Factorii modelatori ai valorilor calculate ale Testului Student în cazul unei regresii liniare cu trei variabile explicative

Florin-Marius Pavelescu[1]

Rezumat. *Pornind de la lucrările şi articolele publicate anterior de către autor, în această lucrare sunt investigate o serie de specificități ale valorilor calculate ale Testului Student în cazul regresiei liniare cu trei variabile explicative. În acest context, se relevă faptul că valențele cognitive ale respectivului model de regresie nu se reduc numai la o interpretare adecvată a semnificației parametrilor estimați în cazul unor modele de tipul funcției de producție Cobb-Douglas cu randamente non-constante şi progres tehnic neîncorporat, funcției de producție Kmenta sau modelului de corectare a erorii, ci şi la o mai bună înțelegere a efectului pe care sporirea numărului de variabile explicative le are asupra creşterii complexității fenomenului colinearității. Este de notat că, în această lucrare, se defineşte colinearitatea ca fiind neîndeplinirea condiției de ortogonalitate între variabilele explicative.*

Astfel, se aduc argumente în favoarea ideii că variabilele explicative pot fi ordonate nu numai din punctul de vedere al intensității corelației cu variabila rezultativă, ci şi al gradului de colinearitate indus de respectivele variabile în ecuația de regresie. De asemenea, se arată că fenomenul colinearității între două variabile explicative se manifestă nu numai în raport cu variabila rezultativă, ci şi în raport de o terță variabilă explicativă. În funcție de tipul de colinearitate dintre primele două variabile explicative generatoare de colinearitate în raport cu cea de a treia se pot identifica variantele regresiei liniare cu trei variabile explicative, respectiv "varianta ordonată" şi "varianta haotică". Pe de altă parte, se demonstrează că valoarea calculată a testului Student este influențată în mod explicit de rădăcina pătrată a factorului de inflamare a dispersiei erorilor (VIF) al variabilelor explicative analizate.

Cuvinte cheie: Testul Student, Rangul variabilelor explicative, Ordinul generatorilor de colinearitate, Factorul de inflamare a dispersiei erorilor şi toleranța,

Clasificarea JEL: C13, C20, C51, C52

[1] Institute of National Economy, Romanian Academy

I. Unele proprietăți algebrice ale Testului Student şi avantajele definirii formei transformate a valorilor calculate ale Testului Student aplicat unei regresii liniare

În metodologia standard de realizare a unei regresii liniare multiple, prin intermediul metodei celor mai mici pătrate, pentru testarea semnificației parametrilor estimați se utilizează testul Student. Însă, pe măsură ce numărul variabilelor se măreşte, valorile calculate ale respectivului test statistic tind să nu mai reflecte legătura care ar trebui să existe, aşa cum remarca D. Belsey, între respectivele valori şi mărimea coeficientului de corelație Pearson dintre variabila rezultativă şi variabila explicativă analizată. Acest fapt se datorează apariției şi intensificării fenomenului de colinearitate, relevat în principal de coeficientul de aliniere la hazardul colinearității, definit inițial în (F.M. Pavelescu, 1986).

Se poate demonstra (F.M: Pavelescu, 2009) că în cazul unei regresii cu n variabile explicative, valoarea calculată a testului Student în forma standard (t_{bnk}) este:

$$t_{bnk} = (m-(n+1))^{(1/2)} * \frac{R(x_k;y)}{(1-R_{ny}^2)^{(1/2)}} * \left(\frac{(R_{jl})_n}{(R_{jl})_{n-1,l\neq k}}\right)^{(1/2)} * T_{nk} \qquad (1), \text{ unde}$$

m= numărul de observații

n= numărul de variabile explicative

$R(x_k;\ y)$= coeficientul de corelație Pearson dintre variabila rezultativă şi variabila explicativă analizată

R^2_{ny}= coeficientul de determinare a regresiei liniare cu n variabile explicative

$$R_n^2 = \sum_{k=1}^{n} R^2(x_k;y) * T_{nk} \qquad (2)$$

$(R_{jl})_n$= determinantul matricii coeficienților de corelație Pearson dintre cele n variabile explicative ale regresiei liniare.

$(R_{jl})_{n-1}$= determinantul matricii coeficienților de corelație Pearson dintre cele n-1 variabile explicative ale regresiei liniare, exclusiv variabila explicativă analizată.

T_{nk}= coeficientul de aliniere la hazardul colinearității

$$T_{nk} = \frac{(R_{j1}, R_{j2}...R_{jk-1}, r_{(xjxk)y}, R_{jk+1}...R_{jn})}{(R_{jk})_n} \qquad j = 1..n, \qquad (3) \text{ unde :}$$

R_{jk} = coeficientul de corelație Pearson dintre variabilele explicative x_j şi x_k.

$$r_{(xjxk)y} = \frac{R(x_j;y)}{R(x_k;y)} \qquad (4) \text{ unde:}$$

$R(x_j;y)$ = coeficientul de corelație Pearson dintre variabila explicativă x_j şi variabila rezultativă.

$R(x_k{:}y)$ = coeficientul de corelație Pearson dintre variabila explicativă x_k şi variabila rezultativă.

