_1 and \hat{Y}_2 and test that the regression coefficient of \hat{Y}_2 is zero.

This test is known as J-test and this test statistic follows $N(0, 1)$ under H_1 .

Fisher and McAleer (1981) have suggested a modification of the J-test which is known as the JA test.
 Under

$$H_1; \text{Plim } \hat{\beta}_2 = \text{Plim} \left(\frac{X_2^l X_2}{n} \right)^{-1} \text{Plim} \left(\frac{X_2^l X_1}{n} \right) \beta_1 + 0.$$

They modified by replacing $\hat{\beta}_2$ as $\tilde{\beta}_2 = \left(X_2^l X_2 \right)^{-1} \left(X_2^l X_1 \right) \hat{\beta}_1$

Where $\hat{\beta}_1 = \left(X_1^l X_1 \right)^{-1} X_1^l Y$.

The JA test procedure is as follows:

- (i) Regress Y on X_1 get $\tilde{Y}_1 = X_1 \hat{\beta}_1$
 (ii) Regress \hat{Y}_1 on X_2 get $\tilde{Y}_2 = X_2 \left(X_2^l X_2 \right)^{-1} X_2^l \hat{Y}_1$
 (iii) Regress Y on X_1 and \tilde{Y}_2 and test the regression coefficient of \tilde{Y}_2 is zero. This is the simple t-test on the regression coefficient of \tilde{Y}_2 .

The J and JA tests are asymptotically equivalent. One can test again by reversing the role of the hypotheses to know the asymmetry of H_1 and H_2 .

Under non-nested hypotheses testing, one may find the following four situations:

		$\alpha = 1$	
		Not Rejected	Rejected
$\alpha = 0$	Not Rejected	Both H_1 and H_2 are not rejected	H_1 rejected H_2 not rejected
	Rejected	H_1 is not rejected H_2 rejected	Both H_1 and H_2 are rejected

Research Article

Both H_1 and H_2 are not rejected \Rightarrow The data are not rich enough to discriminate between the two hypotheses.

Both H_1 and H_2 are rejected \Rightarrow Neither model is useful in explaining the variation is Y .

Either of H_1 and H_2 is rejected \Rightarrow The non-rejected hypothesis may still be brought down by another challenger hypothesis

Here J and JA tests are one degree of freedom tests, whereas the artificially nested F -test is not.

The JA test has relatively low power than J test, when K_1 , the number of parameters in H_1 is larger than K_2 , the number of parameters in H_2 . Thus, one should use the JA test when K_1 is about the same size as K_2 , i.e., the same number of non-overlapping variables. If both H_1 and H_2 are rejected these J and JA tests are inferior to the standard diagnostic tests.

Selection criterion for multiple non-nested linear regression models

The testing of non-nested or separate regression models has gained widespread importance in statistics. The problem of joint tests for non-nested regression models is that of $(m+1)$ competing non-nested linear regression models. Without loss of generality, H_0 may be designated as the null model and H_1, H_2, \dots, H_m may be designated as the alternative models. These non-nested linear regression models may be specified as

$$H_0: Y = X\beta + \epsilon_0, \quad \epsilon_0 \sim (0, \sigma_0^2 I_n) \quad (4.1)$$

$$H_i: Y = Z_i \gamma_i + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma_i^2 I_n) \quad (4.2)$$

$$i = 1, 2, \dots, m$$

Where

Y is $(n \times 1)$ vector of observations on the dependent variable ;

X and Z_i are $(n \times K_0)$ and $(n \times K_i)$ matrices of non-stochastic regressors;

β and γ_i are $(K_0 \times 1)$ and $(K_i \times 1)$ vectors of unknown parameters respectively.

ϵ_0 and ϵ_i are $(n \times 1)$ vectors contain normally, independently and identically distributed random disturbances; and index i represents the i^{th} alternative regression model. A test of H_0 against a single alternative which is represented by H_1 is a paired test for non-nested regression models.

A test of H_0 against the multiple alternatives H_1, H_2, \dots, H_m is a joint test for non-nested regression models.

In testing non-nested hypotheses, it is to assume that the matrices

$$\left(\frac{X^T X}{n} \right) \text{ and } \left(\frac{Z_i^T Z_i}{n} \right) \text{ converge to well defined finite positive definite limits; and } \left(\frac{X^T Z_i}{n} \right) \text{ and } \left(\frac{Z_i^T Z_j}{n} \right)_{i \neq j = 1, 2, \dots, m}$$

coverage to non-zero finite limits.

