AN EXTENSION OF THE METHODOLOGY OF USING THE STUDENT TEST IN CASE OF A LINEAR REGRESSION WITH THREE EXPLANATORY VARIABLES

Author:*
Florin-Marius PAVELESCU

A*bstract. Having in view some algebraical properties of the Student Test statistics, in this paper a new indicator is proposed in order to evaluate the quality of estimation in case of a linear regression with three explanatory variables, when OLS method is used. The respective indicator, named Synthesis of Transformed Form of Student Test statistics (STFST) is defined by taking into consideration the arithmetical mean and the coefficient of variation of Transformed Form of Student Test Statistics. Also, an explanation regarding the deviation of the arithmetical mean of Transformed Form of Student Test Statistics from Fisher Test statistics is given. In this context, a factorial analysis model is proposed in order to reveal the modeling factors contribution in obtaining the Student Test statistics. At the end of the paper a numerical example is presented in order to show the practical use of the proposed analysis methodology.*

Keywords: *Transformed Form of Student Test statistics, Fisher Test statistics, informational energy, factorial analysis*

JEL Classification: *C13, C20, C51, C52*

It is well known that very often in case of multiple linear regressions, when OLS estimation method is used the Student Test statistics are not correlated with the values of Pearson coefficients of correlation between the dependent variable and the analyzed explanatory variable as it happens when we deal with simple linear

* *Florin-Marius PAVELESCU Ph. D., Institute of National Economy, Bucharest, e-mail pavelescu.florin@yahoo.com*

regressions. Also, it is possible that Student Test statistics related to the considered explanatory variables are very dispersed. Therefore, the econometric modelers face a difficult situation, i.e. to validate as significant only a part of considered explanatory variables.

The apparently strange Student Test statistics are a consequence of (multi)collinearity[1], which is manifest in any method of estimation. An important advantage of OLS method is that the impact of (multi)collinearity on the estimated parameters and Student Test statistics can be quantified, if we consider some algebraical properties related to classical statistical tests.

1. Review of algebraical properties of some classic statistical tests statistics

If we consider the demonstrations presented in F.M. Pavelescu (2013) it is possible to express the statistics of the following statistical tests applied to a multiple linear regression with m observations and 3 explanatory variables, such as:

N.B. y = dependent variable, x_k= explanatory variable (k=1…3)

$$F_{m,3} = (\frac{m-4}{3}) \cdot (\frac{R^2_{3y}}{1-R^2_{3y}}) \, , \tag{1}$$

$$t_{bnk} = \left[m-4\right]^{1/2} \cdot \frac{R(x_k\,;y)}{(1-R_{3y})^{1/2}} \cdot (\frac{1}{VIF_{xk}})^{(1/2)} \cdot T_{3k} \, , \tag{2}$$

$$R^2_{3y} = \sum_{k=1}^{3} R^2(x_k\,;y) * T_{3k} \tag{3}$$

$$VIF_{xk} = \frac{1}{1-R^2_{2xk}} \tag{4}$$

$$T_{3k} = \frac{1-p_{jkwmean} \cdot R^2_{2xk}}{1-R^2_{2xk}} \, , \tag{5}$$

where:

$F_{m,3}$= Fisher Test statistics in case of m observations and 3 explanatory variables.

---

[1] *In this paper, the (multi) collinearity is defined as in Glauber, Farrar (1967), i.e. as “departure from orthogonality of the explanatory variables”.*

R^2_{3y}= coefficient of determination for the linear regression with 3 explanatory variables and dependent variable y.

$R(x_k; y)$ = Pearson coefficient of correlation between the explanatory variable x_k and the dependent variable y.

t_{b3k}= standard Student Test statistics in case of a linear regression with m observatons and 3 explanatory variables.

VIF_{2xk} = variance of Inflation Factor related to explanatory variable x_k.

T_{3k}= coefficient of alignment to collinearity hazard related to explanatory variable x_k[1].

R^2_{2xk}= coefficient of determination for the linear regression $x_k = a_2 + \sum_{j=1}^{3} x_{j\, j \neq k}$

$p_{jkwmean}$= weighted arithmetical mean of the ratios p_{jk}

$$p_{jk} = \frac{r_{jk}}{R(x_j; x_k)} \tag{6}$$

$$r_{jk} = \frac{R(x_j; y)}{R(x_k; y)} \tag{7}$$

where:

$R(x_j; y)$ = Pearson coefficient of correlation between explanatory variables x_j and dependent variable.

$R(x_k: y)$ = Pearson coefficient of correlation between explanatory variables x_k and dependent variable.