Este de notat că în F.M. Pavelescu (2010) raportul $r_{(xjxk)y} = \frac{R(x_j;y)}{R(x_k;y)}$ a fost definit drept "coeficient de corelație mediată de variabila rezultativă între variabila explicativă x_j şi variabila explicativă x_k". S-a avut în vedere ierarhizarea variabilelor explicative în funcție de valoarea absolută a coeficienților de corelație Pearson, astfel încât $r_{jk}<1$.

În aceste condiții, se poate defini variabila explicativă primordială a unui model de regresie liniară (x_p) ca fiind acea variabilă explicativă care este cel mai puternic corelată în sens Pearson cu variabila rezultativă. Celelalte variabile rezultative vor putea fi considerate ca având rangul 2…..n în funcție de ordinea dată de valoarea absolută a coeficienților de corelație.

Se poate scrie: $\frac{(R_{jl})_n}{(R_{jl})_{n-1,l\neq k}} = 1 - c_{xk}$, (5) unde

c_{xk}= gradul de colinearitate indus de prezența variabilei explicative x_k în cadrul ecuației de regresie liniară cu n variabile explicative.

În consecință, valoarea calculată a testului Student în formă standard se poate scrie:

$$t_{bnk} = [m-(n+1)]^{1/2} * R(x_k; y) * (\frac{1-c_{xk}}{1-R_n^2})^{1/2} * T_{nk} \quad (6)$$

Formula (6) relevă faptul că valoarea calculată a testului Student în formă standard este dependentă de mărimea absolută a coeficientului de corelație dintre variabila explicativă analizată şi de variabila rezultativă, de numărul gradelor de libertate, precum şi de valoarea coeficientului de determinare a modelului de regresie liniară avut în vedere.

Totodată, valorile calculate ale testului Student sunt dependente şi de intensitatea fenomenului de colinearitate, definită în prezentul studiu prin non-ortogonalitatea variabilelor explicative[2], care are drept efect distorsionarea respectivelor valori. Distorsiunile determinate de colinearitate asupra mărimii parametrilor estimați şi a testului Student sunt reprezentate de valorile coeficientului de aliniere la hazardul colinearității (T_{nk}) şi cele ale gradului de colinearitate îndus de variabila explicativă analizată.

Este important să se aibă în vedere că dacă toate variabilele explicative sunt ortogonale, atunci $c_{xk}=0$ şi $T_{nk}=1$.

Una dintre dificultățile care apar în aprecierea semnificației parametrilor estimați este atunci când valorile coeficientului de aliniere la hazardul colinearității (T_{nk}) sunt negative sau pozitive, dar foarte apropiate de zero. În acest caz, avem de a face cu colinearitatea toxică şi respectiv degradantă.[3]

Se poate observa că **forma standard a testului Student nu permite evidențierea directă a colinearității toxice.** Din acest motiv, în F.M. Pavelescu (2009) a fost propusă realizarea unei modificări a formei standard a valorilor calculate ale testului Student şi care a fost denumită "Testul Student Corectat" (CST). Între valoarea calculată a Testului Student în formă standard aferent variabilei explicative b_{nk} (t_{bnk}) şi cea a Testului Student Corectat (CST_{bnk}) există următoarea relație:

[2] În A. Alin (2010) se arată că (multi)colinearitatea se referă la o relație lineară dintre două sau mai multe variabile care înseamnă implicit şi absența ortgonalității dintre respectivele variabile. Din aceste motive, o modalitate de examinare a colinearității este studiul matricii coeficienților de corelație Pearson dintre variabilele explicative. Dar pentru rigoarea analizei este necesar să se aibă în vedere că între valoarea respectivilor coficienți de corelație şi (multi)colinearitate nu se poate pune în mod automat semnul egalității.

[3] Problematica definirii colinearității toxice şi respectiv degradante în funcție de valorile coeficientului de aliniere la hazardul colinearității, şi de la definițiile date respectivului fenomen de către D. Belsey şi G. Ferrar şi R. Glauber a fost examinată pe larg în F.M. Pavelescu (2011).şi F.M. Pavelescu (2012).

$$CST_{bnk} = \frac{/R(x_k;y)/}{R(x_k;y)} * t_{bnk}$$ (7), echivalent cu:

$$CST_{bxk} = (m-n-1)^{(1/2)} * /R(x_k;y)/* \frac{(1-c_{xk})^{(1/2)}}{(1-R_n)^{(1/2)}} * T_{xk} \qquad (8)$$

Se avea în vedere menţinerea tuturor ipotezelor care definesc forma standard a testului Student. Singura deosebire era utilizarea în cazul Testului Student Corectat a scalarului $\frac{/R(x_k;y)/}{R(x_k;y)}$ pentru evidenţierea directă a semnului coeficientului de aliniere la hazardul colinearităţii şi, pe această cale, detectarea existenţei colinearităţii toxice.

O privire mai atentă a propunerii de detectare directă a existenţei colinearităţii toxice în cadrul valorilor estimate a parametrilor unei regresii liniare relevă faptul că nu avem de a face cu un nou test statistic propriu-zis, ci cu o formă transformată a valorilor calculate ale testului Student. Din acest motiv, în continuarea prezentului studiu se va utiliza noţiunea de "Formă transformată a valorilor calculate ale testului Student" (TFST) în locul noţiunii de „Test Student Corectat".

