TESTING ON THE RECURRENCE OF COEFFICIENTS IN THE LINEAR REGRESSIONAL MODEL

by
Nicoleta Breaz and Daniel Breaz

Abstract. The statistical tests on the coefficients of the linear regressional model are well known. In this paper, we make three new tests, by means of which one can verify if the coefficients are terms of a sequence given by a recurrencial relation namely geometrical and aritmetical progression and a general linear reccurence of specified order.

Introduction.

Let the multiple linear regressional model

$$Y = \sum_{k=1}^{p} \alpha_k X_k + \varepsilon \qquad (1)$$

and a sample data

$y^T = (y_1, y_2, ..., y_n) \in \mathbb{R}^n$,

$$x = (x_1, ..., x_p) = \begin{pmatrix} x_{11} & x_{12} & ... & x_{1p} \\ x_{21} & x_{22} & ... & x_{2p} \\ ... & ... & ... & ... \\ x_{n1} & x_{n2} & ... & x_{np} \end{pmatrix} \in M_{n,p}(\mathbb{R}), n >> p$$

Denoting $\alpha^T = (\alpha_1, \alpha_2, ..., \alpha_p) \in \mathbb{R}^p$, $\varepsilon^T = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n) \in \mathbb{R}^n$, from (1) we obtain the matriceal form

$$y = x\alpha + \varepsilon \qquad (2)$$

The principle of least squares leads to the fitting model

$$y = xa + e$$

with $a^T = (a_1, a_2, ..., a_p) \in \mathbb{R}^p$, $e^T = (e_1, e_2, ..., e_n) \in \mathbb{R}^n$ and

$$\sum_{k=1}^{n} e_k^2 = \min$$

If $rang(x) = p$ then the fitted coefficients are

$$a = (x^T x)^{-1} x^T y$$

We call the clasical regressional linear model the model (2) that satisfies the following conditions.

i) $E(\varepsilon) = \theta, \theta^T = (0,0,...,0) \in \mathbb{R}^n$

ii) $Var(\varepsilon) = \sigma^2 I$,

iii) $\varepsilon - N$ (3)
which can be written as

$$y - N(x\alpha, \sigma^2 I)$$

The following statistical tests are well known.

A. F-test on equality between the coefficients

It is testing the null hypothesis

$$H_0 : \alpha_1 = ... = \alpha_q$$

versus the alternative H_1 :"H_0 – false". For a fixed leavel φ, we find the value

$$f = f_{q-1;n-p;1-\varphi}$$

of a Fisher-Snedecor statistics with $q-1$ and $n-p$ degree of freedom, such that

$$P(F < f | H_0) = 1 - \varphi$$

The null hypothesis is rejected if $f_c > f$ where

$$f_c = \frac{n-p}{q-1} \frac{S_0^2 - S_1^2}{S_1^2}$$

Here and elsewhere in this paper,we use the denotations
S_0^2 - the sum of squared residuals in the reduced model, obtained under H_0
S_1^2 -the sum of squared residuals in the full model, obtained under H_1

B. F-test on signifiance of the coefficients

It is testing the null hypothesis

$$H_0 : \alpha_1 = ... = \alpha_q = 0$$

versus the alternative H_1 :"there exists i_0 from $\{1,2,...,q\}$ such that $\alpha_{i_0} \neq 0$".
For a fixed level φ the test rejects the null hypothesis if $f_c > f$ where

$$f_c = \frac{S_0^2 - S_1^2}{q} \Big/ \frac{S_1^2}{n-p}$$

and $f = f_{q;n-p;1-\varphi}$ is deduced from

$$P(F < f | H_0) = 1 - \varphi$$

Remark. The statistical tests presented above are based on a lemma which provides a Fisher-Snedecor (F) statistics. Using this lemma we make new tests that verifies some recurrential connection between the coefficients.

Lemma. Let be the random vector $\varepsilon^T = (\varepsilon_1,...,\varepsilon_n)$ as

$$\varepsilon - N(\theta;\sigma^2 I)$$

and the matrix $x \in M_{np}(\mathbb{R})$, $A \in M_{pq}(\mathbb{R})$,

$$x_0 = xA \in M_{nq}(\mathbb{R})$$

If

$$Q = I - x(x^T x)^{-1} x^T \text{ and } Q_0 = I - x_0 (x_0{}^T x_0)^{-1} x_0{}^T$$

then the statistics

$$F = \frac{\varepsilon^T Q_0 \varepsilon - \varepsilon^T Q \varepsilon}{rank(Q_0) - rank(Q)} \Big/ \frac{\varepsilon^T Q \varepsilon}{rank(Q)}$$

has a Fisher-Snedecor distribution with $rank(Q_0) - rank(Q)$ and $rank(Q)$ degrees of freedom .