---

[1] *Coefficients of alignment to collinearity hazard act also as a modeling factor for the estimated parameters of a multiple linear regression. It can be demonstrated that in case of a simple linear regression,* $y = a_1 + b_{1k} * x_k$ *, the estimated value for parameter* b_{1k} *is:*

$b_{1k} = \frac{\operatorname{cov}(y; x_k)}{\operatorname{var}(x_k)}$ *, where:*

cov (y;x_k*) = covariance between resultative variable and the explanatory variable* x_k
var (x_k*) = variance of explanatory variable* x_k

If we deal with a multiple linear regression, $y = a_n + \sum_{k=1}^{n} b_{nk} * x_k$ *(k = 1…n), respectively, the estimated value of parameter* b_{nk} *is determinated by the formula:* $b_{nk} = b_{1k} * T_{nk}$ *(F. M. Pavelescu (1986)).*

Formula (2) reveals the fact that Student Test statistics are influenced, among other factors, by the coefficient of alignment to collinearity hazard (T_{3k}). If explanatory variables are strictly orthogonal, i.e. all the Pearson coefficients of correlation are equal to zero, all the coefficients T_{3k} are equal to one[1].

If there are some negative coefficients T_{3k}, we may speak about the harmful collinearity occurrence, emphasized by the presence of "wrong sign(s)" in case of one or more estimated parameters, i.e. contrary to the sign(s) of the Pearson coefficients of correlation between the analyzed explanatory variable(s) and dependent variable (C. Conrad, 2006)).

The standard methodology used for the determination of Student Test statistics does not allow us to highlight at a first sight the existence of the harmful collinearity. Consequently, we may use a transformed form of Student Test statistics (TFST) obtained by multiplying the standard Student Test statistics (t_{bnk}) with the scalar $\frac{|R(x_k;y)|}{R(x_k;y)}$. In other words, Transformed Form of Student Test statistics is defined by formula:

$$TFST_{3k} = [m-4]^{1/2} \cdot \frac{|R(x_k;y)|}{(1-R_{3y})^{1/2}} \cdot (\frac{1}{VIF_{xk}})^{(1/2)} \cdot T_{3k} \qquad (8)$$

It is important to note for $TFST_{bnk}$ that all the requirements related to sampling distributions, critical regions and power functions are the same as in case of standard Student Test. It is important to mention that the respective indicator may be used only in the context of multiple regressions.

2. The correlation between Fisher Test and Student Test statistics and the definition of Synthesis of Transformed Form of Student Test statistics (STFST)

It is well known that in case of a simple linear regression, the square of the Student Test computed value is equal with the Fisher Test computed value.

If we deal with linear regression with three explanatory variables and if we consider the formulae (1) and (2), we are able to detect a relationship between Fisher Test and Transformed Form of Student Test (TFST) statistics, such as:

---

[1] *It is to note that also in case of a linear regression with n explanatory variables, which are strictly orthogonal, all the coefficients of alignment to collinearity hazard are equal to 1.*

$$TFST_{3k} = \sqrt{3 \cdot F_{m,3}} \cdot \frac{|R(x_k; y)|}{(\sum_{k=1}^{3} R^2(x_k; y) \cdot T_{3k})} \cdot \sqrt{(1 - R^2_{2xk})} \cdot T_{3k} \quad (9)$$

Formula (9) highlights the fact that Transformed Form of Student Test statistics are influenced both by common factors, i.e. the number of considered explanatory variables and the Fisher Test statistics, and by specific factors such as $VIF_{(n-1)xk}$ and T_{nk}, which mainly reveal the feature of (multi)collinearity.

It is to note that if the explanatory variables are perfectly orthogonal, $TFST_{3k}$ may be expressed as:

$$TFST_{3k} = \sqrt{3 \cdot F_{m,3}} \cdot \frac{|R(x_k; y)|}{\sum_{k=1}^{3} R^2(x_k; y)} \quad (10)$$

If all the explanatory variables are orthogonal, i.e. $R(x_1;x_2) = R(x_1;x_3) = R(x_2;x_3)=0$, and all the three Pearson coefficients of correlation between explanatory variables and dependent variable are equal, i.e. $R(x_1;y) = R(x_1;y) = R(x_2;y)$, we may obtain:

$$TFST_{3k} = (F_{m,3})^{(1/2)} \quad (11)$$

In other words, the relationship described by formula (11) occurs when there is no collinearity between explanatory variables and the informational energy of Pearson coefficients of correlation between the explanatory variables and the dependent variable ($IER(x_k;y)$) is minimum[1].

On this basis, we may conclude that the deviation from 1 of the ratio $\frac{TFSTAM_{m,3}}{(F_{m,3})^{(1/2)}}$ is an outcome of a certain concentration of the absolute values of the Pearson coefficients of correlation $R(x_k; y)$ and of the intensity of collinearity between the explanatory variables.

[1] *In this paper, the informational energy of Pearson coefficients of correlation between the explanatory variables and dependent variable ($IER(x_k;y)$ may be computed by formula:* $IER(x_k; y) = \frac{\sum_{k=1}^{3} R^2(x_k; y)}{(\sum_{k=1}^{3} R(x_k; y))^2}$

The above-mentioned ratio decreases as the informational energy grows and the collinearity is higher and higher. In fact, the usual situation, which occurs when estimation is made, is $\frac{TFSTAM_{m,3}}{(F_{m,3})^{(1/2)}}$ <1.