Este important de subliniat faptul că la forma transformată a valorilor calculate ale testului Student (TFST) se va face apel doar în cazul testării semnificaţiei parametrilor estimaţi ai regresiilor liniare. În celelalte cazuri unde este necesară utilizarea testului Student se va folosi forma standard a respectivului test[4].

Atât în cazul formei standard cât şi în cazul formei transformate a valorilor calculate ale testului Student, este recomandabilă punerea în evidenţă a factorilor modelatori în conformitate atât cu consideraţiile făcute în lucrări referitoare la studiul colinearităţii, care au devenit "clasice" din punct de vedere teoretic cât şi cu consecinţele asupra valorilor estimate ale parametrilor ecuaţiei de regresie şi a valorilor calculate ale unor teste statistice[5].

---

[4] Utilizarea formei transformate a valorilor calculate ale testului Student (TFST) ar duce la o modificare a metodologiei de estimare a parametrilor regresiilor liniare şi cu deosebire a celor multiple, prin introducerea unei etape suplimentare şi anume:

La începutul procedurii de estimare se calculează valorile parametrilor, concomitent cu determinarea formei transformate a valorilor calculate ale testului Student (TFST).

În cazul încare apar valori negative ale TFST, se poate concluziona că în regresia liniară respectivă există colinearitate toxică şi se poate decide neacceptarea rezultatelor obţinute.

În cazul când nu există valori negative ale TFST se concluzionează că nu fenomenul colinearităţii toxice nu se manifestă. În aceste condiţii, se poate continua procesul de estimare a parametrilor conform metodologiei clasice.

[5] Deoarece valoarea calculată a Testului Student în forma standard sau transformată este un produs al acţiunii mai multor factori, se poate pune problema regrupării termenilor şi în cadrul unor indicatori cu o putere explicativă mai mare.

Astfel, pornind de la consideraţiile expuse în Belsey (1991), potrivit cărora în mod normal valorile calculate ale Testului Student ar trebui să reflecte valorile coeficientului de corelaţie Pearson dintre variabila explicativă analizată şi variabila rezultativă, în F. M. Pavelescu (2012) a fost definit indicatorul "Valoarea de referinţă a Testului Student (Corectat) pentru variabila x_k" ($RV_{studxnk}$) , respectiv: $RV_{studxk} = [m-n-1]^{1/2} * /R(y,x_k)/$

Pe de altă parte, avem în vedere că în D.Ferrar şi R.Glauber (1967) se considera că se poate defini colinearitatea toxică nu numai atunci când semnul parametrilor estimaţi este contrar celui anticipat, ci şi atunci când pătratul coeficientului de corelaţie Pearson dintre variabilele explicative este mai mare decât coeficientul de determinare al regresiei liniare multiple. Pornind de la ipotezele respective, se poate defini indicatorul "Coeficientul de dominare a

Dacă se fac următoarele notații:

$$r_{(nkp)} = \frac{|R(x_k;y)|}{|R(x_p;y)|} \text{ (9) şi } q_{nk} = (\frac{1-c_{xnk}}{1-c_{n\max}})^{(1/2)} \qquad (10),$$

unde:

/R(x_p;y)/= valoarea absolută a coeficientului de corelație Pearson dintre variabila explicativă primordială a modelului de regresie şi variabila rezultativă. (**Prin variabilă explicativă primordială a modelului de regresie se înțelege variabila explicativă care este în mod absolut cel mai puternic corelată cu variabila rezultativă).**

c_{max}= gradul de colinearitate maxim care apare în cadrul ecuației de regresie

se poate scrie:

$$TFST = (m-(n+1))^{(1/2)} * |R(x_k;y)| * (\frac{1-c_{n\max}}{1-R_{ny}^2})^{(1/2)} * r_{(nkp)y} * q_{nk} * T_{nk} \qquad (11)$$

Ecuația (11) are avantajul de a releva faptul că valoarea calculată a formei transformate a valorilor calculate ale testului Student (TFST) este dependentă de factori comuni modelului de regresie liniară şi factori specifici fiecăreia dintre variabilele explicative.

Factorii comuni ai TFST aplicate regresiei liniare analizate sunt:

valoarea de referință (RV_{studn}) $RV_{studn} = (m-(n+1))^{(1/2)} * |R(x_k;y)|$ (12) şi

coeficientul de dominare a gradului de colinearitate maximă de către coeficientul de determinare al regresiei liniare cu n variabile explicative ($Cddc_{nmax}$) $Cddc_{n\max} = \frac{(1-c_{n\max})^{(1/2)}}{(1-R_{ny}^2)}$ (13)

Factorii specifici fiecăreia dintre variabilele explicative ale regresiei liniare avute în vedere sunt rapoartele $r_{(nkp)y}$ şi q_{nk}, precum şi coeficientul de aliniere la hazardul colinearității (T_{nk}). De asemenea, se mai evidențiază faptul că rapoartele r_{xp} şi q_k reprezintă premisele valorilor coeficienților de aliniere la hazardul colinearității (T_{nk}). În fapt, impactul avut de $r_{(xp)y}$ şi q_{nk} asupra mărimii coeficientului T_{nk} este unul mediat de numărul de variabile explicative avute în vedere în cadrul ecuației de regresie. În acest context, un caz interesant al diferențierii TFST aferent fiecărei variabile explicative îl constituie regresia liniară cu trei variabile explicative.