We consider H_0 as the null and it can be tested against any one of the alternatives H_i , or all of them.

In the present study, so far we have proposed paired tests for non-nested linear regression models. These tests may be adapted for testing H_0 joining against H_1, H_2, \dots, H_m .

We frame a comprehensive model as

$$Y = X\beta + \tilde{D}_X \theta + \epsilon \quad (4.3)$$

Where \tilde{D}_X is the $(n \times m)$ matrix with i^{th} column is Internally studentized residual vector $\tilde{\epsilon}_{X_i}$, which is defined as $(n \times 1)$ vector of n internally studentized residuals as

Research Article

$$\tilde{e}_{Xij} = \frac{e_{Xij}^{**}}{\hat{\sigma}_{\xi_i} \sqrt{(1-h_{ij}(X))}} \quad i = 1, 2, \dots, m.$$

$$j = 1, 2, \dots, n.$$

Where

$$\hat{\sigma}_{\xi_i}^2 = \frac{\sum_j e_{Xij}^{**2}}{n - k_i}$$

Here $H_i(X) = ((h_{ijk}))$ is Hat matrix corresponds to the regressions (4.3).

e_{Xi}^{**} is the OLS residual vector obtained from the regression

$$X \hat{\beta} = Z_i \Gamma_i + \xi_i, \quad \xi_i \sim N(0, \sigma_{\xi_i}^2 I_n) \quad (4.4)$$

i.e., $X \hat{\beta}$ is regressed on Z_i

We may write e_{Xi}^{**} as

$$e_{Xi}^{**} = [I - Z_i (Z_i^T Z_i)^{-1} Z_i^T] X \hat{\beta}, \quad i = 1, 2, \dots, m.$$

The validity of model under H_0 can be tested by testing the significance of θ from zero; the proposed test statistic is given by

$$J^{**} = \frac{Y^T \tilde{D}_X [\tilde{D}_X^T \tilde{D}_X]^{-1} \tilde{D}_X^T Y}{\hat{\sigma}_\epsilon^2} \quad (4.5)$$

Where $\hat{\sigma}_\epsilon^2$ is the estimator of error variance to be obtained from comprehensive model (4.3).

Under H_0 , the test statistic J^{**} follows χ^2 distribution with m degrees of freedom.

Remarks : 1 Consider the comprehensive model as

$$Y = X \beta + \tilde{D}_Z \delta + u, \quad u \sim N(0, \sigma_u^2 I_n) \quad (4.6)$$

Where \tilde{D}_Z is the $(n \times m)$ matrix with i^{th} column is Internally studentized residual vector \tilde{e}_{Z_i} . Which is defined as $(n \times 1)$ vector if n internally studentized residuals as

$$\tilde{e}_{Zij} = \frac{e_{Zij}^{**}}{\hat{\sigma}_{\omega_i} \sqrt{(1-h_{ij}(Z))}} \quad i = 1, 2, \dots, m.$$

$$j = 1, 2, \dots, n.$$

Where

$$\hat{\sigma}_{\omega_i}^2 = \frac{\sum_j e_{Zij}^{**2}}{n - k_0}$$

Here $H_i(Z) = ((h_{ijk}))$ is Hat matrix corresponds to the regressions (4.6).

e_{Zi}^{**} is the OLS residual vector obtained from the regression

$$Z_i \hat{\gamma}_i = X \eta + W, \quad W \sim N(0, \sigma_w^2 I_n) \quad (4.7)$$

i.e., $Z_i \hat{\gamma}_i$ is regressed on X .

Research Article

We write $e_{Z_i}^{**}$ as

$$e_{Z_i}^{**} = [I - (X'X^{-1})X']Z_i \hat{\gamma}_i, \quad i = 1, 2, \dots, m. \quad (4.8)$$

The validity of model under H_0 can be tested by testing the significance of δ from Zero. The proposed statistic is given by

$$J^{***} = \frac{Y' \tilde{D}_Z [\tilde{D}_Z' \tilde{D}_Z]^{-1} \tilde{D}_Z' Y}{\hat{\sigma}_u^2} \sim \chi_m^2 \quad (4.9)$$

Where $\hat{\sigma}_u^2$ is estimator of error variance to be obtained from comprehensive model (4.6).