It is important to note that under the condition of strict orthogonality of explanatory variables and minimum (IER(x_k, y) it is possible to obtain the minimum value of the coefficient of variation of the transformed form of student test statistics ($TFSTCV_{m,3}$). In this case $TFSTCV_{m,3}$=0.

If all the coefficients of alignment to collinearity hazard are positive, $TFSTCV_{m,3}$ takes on values comprised between 0 and $\sqrt{2}$. If there are negative coefficients of alignment to collinearity hazard, $TFSTCV_{m,3}$ takes on higher values, even bigger than $\sqrt{2}$.

The computation formula for $TFSTCV_{m,3}$ is sensibly more complicated in comparison with the case of linear regression with two explanatory values[1]. The difficulty mentioned above is given by the fact that the computation of the indicator involves six different Pearson coefficients of correlation which may be distributed in six situations, because, as we shall see further, the explanatory variables act as descriptors of the dependent variable behaviour and generators of collinearity at the same time.

---

[1] *For the case of a linear regression with two explanatory variables the computation formula for $TFSTCV_{m,2}$ is given in F. M. Pavelescu (2012):*

$$TFSTCV_{m,2} = \frac{(1+R(x_p x_s))\cdot(1-\frac{|R(x_s;y)|}{|R(x_p;y)|})}{(1-R(x_p;x_s))\cdot(1+\frac{|R(x_s;y)|}{|R(x_p;y)|})}$$

The formula mentioned above shows that a high value of the coefficient of variation of the Transformed Form of Student Test statistics is determinated both by an intense correlation between the considered explanatory variables and by the differentiation of the absolute values of the Pearson coefficients of correlation $R(x_p;y)$ and $R(x_s;y)$.

It is important to have in mind that $|R(x_s;y)| < |R(x_p;y)|$. *Therefore, x_p= main explanatory variable and x_s= secondary explanatory variable in the linear regression with two explanatory variables (cf. F. M. Pavelescu (2012).*

3. A model of factorial analysis of the Transform Form of Student Test statistics

The advantage of defining Transformed Form of Student Test statistics consists not only in the possibility of a rapid determination, but also in the possibility to provide a hierarchy of the explanatory variables from two points of view: a) descriptors of the behaviour of the dependent variable; and b) generators of collinearity.

We may establish the ranks of the explanatory variables considering the absolute values of the Pearson coefficient of correlation between the explanatory variables and the dependent variable. Therefore, the primordial (main) explanatory variable is considered to be the explanatory variable which has the highest absolute value of the Pearson coefficient of correlation with the dependent variable. The explanatory variable of rank 3 is that explanatory variable which has the smallest absolute values of the Pearson coefficient of correlation with the dependent variable.

The order of generators of collinearity is given by the values of the Variance Inflation Factor (VIF). In other words, the main generator of collinearity in a linear regression with three explanatory variables in our case is the explanatory equation with the highest VIF, while the generator of collinearity of order 3 is the explanatory variable with the smallest VIF[1].

[1] *In F. M. Pavelescu (2013) it is also demonstrated that in case of a linear regression with three explanatory variables that the order of generators of collinearity is given both by the values of VIF related to each explanatory variables and by the absolute values of the Pearson coefficients between explanatory variables. If we consider this method for the determination of the order of generators of collinearity it is important to make some transformation of the matrix of Pearson coefficients of correlation between the explanatory variables in order to obtain the first two Pearson coefficients of correlation as positive.*
In this context, $R(x_1;x_2)>R(x_1;x_3)>R(x_2;x_3)$ where x_1= main generator of collinearity, x_2= generator of collinearity of order 2, x_3= generator of collinearity of order 3.
If we note $K_1 = \frac{R(x_1;x_3)}{R(x_1,x_2)}$ *and* $K_2 = \frac{R(x_2;x_3)}{R(x_1,x_3)}$ *, we have:*

$$R^2_{2x1} = R^2(x_1;x_2)\cdot(1+\frac{(K_1 - R(x_1;x_2)\cdot K_1\cdot K_2)^2}{1-R^2(x_1;x_2)\cdot K_1^2\cdot K_2^2}),$$

$$R^2_{2x2} = R^2(x_1;x_2)\cdot(1+\frac{(K_1\cdot K_2 - R(x_1;x_2)\cdot K_1\cdot K_2)^2}{1-R^2(x_1;x_2)\cdot K_1^2}),$$

As we have seen before, in case of a multiple linear regression, the Transformed Form of Student Test statistics is, at first sight, determined by the square root of degrees of freedom and the absolute values of the Pearson coefficients of correlation, $R(x_k, y)$ and $R(x_k; x_j)$, respectively.