---

gradului de colinearitate de către coeficientul de determinare al ecuației de regresie relativă la varibila rezultativă y" (Cddc), respectiv: $Cddc = \frac{(1-R^2(x_p;x_s))^{(1/2)}}{(1-R^2_{2y})^{(1/2)}}$. Este de remarcat că în cazul regresiei liniare cu două variabile explicative, gradul de colinearitate indus de adăugarea unei noi variabile este identic, indiferent dacă respectiva variabilă este principală sau secundară. Totodată, se observă că gradul de colinearitate indus este egal cu coeficientul de determinare corespunzător ecuației de regresie simplă dintre cele două variabile explicative.

În consecință, pentru determinarea valorilor calculate ale formei tranformate a Testului Student aplicat unei regresii liniare cu două variabile explicative se poate folosi formula: $CST_{x2k} = RV_{studx2k} * Cddc * T_{2k}$.

II. Specificitatea valorilor calculate ale Testului Student în cazul regresiei liniare cu trei variabile rezultative.

Pentru asigurarea coerenței demersului investigativ al factorilor modelatori ai valorilor calculate ale testului Student în contextul unei regresii cu trei variabile explicative, se va avea în vedere cazul în care toți coeficienții de corelație Pearson dintre variabila rezultativă şi variabilele explicative ($R(x_k;y)$) sunt pozitivi. Dacă există coeficienți $R(x_k;y)$ negativi, se vor efectua, în cadrul ecuațiilor prin care sunt determinați coeficienții de aliniere la hazardul ccolinearității, transformări algebrice pentru asigurarea premiselor arătate anterior. În acest fel, sunt create premisele obținerii formei transformate a valorilor calculate ale testului Student.

Dacă un coeficient de corelație Pearson între două variabile explicative este negativ, atunci este recomandabil să se realizeze o transformare algebrică în cadrul matricii coeficienților de corelație, astfel încât semnul negativ să fie ataşat coeficientului de corelație Pearson cu valoarea absolută cea mai mică.

Calculul coeficienților de corelație mediată de variabila rezultativă dintre variabilele explicative şi variabila explicativă primordială a modelului de regresie liniară permite determinarea rangurilor variabilelor explicative. Astfel, variabila explicativă primordială va avea rangul 1, iar variabila explicativă cel mai slab corelată cu variabila rezultativă va avea rangul n.

Pe de altă parte, valorile rapoartelor q_{3k} permit evaluarea şi ordonarea variabilelor explicative în calitate de „generatori de colinearitate"[6] Astfel, variabila explicativă cu cel mai ridicat grad de colinearitate indus în ecuația de regresie (c_{max}) va fi denumită "generator principal de colinearitate", în timp ce variabila explicativă cu cel mai scăzut grad de colinearitate indus în ecuația de regresie va reprezenta generatorul de colinearitate de ordinul n.

Gradul de colinearitate indus de cele trei variabile explicative în ecuația de regresie este dependent în principal de raportul în care se află coeficienții de corelație Pearson dintre variabilele explicative ($R(x_j;x_k)$) Dacă se notează :

R_{3max}= Coeficientul de corelație Pearson dintre variabilele explicative cu valoarea absolută cea mai ridicată.

K_1= raportul dintre mărimea coeficientului de corelație Pearson aflat pe locul doi din punct de vedere al valorii absolute şi cea a coeficientului de corelație cu valoarea absolută cea mai ridicată ($0<K_1<1$)[7].

K_2= raportul dintre mărimile coeficienților de corelație Pearson aflați pe locul trei şi respectiv doi din punct de vedere al valorii absolute ($K_2<1$)[8].

---

[6] În F. M. Pavelescu (2004c) şi (2010b) a fost utilizată pentru ceea ce în prezenta lucrare se defineşte ca "generatori de colinearitate" denumirea de "atractori ai colinearității". Dar noțiunea de "atractor" ar însemna existența unui proces dinamic care ar tinde spre o anumită stabilitate. În cazul de față, nu este vorba decât de o estimare de natură statică a unor parametri.

[7] Este de notat că K_1 reprezintă coeficientul de corelație mediată de generatorul principal al colinearității dintre generatorul de ordinul 3 şi generatorul de ordinul 2

[8] Se observă că K_2 reprezintă coeficientul de corelație mediată de generatorul colinearității de ordinul 3 dintre generatorul de ordinul 2 şi generatorul principal al colinearității.

se obțin gradele de colinearitate induse de fiecare dintre variabilele explicative şi anume:
pentru generatorul principal de colinearității (c_{3max})

$$c_{3\max} = \frac{R^2_{3\max} * (1 - 2 * K_1^2 * K_2 + K_1^2)}{1 - R^2_{3\max} * K_1^2 * K_2^2} \quad (14),$$

echivalent cu:

$$c_{3\max} = R^2_{3\max} * (1 + \frac{(K_1 - R_{3\max} * K_1 * K_2)^2}{1 - R^2_{3\max} * K_1^2 * K_2^2}) \quad (15)$$

Dacă se are în vedere formula de calcul a coeficientului de determinare în cazul unei regresii liniare cu două variabile explicative prezentată în F.M. Pavelescu (2004a)[9] Se observă că:

$$c_{3\max} = R^2_{cp} \quad (16)$$

unde:

R^2_{2cp}= coeficientul de determinare al regresiei liniare $X_{cp} = \alpha_{cp} + \beta_{cp} * X_{c2} + \gamma_{cp} * X_{c3}$, unde:

X_{cp}= variabila explicativă care reprezintă generatorul principal de colinearitate în ecuația de regresie liniară cu trei variabile explicative.