2. Consider the comprehensive model as

$$Y = X\beta + \tilde{D}\alpha + v, \quad v \sim N(0, \sigma_v^2 I_n) \quad (4.10)$$

Where \tilde{D} is the $(n \times m)$ matrix with i^{th} column is the difference between the Internally Studentized residual vectors

say $(\tilde{e}_{X_i} - \tilde{e}_{Z_i}), i = 1, 2, \dots, m.$

The validity of model under H_0 can be tested by testing the significance of α from zero. The proposed statistic is given by

$$\bar{J} = \frac{Y' \tilde{D} [\tilde{D}' \tilde{D}]^{-1} \tilde{D}' Y}{\hat{\sigma}_v^2} \sim \chi_m^2 \quad \dots (4.11)$$

Where $\hat{\sigma}_v^2$ is the estimator of error variance to be obtained from the comprehensive model (4.10).

3. The consistency of the aforementioned joint tests for non-nested linear regression models can be proved by using the method given by Dastoor and Mc Aleer (1987).

Modified criterion for non-nested linear regression models using j-regression model

Consider two non-nested linear regression models,

$$H_1 : Y = X\beta + \epsilon_1, \quad \epsilon_1 \sim N(0, \sigma_1^2 I_n) \quad (5.1)$$

$$H_2 : Y = Z\gamma + \epsilon_2, \quad \epsilon_2 \sim N(0, \sigma_2^2 I_n) \quad (5.2)$$

Where Y is $(n \times 1)$, X is $(n \times K_1)$, Z is $(n \times K_2)$, ϵ_1 and ϵ_2 are $(n \times 1)$; β is $(K_1 \times 1)$ and γ is $(K_2 \times 1)$ vectors.

For testing the empirical adequacy of H_1 against H_2 we consider an artificial compound model (Known as J. Regression model) as

$$H_3 : Y = (1 - \alpha)X\beta + \alpha e_Z + u, \quad u \sim N(0, \sigma_u^2 I_n) \quad (5.3)$$

Where α is a nesting parameter;

e_Z is the OLS residual obtained from the regression

$$Y = Z\gamma + \epsilon_2.$$

$$\begin{aligned} \text{We have} \quad e_Z &= Y - \hat{Y}_Z = Y - Z(Z'Z)^{-1}Z'Y \\ &= [I - Z(Z'Z)^{-1}Z']Y \\ e_Z &= P_Z Y \end{aligned}$$

Research Article

$$\therefore H_3 : Y = (1-\alpha) X\beta + \alpha P_Z Y + u$$

The OLS estimator of α is given by

$$\hat{\alpha} = [Y^{\perp} P_Z P_X P_Z Y]^{\perp} [Y^{\perp} P_Z P_X Y] \tag{5.4}$$

Where

$$P_X = [I - X (X^{\perp} X)^{\perp} X^{\perp}]$$

$\hat{\alpha}$ is a ratio of quadratic forms in the n-dimensional normal vector Y.

Michelis and Stengos (1993) have obtained an expression for the exact mean of the estimator of the nesting parameter $\hat{\alpha}$ under H_1 as

$$E(\hat{\alpha}) = \int_0^{\infty} |D_{0t}|^{-1/2} [\text{tr}(N D_{0t}^{\perp}) + H^{\perp} D_{2t} H] \exp\left\{-\frac{1}{2} H^{\perp} D_{1t} H\right\} dt$$

(5.5)

Where

$$H = X \beta$$

$$D_{0t} = [I + 2t \sigma_1^2 D]$$

$$D_{1t} = 2t D D_{0t}^{\perp}$$

$$D_{2t} = D_{0t}^{\perp} N D_{0t}^{\perp}$$

Here,

$$N = P_Z P_X$$

$$D = P_Z P_X P_Z P_X$$

It should be noted that under H_1 , $\alpha = 0$ but $E(\hat{\alpha}) \neq 0$. Following the results of Ullah (1990), one can derive the higher order moments of $\hat{\alpha}$ and also its exact distribution.

The expression $E(\hat{\alpha})$ suggests that under the null hypothesis H_1 , the exact mean of $\hat{\alpha}$ depends on all the parameters and regressors of the two linear regression models. Asymptotically, $\hat{\alpha}$ converges to its true value zero, but in finite samples it has a nonzero mean.

Under the null hypothesis that $\alpha = 0$, the simple t-statistic for the significance of $\hat{\alpha}$, has zero mean asymptotically and a standard normal distribution.

Under the null hypothesis that H_1 is correct, $\hat{\alpha}$ converges asymptotically to its true value of zero, since in large sample the correlation between the test regressor $P_Z Y$ and the errors of the compound model u , vanishes.

In finite samples, this correlation leads to a non zero mean of $\hat{\alpha}$ and hence non zero mean of the corresponding t-statistic for $\alpha=0$.