On the other hand, it is important to consider also the distribution of Pearson coefficients of correlation between the explanatory variables ($R(x_j, x_k)$) in comparison with the hierarchy established by the absolute values of Pearson coefficients of correlation between the explanatory variables and the dependent variable ($R(x_k, y)$)[1].

Consequently, we consider that building and practical use of a factorial analysis model of the Transformed Form of Student Test statistics would act for a better understanding of results obtained during parameters estimation of linear regressions with three explanatory variables.

Therefore, we may consider that Transformed Form of Student Test statistics are modelled by:

A) the premise of the respective indicators;

B) the differentiation of the absolute values of Pearson coefficients of correlation between the explanatory values and the dependent variable ($R(x_k, y)$);

$$R^2_{2x3} = R^2(x_1;x_2) \cdot K_1^2 \cdot (1 + \frac{(K_2 - R(x_1;x_2))^2}{1 - R^2(x_1;x_2)})$$

We may observe that $R^2_{2x1} > R^2_{2x2} > R^2_{2x3}$.

[1] *In F. M. Pavelescu (2010) we thoroughly analyzed the modeling factors of the coefficients of alignment to collinearity hazard in case of a linear regression with three explanatory variables, which essentialy influence the features of Transformed Form of Student Test statistics. The advantage of the respective linear regression is that it is relatively easy to highlight the modeling factors of the coefficients of alignment to collinearity hazard, i.e. A) the trajectory of differentiation of the Pearson coefficients of correlation between the explanatory variables and the dependent variable, B) the type of collinearity between the explanatory variables, C) the distribution of ranks of the descriptors of dependent variables behaviour in comparison with the distribution of the order of the generators of collinearity.*

In this context, it is possible to have six distributions of ranks of descriptors of dependent variable behaviour in comparison with the order of generators of collinearity. It is to note that the number of such situations is equal to 3!

C) the differentiation of Pearson coefficients of correlation between the explanatory variables ($R(x_j, x_k)$) under the condition of coincidence between the ranks of explanatory variables and the order of generators of collinearity;

D) the real distribution of Pearson coefficients of correlation between the explanatory variables ($R(x_j, x_k)$).

In order to obtain the proposed model of factorial analysis, it is necessary to define a series of indicators which are able to highlight the modelling factors mentioned above.

Therefore, the premises for the Transformed Form of Corrected Student Test statistics may be considered: a) the square root of degrees of freedom; b) absolute value of the Pearson coefficient of correlation between the main (primordial) explanatory variable of the considered linear regression ($/R(x_p, y)/$); c) the Pearson coefficient of correlation between explanatory variables with the highest absolute value (R_{3max}).

If we consider that all the Pearson coefficients of correlation between the explanatory variables are equal to ($/R(x_p, y)/$ and all the Pearson coefficients of correlation between the explanatory variables are equal to R_{3max} it is possible to define the indicator "**Premise of Transformed Form of Student Test statistics**" ($TFST_{m3premis}$), which may be computed by the formula:

$$TFST_{m3\,premis} = (m-4)^{(1/2)} \cdot \left|R(x_p;y)\right| \cdot \left(\frac{1-R_{3\max}}{1+2\cdot R_{3\max} - 3\cdot R(x_p;y)}\right)^{(1/2)} \quad (12)$$

The second indicator which we have to define in order to make the proposed factorial analysis possible is "Transformed Form of Student Test statistics in conditions of real differentiation of the **Pearson coefficients of correlation between the explanatory variables and dependent variable and non-differentiation of the Pearson coefficients of correlation between the explanatory variables**" ($TFST_{m,3diff/R(xk,y)/}$). The indicator considers the real Pearson coefficients of correlation between the explanatory variables and maintain the premise that all the Pearson coefficients of correlation between the explanatory variables are equal to R_{3max}.

Therefore, we may use the computation formula:

$$TFST_{m3diff\,/R(xk,y)/} = (m-4)^{(1/2)} \cdot \left|R(x_k;y)\right| \cdot \left(\frac{1-R_{3\max}}{1+2\cdot R_{3\max} - 3\cdot R(x_p;y)}\right)^{(1/2)} \quad (13)$$

The third indicator needed by the analysis methodology is "**Transformed Form of Student Test statistics in conditions of concordance between the ranks of explanatory variables and the orders of generators of collinearity**" ($TFST_{m,3concord}$). The indicator implies the computation of the Transformed Form of Student Test having in view the real values of the Pearson coefficients of correlation between the explanatory variables and dependent variable ($R(x_k, y)$, on the one hand, and the real values of Pearson coefficients of correlation between the explanatory variables ($R(x_j, x_k)$ (re-)arranged under the condition that the ranks of explanatory variables seen as descriptors of dependent variable are identical to the order of generators of collinearity.

Having in view the above-mentioned indicators, it is possible to detect the contribution modelling factors to Transformed Form of Student Test statistics:

A) The premises of the Transformed Form of Student Test statistics ($TFST_{m3premis}$).