X_{c2}= variabila explicativă care reprezintă generatorul de colinearitate de ordinul 2 în ecuația de regresie liniară cu trei variabile explicative.

X_{c3}= variabila explicativă care reprezintă generatorul de colinearitate de ordinul 3 în ecuația de regresie liniară cu trei variabile explicative.

pentru generatorul de ordinul al doilea al colinearității (c_{32})

$$c_{32} = \frac{R^2_{3\max} * (1 - 2 * R_{\max} * K_1^2 * K_2 + K_1^2 * K_2^2)}{1 - R^2_{3\max} * K_1^2} \quad (18).$$

Expresia (18) poate fi scrisă şi sub forma:

$$c_{32} = R^2_{3\max} * (1 + \frac{(K_1 * K_2 - R_{3\max} * K_2)^2}{1 - R^2_{3\max} * K_1^2}) \quad (19), \text{ echivalent cu:}$$

$c_{32} = R^2_{2c2}$ (20), unde:

R^2_{2c2}= coeficientul de determinare al regresiei liniare $X_{c2} = \alpha_{c2} + \beta_{c2} * X_{cp} + \gamma_{c2} * X_{c3}$

pentru generatorul de colinearitate de ordinul al treilea (c_{33})

---

[9] În F.M. Pavelescu (2004a) se demonstrează faptul că în cazul unei regresii liniare cu două variabile explicative (R^2_y) coeficientul de determinare poate fi scris şi sub forma: $R^2_{2y} = R^2(x_p; y) * (1 + \frac{(r_{2(sp)y} - R(x_p; x_s))^2}{1 - R^2(x_p; x_s)})$, unde :

$R(x_p;y)$ = coeficientul de corelație dintre variabila explicativă primordială şi variabila rezultativă.

$R(x_p;x_s)$ = coeficientul de corelație dintre variabila explicativă primordială şi variabila explicativă secundară.

$R_{2(sp)y}$= coeficientul de corelație mediată de variabila rezultativă dintre variabila explicativă primordială şi variabila explicativă secundară

$$c_{33} = \frac{R^2_{3\max} * K_1^2 * (1 - 2 * R_{\max} * K_2 + K_2^2)}{1 - R^2_{3\max}} \qquad (21)$$

Expresia (21) poate fi scrisă $c_{33} = R^2_{3\max} * K_1^2 * (1 + \frac{(K_2 - R_{3\max})^2}{1 - R^2_{3\max}})$ (22),

echivalent cu $c_{32} = R^2_{c3}$ (23),

Unde R^2_{2c3}= coeficientul de determinare al regresiei liniare $X_{c3}= \alpha_{c3}+\beta_{c3}*X_{cp}+\gamma_{c3}*X_{c2}$

În aceste condiţii, se poate face legătura între valoarea calculată a testului Student şi unul dintre indicatorii utilizaţi în literatura de specialitate pentru evaluarea intensităţii fenomenului de colinearitate şi anume "Factorul de inflamare al dispersiei erorilor" (VIF)[10]. De asemenea, se poate utiliza şi inversul indicatorului amintit anterior, denumit "Toleranţă" (TOL)[11].

Cu alte cuvinte, în cazul unei regresii liniare cu trei variabile explicative, valorile calculate ale testului Student în forma standard (t_{bnk}) pot fi determinate şi cu ajutorul formulei:

$$t_{bnk} = (m - (n+1)^{(1/2)} * R(x_k; y) * (\frac{TOL(x_k)}{1 - R^2_{ny}})^{(1/2)} * T_{nk} \qquad (24).$$

În forma transformată a valorilor calculate ale testului Student (TFST) formula de calcul utilizată devine:

$$TFST_{bnk} = (m - (n+1))^{(1/2)} * |R(x_k; y)| * (\frac{TOL(x_k)}{1 - R^2_{ny}})^{(1/2)} * T_{nk} \qquad (25).$$

De asemenea, se poate veni în sprijinul ipotezei formulate de către Farrar şi Glauber în 1967 potrivit căreia o intensitate mai mare a corelaţiei dintre variabila explicativă analizată şi ansamblul celorlate variabile explicative decât cea a variabilei rezultative în raport cu ansamblul variabilelor explicative, creează premisele apariţiei colinearităţii toxice.