MAIN RESULTS

The following results are obtained on the classical regressional model defined by (3).

The aritmetical progression test

We consider the null hypothesis

$$H_0 : \alpha_2 = \alpha_1 + r, \alpha_3 = \alpha_1 + 2r, ..., \alpha_p = \alpha_1 + (p-1)r, \ \ p > 2$$

versus the alternative H_1 : ” H_0 false “

Denoting

$$x_{H_0}^T = (x_1 + x_2 + ... + x_p, x_2 + 2x_3 + ... + (p-1)x_p)$$

$$\alpha_{H_0}^T = (\alpha_1, r) = (\alpha_1, \alpha_2 - \alpha_1)$$

we have the full model

$$y = x\alpha + \varepsilon \ \text{ (under } H_1\text{)}$$

and a reduced one

$$y = x_{H_0}\alpha_{H_0} + \varepsilon \ \text{ (under } H_0\text{)}$$

We note that for $r = 0$ is obtained the classical test A.

It can be proved the following proposition:

1.Proposition

We have

$$x_{H_0} = xA,$$

where $A \in M_{p,2}$

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ \dots & \dots \\ 1 & p-1 \end{pmatrix}$$

2.Proposition

Let be the matrix

$$Q = I - x\left(x^T x\right)^{-1} x^T \text{ and } Q_0 = I - x_{H_0}\left(x_{H_0}^T x_{H_0}\right)^{-1} x_{H_0}^T$$

and the random variable

$$S_0^2 = \varepsilon^T Q_0 \varepsilon = \|e_0\|^2 \text{ and } S_1^2 = \varepsilon^T Q \varepsilon = \|e\|^2$$

The random variable given by

$$F = \frac{S_0^2 - S_1^2}{p-2} \Big/ \frac{S_1^2}{n-p} - F(p-2, n-p),$$

has a Fisher-Snedecor distribution with $p-2$ and $n-p$ degrees of freedom.

Proof.

If in lemma we use $q = 2$ and matrix A from proposition 1., we obtain

$$F = \frac{\varepsilon^T Q_0 \varepsilon - \varepsilon^T Q \varepsilon}{rank(Q_0) - rank(Q)} \Big/ \frac{\varepsilon^T Q \varepsilon}{rank(Q)} \quad - F\left(rank(Q_0) - rank(Q), rank(Q)\right),$$

Moreover

$$rank(Q_0) = n - 2$$

$$rank(Q_0) - rank(Q) = n - 2 - n + p = p - 2$$

so the result holds.

Testing

- $\varphi \to f \quad f = f_{p-2;n-p;1-\varphi} : P\left(F < f \middle| H_0\right) = 1 - \varphi$
- $f_c = \frac{S_0^2 - S_1^2}{p-2} \Big/ \frac{S_1^2}{n-p}$
- if $f_c > f$ then the H_0 is rejected.

The geometrical progression test

We consider the null hypothesis

$$H_0 : \alpha_2 = \alpha_1 r,\ \alpha_3 = \alpha_1 r^2, ..., \alpha_p = \alpha_1 r^{p-1},\ p > 1$$

versus the alternative H_1 : ” H_0 false “

Denoting

$$x_{H_0}^T = \left(x_1 + r x_2 + ... + r^{p-1} x_p\right),\quad r = \frac{\alpha_3}{\alpha_2}$$

$$\alpha_{H_0}^T = \alpha_1 .$$

we have the full model

$$y = x\alpha + \varepsilon \text{ (under } H_1\text{)}$$

and a reduced one

$$y = x_{H_0}\alpha_{H_0} + \varepsilon \text{ (under } H_0\text{)}$$

We note that for $r = 0$ is obtained the classical test B and for $r = 1$ is obtained the clasical test A.

It can be proved the following proposition:

3.Proposition

We have

$$x_{H_0} = xA,$$

where $A \in M_{p,1}$

$$A = \begin{pmatrix} 1 \\ r \\ r^2 \\ ... \\ r^{p-1} \end{pmatrix}$$

4.Proposition

With simillar notations we have

$$F = \frac{S_0^2 - S_1^2}{p-1} \Bigg/ \frac{S_1^2}{n-p} - F\left(p-1, n-p\right),$$

Proof.