B) Contribution of the real differentiation of the Pearson coefficients of correlation between the explanatory variables and dependent variable ($\Delta Diff\ /\ R(x_k; y)\ /$)

$$\Delta Diff\ /\ R(x_k; y)\ / = TFST_{m,3diff\ /\ R(x_k; y)/} - TFST_{m,3\,premis} \qquad (14)$$

C) Contribution of the real differentiation of the Pearson coefficients of correlation between the explanatory variables ($\Delta DiffR(x_j; x_k)$)

$$\Delta DiffR(x_j; x_k) = TFST_{m,3concord} - TFST_{m,3diff\ /\ R(xk,y)/} \qquad (15)$$

D) Contribution of the non-concordance of the ranks of explanatory variables and the order of generators of collinearity (Δnc_{evgc})

$$\Delta nc_{evgc} = TFST_{m,3xk} - TFST_{m,3concord} \qquad (16)$$

The above-mentioned contributions are determinated for each explanatory variable and also for the three indicators involved in the Synthesis of the Transformed Form of Student Test statistics.

Consequently, if the harmful collinearity is manifest it is possible to determine its primary cause. In other words, we may speak about:

a) harmful collinearity primarily determinated by the differentiation of the Pearson coefficients of correlation between the explanatory variables and dependent variable;

b) harmful collinearity primarily determinated by real differentiation of Pearson coefficients of correlation between the explanatory variables in the context of a standard distribution of those correlation coefficients; and

c) harmful collinearity primarily determinated by the non-concordance between the ranks of descriptors of dependent variables and the order of generators of collinearity.

4. A numerical example. Factorial analysis of the Transformed Form of Student Test statistics in case of estimated parameters of Cobb-Douglas production function with non-constant return to scale and disembodied technical progress in Romania during the 1960-1979 period

In order to illustrate the possibility of practical use of the proposed factorial analysis model for the Transformed Form of Student Test statistics, we have estimated the parameters and computed the values of the above-mentioned test of the Cobb-Douglas production function with non-constant return to scale and disembodied technical progress in the case of Romania during the 1960-1979 period. The above-mentioned production function has the form: $\ln Y = \ln A_3 + g_3 \cdot t + a_3 \cdot \ln FK + b \cdot \ln L$, where:

ln Y= natural logarithm of output (national income) indices;

t= factor time;

ln FK= natural logarithm of fixed capital indices;

ln L = natural logarithm of employed population indices;

A_3 = natural logarithm of residual factor (to be estimated);

g_3= rate of disembodied technical progress (to be estimated);

a_3 = elasticity of output related to fixed capital (to be estimated);

b_3 = elasticity of output related to employed population (to be estimated);

The estimation of parameters has given the following results:

lnY= -0.25670+0.091364*t-0.43293*ln FK+7.721294*lnL R^2_{3Y}=0.998745
(-1.99785) (8.554177) (-2.61557) (4.173538) $F_{20;3}$=4243.675

N.B. In the brackets there are presented the standard Student Test statistics.

R^2_{3Y}= coefficient of determination, $F_{20;3}$= Fisher Test statistics

The examination of the matrix of Pearson coefficients of correlation between the explanatory variables and dependent variable ($R(x_k; y)$, on the one hand, and the Pearson coefficients of correlation between each explanatory variable ($R(x_j, x_k)$) shows that all the indicators are positive (Table 1). In these conditions, the Transformed Form of Student Test statistics is identical with the standard Student Test statistics. Also, it is to note that within the linear regression we have estimated the main (primordial) explanatory variable is factor time (t), while the rank of fixed capital is 2 and the rank of employed population is 3 because we have:

R(t;Y)= 0.9987, R(FK; Y)= 0.9959, R(L, Y) = 0.9945

Table 1 - Matrix of Pearson coefficients of correlation for the linear regression $\ln Y = \ln A_3 + g_3 \cdot t + a_3 \cdot \ln FK + b \cdot \ln L$ **in case of Romania's economy during 1960-1979**

	LnY	T	LnFK	lnL
LnY	1.0000	0.9987	0.9959	0.9945
T	0.9987	1.0000	0.9972	0.9922
LnFK	0.9959	0.9972	1.0000	0.9953
LnL	0.9945	0.9922	0.9953	1.0000

In this context, we are able to compute r_{21}, $r_{32,}$ K_1, K_2.

We have obtained r_{21}=0.9972, r_{32}= 0.9986, K_1=0.9981 K_2=0.9969

On this basis, we may conclude that we deal with an underexponential differentiation of the Pearson coefficients of correlation between the explanatory variables and the dependent variable, on one hand, and with a harmful collinearity between the main generator of collinearity and the generator of collinearity of order 2 related to the generator of collinearity of order 3.

Also, it may be observed that the ranks of descriptors of dependent variables behaviour are not identical with the orders of generators of collinearity.