Pe de altă parte, se poate demonstra (F.M. Pavelescu, 2011) că valorile coeficienţilor de aliniere la hazardul colinearităţii (T_{nk}) sunt determinate de: **a) valoarea coeficientului R_{max} b) ritmul şi traiectoria de diferenţiere a coeficienţilor de corelaţie mediată de variabila rezultativă între variabilele explicative, c) tipul de colinearitate relativă mediată de generatorul de colinearitate de ordinul 3 între generatorul principal şi generatorul de colinearitate de ordinul 2, d) distribuţia ordinelor generatorilor de colinearitate în raport cu rangurile variabilelor explicative.**

---

[10] Calculul factorului de inflamare a dispersiei erorilor (variance inflation factor-VIF) presupune efectuarea unor regresii liniare a fiecăreia dintre variabilele explicative ale modelului în raport cu celelalte variabile explicative. Se calculează coeficientul de determinaţie aferent respectivelor regresii şi apoi se determină VIF pentru fiecare dintre variabilele explicative (X_j), conform formulei: $VIF(X_j) = \frac{1}{1 - R^2_{xj}}$.

De regulă, se consideră că dacă $VIF(X_j) \geq 5$, se consideră că multicolinearitatea este ridicată.

[11] Avînd în vedere relaţia dintre $VIF(X_j)$ şi $TOL(X_j)$ se poate deduce că multicolinearitatea poate fi considerată ca fiind ridicată dacă $TOL(X_j)<0.2$, echivalent cu $R^2_{xj}>0.8$.

Astfel, **creşterea valorii coeficientului R_{nmax}** este de natură să polarizeze valorile coeficienților T_{nk} sau să crească probabililitatea apariției unor valori negative ale acestora.

Ritmul mediu de diferențiere a coeficienților de corelație mediată de variabila rezultativă între variabilele explicative, respectiv media geometrică a coeficienților de corelație mediată r_{21} şi r_{32}, crează premisele pentru o anumită polarizare a coeficienților T_{nk}. Dar contribuția efectivă la polarizarea respectivilor coeficienți este dată de t**raiectoria diferențierii coeficienților de corelație mediată de variabila rezultativă între variabilele explicative**.

Se pot identifica trei traiectorii de diferențiere a coeficienților de corelație mediată de variabila rezultativă între variabilele explicative: a) exponențială, dacă $r_{21}=r_{32}$, b) subexponențială dacă $r_{21}>r_{32}$, şi c) supraexponențială dacă $r_{21}<r_{32}$.

În F. M. Pavelescu (2005 şi 2010 b) se arată că determinantul matricii coeficienților de corelație Pearson dintre variabilele explicative det(R_{jk}) poate fi exprimat prin prisma generatorilor de colinearitate, respectiv:

a) prin prisma generatorului principal de colinearitate (x_{cp})

$$\det(R_{jk}) = 1 - R_{3\max}^2 K_1^2 K_2^2 - R_{3\max} * A - R_{3\max} K_1 * B \tag{26}$$

b) prin prisma generatorului de colinearitate de ordinul al doilea (x_{c2})

$$\det(R_{jk}) = 1 - R_{3\max}^2 K_1^2 - R_{3\max} * A - R_{3\max} * K_1 * K_2 * C \tag{27}$$

prin prisma generatorului de colinearitate de ordinul al treilea (x_{c3})

$$\det(R_{jk}) = 1 - R_{3\max}^2 - R_{3\max} * K_1 * B - R_{3\max} * K_1 * K_2 * C, \tag{28}$$ unde:

$$A = R_{3\max} - R_{3\max}^2 * K_1^2 * K_2 \tag{29}$$

$$B = R_{3\max} * K_1 - R_{3\max}^2 * K_1 * K_2 \tag{30}$$

$$C = R_{3\max} * K_1 * K_2 - R_{3\max}^2 * K_1 \tag{31}$$

Se observă că A>B>0 şi B>C>0 dacă $K_2>R_{3max}$ şi că A>B>0 , iar C<0, dacă $K_2< R_{3max}$

În condițiile unei nediferențieri a coeficienților de corelație mediată de variabila explicativă, respectiv $r_{12}=r_{23}=1$, valorile coeficienților de aliniere la hazardul colinearității corespunzători celor trei generatori de colinearitate sunt:

$$T_{xcp} = \frac{(1 - R_{\max} * K_1 * K_2) * (1 + R_{\max} * K_1 * K_2 - R_{\max} - R_{\max} * K_1)}{(R_{jk})_3} \tag{32}$$

$$T_{xc2} = \frac{(1 - R_{\max} * K_1) * (1 + R_{\max} * K_1 - R_{\max} - R_{\max} * K_1 * K_2)}{(R_{jk})_3} \tag{33}$$

$$T_{xc3} = \frac{(1 - R_{\max}) * (1 + R_{\max} - R_{\max} * K_1 - R_{\max} * K_1 * K_2)}{(R_{jk})_3} \tag{34}$$

Rezultă că, în cazul nediferențierii coeficienților de corelație mediată de variabila rezultativă dintre variabilele explicative, valori negative ale coeficienților de aliniere la hazardul colinearității pot apare numai la generatorul principal de colinearitate, doar când $K_2<R_{3max}$.