In lemma we use $q = 1$ and matrix A from proposition 3.

Moreover

$$rang\left(Q_0\right) = n - 1$$

$$rang(Q_0) - rang(Q) = n - 1 - n + p = p - 1$$

so the result holds.

Testing

- $\varphi \to f \quad f = f_{p-1;n-p;1-\varphi} : P(F < f | H_0) = 1 - \varphi$
- $f_c = \dfrac{S_0^2 - S_1^2}{p-1} \Big/ \dfrac{S_1^2}{n-p}$
- if $f_c > f$ then the H_0 is rejected.

Testing on the linear recurrence of $p-2$ order, $p > 2$

We consider the null hypothesis

H_0 :"$\alpha_1, ..., \alpha_p$ are terms of a sequence given by a linear recurence of p - 2 order

namely $x_{n+p-2} = k_1 x_n + k_2 x_{n+1} + ... + k_{p-2} x_{n+p-3} + k_{p-1}$"

versus the alternative H_1 : " H_0 false "

The hypothesis H_0 can be written as

$$\begin{cases} \alpha_{p-1} = \alpha_1 k_1 + ... + \alpha_{p-2} k_{p-2} + k_{p-1} \\ \alpha_p = (\alpha_2 k_1 + ... + \alpha_{p-2} k_{p-3}) + k_{p-2}(\alpha_1 k_1 + ... + \alpha_{p-2} k_{p-2} + k_{p-1}) + k_{p-1} \end{cases}$$

Denoting

$$x_{H_0}^T = (x_1 + k_1 x_{p-1} + k_{p-2} k_1 x_p,\ x_2 + k_2 x_{p-1} + k_1 x_p + k_{p-2} k_2 x_p, ...,$$
$$x_{p-2} + k_{p-2} x_{p-1} + k_{p-3} x_p + k_{p-2} k_{p-2} x_p,\ x_{p-1} + x_p + k_{p-2} x_p)$$

$$\alpha_{H_0}^T = (\alpha_1, \alpha_2, ..., \alpha_{p-2}, k_{p-1}).$$

we have the full model

$$y = x\alpha + \varepsilon \text{ (under } H_1)$$

and a reduced one

$$y = x_{H_0} \alpha_{H_0} + \varepsilon \text{ (under } H_0)$$

It can be proved the following proposition:

5.Proposition

We have

$$x_{H_0} = xA,$$

where $A \in M_{p,p-1}$

$$A = \begin{pmatrix} 1 & 0 & ... & 0 & 0 \\ 0 & 1 & ... & 0 & 0 \\ ... & ... & ... & ... & ... \\ 0 & 0 & ... & 1 & 0 \\ k_1 & k_2 & ... & k_{p-2} & 1 \\ k_1 k_{p-2} & k_1 + k_2 k_{p-2} & ... & k_{p-3} + k_{p-2}^2 & 1 + k_{p-2} \end{pmatrix}$$

6.Proposition
With simillar notations we have

$$F = \frac{S_0^2 - S_1^2}{1} \Big/ \frac{S_1^2}{n-p} - F(1, n-p),$$

Proof.
In lemma we use $q = p - 1$ and matrix A from proposition 5.
Moreover

$$rang(Q_0) = n - p + 1$$

$$rang(Q_0) - rang(Q) = n - p + 1 - n + p = 1$$

so the result holds.

Testing

- $\varphi \to f \;\; f = f_{1;n-p;1-\varphi} : P(F < f | H_0) = 1 - \varphi$
- $f_c = \frac{S_0^2 - S_1^2}{1} \Big/ \frac{S_1^2}{n-p}$
- if $f_c > f$ then the H_0 is rejected.

7.Remark
The last test can be generalised for a linear recurrence of l order, $1 \le l \le p - 2$.

REFERENCES

1.Stapleton J.H.-Linear Statistical Models, A Willey-Interscience Publication, Series in Probability and Statistics, New York,1995.
2.Saporta G.-Probabilités, analyse des données et statistique, Editions Technip, Paris, 1990

Authors:

Nicoleta Breaz- „1 Decembrie 1918” University of Alba Iulia, str. N. Iorga, No. 13, Departament of Mathematics and Computer Science, email: nbreaz@lmm.uab.ro

Daniel Breaz- „1 Decembrie 1918” University of Alba Iulia, str. N. Iorga, No. 13, Departament of Mathematics and Computer Science, email: dbreaz@lmm.uab.ro