Conclusions

In the present research study, an attempt has been made by proposing some criteria for selection of non-nested linear statistical models by using internally studentized residuals. Unlike most previous work on model selection based on usual R^2 , \bar{R}^2 and C_p - criteria, here, we adopted the classical hypothesis testing framework and proposed some procedures for choosing between non-nested linear statistical models.

Recently, the testing of non-nested linear regression models has gained widespread importance in applied statistics. The various procedures for testing pairs of non-nested linear regression models (Paired tests) as well as testing a null model jointly against multiple non-nested alternatives (Joint tests) have been proposed under the present research work.

The criteria proposed under the present study allow one to test the truth of a possibly linear regression model, when there exists a non-nested alternative linear model. The latter need not be true, and need not even be a hypothesis, which the statistician would maintain.

The proposed techniques have been developed by using OLS and studentized residuals. This kind of research can be further extended by considering various problems such as multicollinearity, heteroscedasticity and autocorrelation along with the proposed techniques for model selection.

Research Article

Under the proposed research work, we consider only univariate linear regression models. The present research study can be generalized for multivariate linear regression models or sets of linear regression models.

REFERENCES

- Atkinson AC (1970).** A Method for Discriminating Between Models. *Journal of the Royal Statistical Society, Series B* **32** 323-345.
- Berger J and Pericchi L (1996).** The Intrinsic Bayes Factor for Model Selection and Prediction. *Journal of the American Statistical Association* **91**(433) 109-122.
- Chow GC (1957).** Demand for Automobiles in the United States: A Study in Consumer Durables, North-Holland, Amsterdam.
- Chow GC (1980).** The Selection of Variates for Use in Prediction: A Generalization of Hotellings Solution. In: *Quantitative Econometrics and Development* edited by LR Klien, M Nerlove and SC Tsiang (Academic, New York) 105-114.
- Cook RD and Weisberg S (1982).** *Residuals and Influence in Regression* (Chapman and Hall, New York).
- Cox DR (1961).** Tests of Separate Families of Hypotheses. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability 1* (University of California Press, Berkeley).
- Cox DR (1962).** Further Results on Tests of Separate Families of Hypothesis. *Journal of the Royal Statistical Society, Series B* **24** 406-424.
- Dastoor NK and Mc Aleer M (1987).** On the Consistency of Joint and Paired Tests for Non-Nested Regression Models. *Journal of Quantitative Economics* **3** 65-84.
- Davidson R and MacKinnon JG (1981).** Several Tests for Model Specification in the Presence of Alternative Hypotheses. *Econometrica* **49** 781-793.
- Deaton AS (1982).** Model Selection Procedures, or Does the Consumption Function Exist? Chapter 5 In: *Evaluating the Reliability of Macro-Economic Models* edited by GC Chow and P Corsi (Wiley, London).
- Fisher GR and McAleer M (1981).** Alternative Procedures and Associated Tests of Significance for Non-nested Hypotheses. *Journal of Econometrics* **16** 103-119.
- Giri D (2006).** Some New Selection Techniques for Linear Statistical Models. Unpublished Ph.D., Thesis, S.V. University, Tirupati, Andhra Pradesh State, India
- Godfrey LG (1983).** Testing Non-Nested Models After Estimation by Instrumental Variables or Least Squares. *Econometrica* **51** 355-365.
- Godfrey LG (1983).** Tests of Non-Nested Regression Models. *Journal of Econometrics* **21** 133-154.
- Hansen MH and Yu B (2001).** Model Selection and the Principle of Minimum Description Length. *Journal of the American Statistical Association* **96** 746-774.
- Harvey A (1990).** The Econometric Analysis of Time Series. *The MIT Press*, 2d edition (Cambridge Mass) Chapter 5.
- Hotelling H (1940).** The Selection of Variates for Use in Prediction with Some Comments on the General Problem of Nuisance Parameters. *The Annals of Mathematical Statistics* **11** 271-283.
- Mallows CL (1973).** Some Comments on Cp. *Technometrics* **15** 661-676.
- Pesaran MH and Deaton AS (1978).** Testing Nonnested Nonlinear Regression Model. *Econometrica* **46** 677-694.
- Quandt RE (1974).** A Comparison of Methods for Testing Non-nested Hypotheses. *Review of Economics and Statistics* **56** 92-99.
- Ullah A (1990).** Finite Sample Econometrics: A Unified Approach. In: *Contributions to Econometric Theory and Applications* edited by Carter et al., (Springer Vertag).
- Vuong QH (1989).** Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica* **57**(2) 307-333.