Deoarece raportul K_2 reprezintă coeficientul de corelație mediată de generatorul de colinearitate de ordinul 3 dintre generatorul principal şi generatorul de colinearitate de ordinul al doilea, rezultă că în **fapt colinearitatea dintre două variabile explicative se poate defini nu numai în**

relație cu variabila rezultativă, ci şi cu o terță variabilă explicativă. În consecință, se pot identifica şi în cazul raportării la o terță variabilă explicativă tipurile de colinearitate puse cu claritate în evidență în cazul unei regresii liniare cu două variabile explicative în F. M. Pavelescu (2012)[12] şi anume:

a) colinearitate non-toxică relativă la o terță variabilă explicativă, dacă $1>K_2>R_{3max}$

b) colinearitate toxică relativă la o terță variabilă explicativă, dacă $R_{3max}>K_2>0$

c) anticolinearitate relativă la o terță variabilă explicativă, dacă $K_2<0$

Dacă 1>K2>R3max atunci dispersia valorilor determinanților A, B şi C este mai mică. În consecință, din punct de vedere teoretic, există o probabilitate mai mică de apariție a unor valori negative ale coeficienților Tnk. Drept urmare, această situație ar putea reprezenta "varianta ordonată", în ceea ce priveşte regresia liniară cu trei variabile explicative.

Dacă K2<R3max se poate vorbi de o "variantă haotică" a regresiei liniare cu trei variabile explicative, în care valorile determinanților A, B şi C sunt sensibil mai dispersate comparativ cu varianta ordonată. În aceste condiții, există premisele ca valorile coeficienților Tnk să prezinte un grad mai mare de deviație de la valoarea de referință a testului Student.

Deviația de la valoarea de referință a indicatorului TFST nu este determinată doar de tipul de colinearitate dintre generatorul principal şi generatorul de ordinul 2 în raport cu cea de a treia variabilă explicativă, ci şi de distribuția generatorilor de colinearitate în raport cu rangurile variabilelor explicative. Din punct de vedere teoretic, există 6 situații posibile de distribuție a variabilelor explicative în funcție de rang (al intensității corelației cu variabila rezultativă) şi ordinul generatorilor de colinearitate (tabelul nr. 1).

Pentru evitarea apariției semnului negativ al coeficientului Tnk şi implicit a colinearității toxice, cea mai favorabilă situație este atunci când există o concordanță între rang şi ordinul generatorilor de colinearitate pentru toate cele trei variabile explicative (cazul 1). De asemenea, în acest caz, se asigură premisele unei variații mai mici a TFST corespunzătoare diferitelor variabile explicative.

[12] În F.M. Pavelescu (2012) s-a demonstrat că tipul de colinearitate în cazul unei regresii liniare cu două variabile exlicative se poate stabili prin prisma valorii coeficientului de corelație mediată de variabila rezultativă între cele două variabile explicative ($r_{(sp)y}$), respectiv:

colinearitate laxă (acceptabilă) dacă $\frac{2*R(x_p;x_s)}{1+R^2(x_p;x_s)} < r_{(sp)y} < 1$

colinearitate degradantă dacă $R(x_p;x_s) < r_{(sp)y} < \frac{2*R(x_p;x_s)}{1+R^2(x_p;x_s)}$

colinearitate toxică dacă $0 < r_{(sp)y} < R(x_p;x_s)$

anticolinearitate dacă $r_{(sp)y} < 0$

Tabelul nr.1

Situații posibile in cazul unei regresii liniare trifactoriale

Cazul		Rangul variabilei explicative		
		Primordială	Rang 2	Rang 3
1	Ordin generator colinearitate	Principal	2	3
2	Ordin generator colinearitate	Principal	3	2
3	Ordin generator colinearitate	2	Principal	3
4	Ordin generator colinearitate	2	3	Principal
5	Ordin generator colinearitate	3	Principal	2
6	Ordin generator colinearitate	3	2	Principal

Situația care sporeşte cel mai mult din punct de vedere teoretic şansele apariției colinearității toxice este în cazul 6, când există o inversare a rangului şi ordinului de colinearitate între variabila primordială şi variabila de rang 3. În acest caz, probabilitatea apariției unui coeficient T_{3k} negativ se majorează dacă între generatorul principal şi generatorul de colinearitate de ordinul al doilea există o colinearitate toxică în raport cu cea de a treia variabilă explicativă.

In evaluarea factorilor care determină valori negative ale coeficientului T3k este necesar să se aibă în vedere şi tipul colinearității în raport cu variabila rezultativă dintre variabila explicativă analizată şi fiecare dintre celelalte variabile explicative.

III. Un exemplu numeric de analiză a valorilor calculate ale Testului Student aplicat unei regresii liniare cu trei variabile explicative

Pentru ilustrarea acțiunii factorilor modelatori ai valorilor calculate ale testului Student în forma standard sau forma transformată identificați anterior, au fost estimați parametrii unui model de regresie liniară referitor la evoluția valorii importului României, exprimat în milioane euro (M_t) în perioada 1991-2009.

Variabile explicative care au fost luate în considerare sunt:

a) valoarea importului din anul precedent (M_{t-1}),

b) indicele consumului privat din anul precedent ($ICHc_{t-1}$) şi

c) indicele formării brute a capitalului fix din anul curent ($IGFCFc_t$).

Cu alte cuvinte modelul estimat este: $M_t=c(1)+c(2)*M_{t-1}+c(3)*ICHc_{t-1}+c(4)*IGFCFc_t$

Rezultatul estimării ecuației de regresie menționate anterior a relevat un coeficient de determinare (R^2_{3Mt}) de 0,91238. Valorile estimate pentru fiecare parametru, precum şi valorile calculate ale testului Student şi ale probabilității ipotezei nule sunt prezentate în tabelul 2.