In order to determine Variance Inflation Factor (VIF), the following linear regressions were estimated:

t= 0.841358 + 12.41424*ln FK+ -7.58557*lnL R^2_{2t}= 0.994462

(4.033005) (5.517697) (-0.18073) $F_{20;2}$= 1526.454

lnFK= -0.04803+ 7.044957*lnL+ 0.05169*t R^2_{2FK}= 0.996656

(-3.24615) (3.347431) (5.517697) $F_{20;2}$= 2533.312

lnL= 0.001297 - 0.00025*t+ 0.056392*ln FK R^2_{2L}= 0.990685

(0.783892) (-0.18073) (3.347431) $F_{20;2}$= 0.990685

On this basis, the Variance Inflation Factor can be computed for each explanatory variable: VIF_t= 180.5828, VIF_{FK} = 299.0367, VIF_L= 107.3536. It is to note that we face a very stressed collinearity.

According to the analysis methodology presented above, we firstly have to determine **the premises of the Transformed Form of Student Test statistics ($T_{m,3premis}$).**

The above-mentioned indicator is equal to 2.92344. $TFST_{premis}$ is high because the Pearson coefficient of correlation between the main (primordial) explanatory variable and the dependent variable has a very high absolute value (0.9987) and is bigger than the highest absolute value of the Pearson coefficient of correlation between the explanatory variables (R_{3max}).

Fisher Test statistics computed in these case ($F_{20;3premis}$) is equal to 7270.586. Consequently, the ratio $\frac{TFST_{20;3\,premis}}{\sqrt{F_{20;3\,premis}}}$ is equal to 0.034285. The very small value of the respective ratio is due to the very intense correlation between the explanatory variables which was taken into consideration.

If the real differentiation of the Pearson coefficients of correlation between the explanatory variables and the dependent variable R(x_k, y) is considered and the assumption regarding the non-differentiation of the Pearson coefficients of correlation between the explanatory variables (R(x_j, x_k)) is maintained, it is possible to obtain:

$TFST_{tdiff/R(xk,y)/}$=1.106402, $TFST_{FKdiff/R(xk,y)/}$=1.103343, $TFST_{Ldiff/R(xk,y)/}$=1.101787

$TFSTAM_{20;3diff\,R(xk,y)/}$=1.103844, $F_{20;3\,diff\,R(xk,y)/}$=983.4529, $\frac{TFST_{20;3diff\,/\,R(xk,y)/}}{\sqrt{F_{20;3diff\,/\,R(xk,y)/}}}$ = 0.035199

$TFSTCV_{20;3diff\,R(xk,y)/}$ = 0.01737.

If we take into consideration the real differentiation of the Pearson coefficients of correlation R(x_k, y) and R(x_j, x_k), but we adopt the assumption of the standard distribution of the Pearson coefficients of correlation R(x_j, x_k), the following results are obtained:

$TFST_{tdiffR(xj,xk)}$=4.265698, $TFST_{FKdiff\,R(xj,xk)}$=0.031453, $TFST_{Ldiff\,R(xj,xk)}$= 0.411554

$TFSTAM_{20;3diff\,R(xj,xk)}$=1.569568, $F_{20;3\,diff\,R(xj,xk)}$=2068.063, $\frac{TFST_{20;3diffR(xj,xk)}}{\sqrt{F_{20;3diffR(xj,xk)}}}$ = 0.034514

$TFSTCV_{20;3diff\,R(xj,xk)}$= 1.218651

It is to note that in conditions of real differentiation of the Pearson coefficients of correlation between the explanatory variables and their standard distribution, i.e. when there is a strict concordance between the ranks of descriptors of dependent variable behaviour and the order of generators of collinearity, we face an important increase of dispersion of the TFST statistics. On the other hand, it is important to observe that the harmful collinearity does not yet occur.

If the real differentiation and distribution of Pearson coefficients R(x_k, y) and R(x_j, x_k) is considered, we obtain the following results:

$TFST_{t20;3}$= 8.554177, $TFST_{FK20;3}$= -2.61557, $TFST_{L20;3}$= 4.173538

$TFSTAM_{20;3}$= 3.370715, $F_{20;3}$=4243.975, $\frac{TFST_{20;3}}{\sqrt{F_{20;3}}}$ = 0.514741

$TFSTCV_{20;3}$= 1.363281

It is important to observe that in this context harmful collinearity is manifest in case of the explanatory variable FK. In other words, we may conclude that the primary cause of the negative coefficient of alignment occurrence is the non-concordance between the ranks of descriptors of dependent variable behaviour and the order of generators of collinearity .

In this context, both TFSTAM and CVTFST grow significantly in comparison with the standard distribution of $R(x_j, x_k)$, from 1.569568 to 3.370715 in case of TFSTAM and from 1.218651 to 1.363281 in case of CVTFST. Also, the ratio $\frac{TFST_{20;3}}{\sqrt{F_{20;3}}}$ increases from 0.34514 to 0.514741.

Having in view the computation presented above we are able to emphasize the contributions of modeling factors to $TFST_{xk}$, $TFSTAM_{20;3}$ $\frac{TFST_{20;3}}{\sqrt{F_{20;3}}}$ and $CVTFST_{20;3}$.

Therefore, it can be observed that the differentiation of the Pearson coefficients of correlation between the explanatory variables and dependent variable acts in favour of a sensible diminution of $TFST_{xk}$ and $TFSTAM_{20;3}$ (Table 2).

Table 2 -- Contributions of modeling factors to

$TFST_{xk}$, TFSTAM20;3 $\frac{TFST_{20;3}}{\sqrt{F_{20;3}}}$ and $CVTFST_{20;3}$ statistics

Indicator	$TFST_{m3premis}$	ΔDiff $(R(x_k,y))$	ΔDiffstd$(R(x_k, xj))$	Δnc_{evgc}
t	2.92345	-1.81704	3.15930	4.28848
FK	2.92345	-1.82010	-1.07189	-2.64702
L	2.92345	-1.82166	-0.69023	3.76198
$\frac{TFST_{20;3}}{\sqrt{F_{20;3}}}$	0.034285	0.000914	-0.00068	0.017227
$TFSTAM_{20;3}$	2.92345	-1.81960	0.46572	1.80115
$CVSTAM_{20,3}$	0.00000	0.00192	1.21691	0.14463

The differentiation of the Pearson coefficients between explanatory variables $(R(x_j, x_k)$ in conditions of standard distribution determinates a sensible increase in TFST related to the main (primordial) explanatory variable and an important decrease in the indicator related to the other two explanatory variables. This evolution have the most important contribution to the final value of $CVTFST_{20;3}$.

The real distribution of the Pearson coefficients between explanatory variables ($R(x_k$, xj) has the most important contribution to the modification of the TFST related to each explanatory variable in comparison with the premises of the respective indicator.

It is to note that in case of $TFSTAM_{20;3}$ the difference between the value obtained in real conditions and the premise value is practically determinated by the differentiation of the Pearson coefficients of correlation $R(x_k, x_j)$ in standard conditions, because the decrease determinated by the differentiation of the Pearson coefficients of correlation $R(x_k, y)$ is compensated in a proportion of 98.99% by the increase generated by the real distribution of the Pearson coefficients of correlation $R(x_j, x_k)$.

On the other hand, it is important to observe that the arithmetical mean of the Transformed Form of Student Test statistics (TFSTAM), 3.370715, is very small in comparison with the square root of Fisher Test statistics, 65.1435. Consequently, the ratio $(TFSTAM/(F_{m,n})^{(1/2)})$ is equal to 5.17%. The very low level of the ratio mentioned above is mainly a consequence of the very high degree of collinearity quantified by VIF.

5. Conclusions

A deep examination of the algebraical properties of the OLS method highlights that the Student Test statistics are correlated with the Fisher Test statistics not only in case of simple linear regressions, but also when we deal with a linear regression with three explanatory variables.

If the computation formula of the Student Test demonstrated in this paper is taken into account in case of a linear regression with three explanatory variables, the most significant correlation between the Student Test and Fisher Test statistics is obtained when the arithmetical mean of the Transformed Form of Student Test statistics ($TFSTAM_{m,n}$) is considered.

The departure of $TFSTAM_{m,n}$ from the square root of the Fisher Test statistics is determinated, on the one hand, by the intensity of collinearity between the explanatory variables, highlighted by the values of variance inflation factor and the coefficient of alignment to collinearity hazard and, on the other hand, by the dispersion of the coefficients of determination recorded in case of simple linear regressions of the dependent variable at each explanatory variable.

Therefore, we may conclude that the ideal condition for running a multiple linear regression appears to be not only the orthogonality of explanatory variables, but also the strict equality between the Pearson coefficients of correlation between the dependent variable and each of the explanatory variables. In such a case, $TFSTAM_{m,n}$ is equal to the square root of the Fisher Test statistics. Also, it is important to mention that, in this context, the coefficient of variation of the Transformed Form of Student Test statistics ($TFSTCV_{m,3}$) is equal to 0.

In practice, the above-mentioned situation is unlikely to occur. Under these conditions, the definition and use of the Synthesis of the Transformed Form of Student Test statistics may act as a useful tool for creating a comprehensive image of the significance of the estimated parameters. In fact, the arithmetical mean of the Transformed Form of Student Test statistics ($TFSTAM_{m,3}$) can be considered as an indicator of the significance of the estimated parameters in a linear regression seen as a whole. The respective arithmetical mean has to be seen in connection with the Fisher Test statistics ($F_{m,n}$) and with the Coefficient of variation of the Transformed Form of Student Test statistics ($CVTFST_{m,n}$). We may consider that a good quality of estimated parameters occurs not only when there is no negative coefficient of alignment to collinearity hazard (T_{nk}) and the $TFSTAM_{m,3}$ is high enough to exceed the critical values of Student Test, but also when the ratio ($TFSTAM_{m.3}/(F_{m,3})^{(1/2)}$) is higher enough, over 50%, and $CVTFST_{m,3}$ is relatively low.

In explaining the Transformed Form of Student Test statistics ($TFST_{bnk}$) recorded in a linear regression with three explanatory values it is very important to consider all the modeling factors from algebraical viewpoint of the respective statistics. The factorial analysis model presented in this paper emphasizes the premises of $TFST_{b3k}$. It highlights the major role of the differentiation of Pearson coefficients of correlation ($R(x_k\ y)$, and $R(x_j,\ x_k)$, relative to the premises of $TFST_{m,3}$. Also, the features of the distribution of the Pearson coefficients of correlation ($R(x_j\ x_k)$) can favour or not the occurrence of harmful collinearity. Theoretically, a non-concordance between the ranks of descriptors of the dependent variable behaviour and the order of generators of collinearity acts in favour of harmful collinearity occurrence, as it was shown by the numerical example presented in this paper.

References

1. Allen Ch., Hall St. (1997)- *Macroeconomic Modeling in a Changing World*, John Wiley & Sons, New York.
2. Aylin A. (2010) – "Multicollinearity. Focus Article", *Wiley Interdisciplinary Reviews*: *Computational Statistics*, Vol. 2, Issue 3.

3. Belsey D. (1976)– "Multicollinearity: Diagnosing its presence and assessing the potential damage it causes to least square estimation", *NBER Working Paper no. 154, Cambridge, Massachusetts, October.*

4. Belsey D. (1991)- "Conditional diagnostics: collinearity and weak data in regression". *Wiley Series in Probability*. John Wiley, New York.

5. Conrad C. (2006)-"Applied Regression Analysis. Lectures Notes", Department of Economics, Pomona College, Claremont, California, spring.

6. Dougherty C. (2002)- *Introduction to econometrics*, Second Edition, Oxford University Press.

7. D. Douglas, D.Clader, J.Cristy, P. Michaelis, D.Belsey (2003)-" Test about harmful collinearity among predictor variables used in modeling global temperature", *Climate Research vol. 24, June.*

8. Farrar D., Glauber R. (1967)- "Multicollinearity in regression analysis: The problem revisited", *Review of Economics and Statistics,* February.

9. Ferguson C. E. (1979) – *The Neoclassical Theory of Production and Distribution*, Cambridge, New York, Melbourne, Cambridge University Press.

10. W. Greene (1993) - *Econometric Analysis*, Mcmillan Publishing Company, New York.

11. Grilichs Z, Ringstad V. (1971)- *Economies of Scale and the Form of the Production Function*, North- Holland Publishing Company, Amsterdam.

12. Isaic - Maniu, C. Mitruț, V. Voineagu (1996)- *Statistica pentru managementul afacerilor* ("Statistics for Business Management"), Economic Publishing House, Bucharest. (in Romanian).

13. L. Klein (1962)- *An introduction to econometrics*, Prentice Hall.

14. Pavelescu F. M. (2010) – "An Extensive Study on the Disturbances Generated by Collinearity in a Linear Regression Model with Three Explanatory Variables", *Romanian Journal of Economics* no.2 (40).

15. Pavelescu F. M. (2012)- Interdependența dintre tipul de colinearitate şi valorile calculate ale testului Student în cazul unei regresii liniare cu două variabile explicative ("Interdependence between the type of colliniarity and the computed values of Student test in case of a linear regression with two explanatory variables"), *Caietul de Studiu nr. 25/2012 al Seminarului de Modelare Macroeconomică al INCE.*

16. Pavelescu F. M. (2013) – Factorii modelatori ai valorilor calculate ale Testului Student în cazul unei regresii liniare cu trei variabile explicative, (Modelling factors of the computed values of Student Test in case of a linear regression with three explanatory variables) *Caietul de Studiu nr. 26/2013 al Seminarului de Modelare Macroeconomică al INCE.*

17. E. Pecican - *Econometrie*, Editura All, Bucureşti, 1994 (Econometrics, ALL Publishing House, Bucharest, 1994).

18. E. Pecican - *Macroeconometrie. Politici economice guvernamentale şi econometrie*. Editura Economică Bucureşti, 1996 (Macroeconometrics. Gouvernamental economic policies and econometrics. Economica Publishing House, Bucharest, 1996).

19. S.D.Sylvey (1969) – "Multicollinearity and imprecise estimation", *Journal of Royal Statistical Society, Series B,* Vol.31.

20. G.W. Stewart (1973) – *Introduction to matrix computation*, Academic Press, New York.

21. R.Wiliams (2008) - *Sociology Graduate Statistics*, University of Notre Dame, Indiana.