Tabelul nr. 2

Valorile estimate ale parametrilor şi valorile calculate ale testului Student pentru un model de regresie referitor la evoluţia importului României în perioada 1991-2009

Parametru	Valoare estimată parametru	Valoare calculată test Student	Probabilitate ipoteză nulă
C(1)	-25,3321	-1,83214	0,086859
C(2)	0,8358	8,290522	5,53E-07
C(3)	6,1985	0,447150	0,661156
C(4)	23,0834	4.224456	0,000736

Valorile calculate ale testului Student coincid cu forma transformată a valorilor calculate ale testului Student, datorită faptului că toţi parametrii estimaţi pentru variabilele explicative au acelaşi semn cu coeficienţii de corelaţie Pearson dintre variabila rezultativă (M_t) şi variabilele explicative (sunt toţi pozitivi). Valorile absolute ale coeficienţilor de corelaţie Pearson dintre variabila rezultativă şi variabilele explicative ($R(x_k;y)$, prezentate în tabelul nr. 3, arată că variabila explicativă primordială este M_{t-1}, variabila explicativă de rangul 2 este $ICHc_{t-1}$, iar variabila explicativă de rangul 3 este $IGFCF_{t.}$

Tabelul nr. 3

Matricea coeficienţilor de corelaţie Pearson dintre variabila rezultativă şi variabilele explicative ale unui model de regresie referitor la evoluţia importului României în perioada 1991-2009

	M_t	M_{t-1}	$ICHc_{t-1}$	$IGFCFc_t$
M_t	1	0.899258	0.613494	0.424475
Rangul variabilelor explicative	…	1	2	3
M_{t-1}	0.899258	1	0.65015	0.116792
$ICHc_{t-1}$	0.613494	0.65015	1	0.085230
$IGFCF_t$	0.424475	0.116792	0.08523	1
Ordinul generatorilor de colinearitate	…	1	2	3

Se observă că valorile absolute ale coeficienţilor de corelaţie Pearson dintre variabilele explicative relevă faptul că există o concordanţă între rangurile variabilelor explicative şi ordinele generatorilor de colinearitate.

Pe baza coeficienţilor Pearson se pot obţine coeficienţii de corelaţie mediată, respectiv $r_{3(kp)y}$, K_1 şi K_2.

$r_{3(2p)y}$= 0.682223; $r_{3(3p)y}$=0,472028; $r_{3(32)y}$=0.691898

K_1= 0,179639; K_2= 0.729760

N.B. *Au fost avute în vedere următoarele notaţii referitoare la coeficienţii de corelaţie mediată de variabile rezultativă* M_t= y, M_{t-1}= x_p $ICHc_{t-1}$= x_2 $IGFCFc_t$= x_3

Determinarea valorilor coeficienţilor de corelaţie mediată dintre variabilele explicative permite să se releve faptul că traiectoria de diferenţiere a coeficienţilor de corelaţie Pearson dintre variabila rezultativă şi variabilele explicative ($R(x_k;y)$ este cvasi-exponenţială, cu o uşoară tendinţă de subexponențialitate. Este de menționat faptul că diferențierea se face în ritm rapid, media geometrică a coeficienților $r_{3(k-1k)y}$ fiind de 0,6874.

Se observă că în raport cu variabila rezultativă, între variabilele explicative nu există colinearitate toxică. Nici în raport cu cea de a treia variabilă explicativă, între generatorul principal şi generatorul de colinearitate de ordinul al doilea nu există colinearitate toxică. Cu alte cuvinte, modelul estimat se încadrează în cazul "ordonat" al regresiei liniare cu trei variabile explicative.

Calculul gradului de colinearitate indus de fiecare dintre variabilele explicative prin raportul, conduce la următoarele rezultate:

c_{Mt-1}= 0,42649, $c_{ICHct-1}$= 0,42278, $c_{IGFCFct}$= 0,0,1379.

Efectuarea de regresii liniare în care fiecare dintre cele trei variabile explicative este considerată ca fiind variabilă rezultativă, iar celelate două ca fiind variabile explicative relevă faptul că se confirmă egalitatea $c_{3xk}=R^2_{2xk}$

Rezultatele regresiilor liniare dintre cele trei variabile explicative sunt următoarele:

M_{t-1}= -67,8923+4,3947*$IGFCFc_t$+88,9636*$ICHc_{t-1}$ R^2_{2Mt-1}= 0,42649
(-2,2787) (0,3254) (3,3938)

$ICHc_{t-1}$= 0,8979+0,0047* M_{t-1}+0,0049* $IGFCFc_t$ R^2_{2Mt-1}= 0,42272
(8,2694) (3,3938) (0,0493)

IGFCFc= 0,9843+ 0,0313* $ICHc_{t-1}$+0,0015* M_{t-1} R^2_{2Mt-1}= 0,01379
(1,6889) (0,0493) (0,3254)

N.B. *În paranteze sunt prezentate valorile calculate ale testului Student*.

În aceste condiții, pot fi identificați factorii modelatori ai formei transformate a valorilor calculate ale Testului Student referitor la fiecare dintre cele trei variabile explicative avute în vedere în modelul de regresie (tabelul nr. 4)

Tabelul nr. 4

Factorii modelatori ai valorilor calculate ale Testului Student pentru un model de regresie referitor la evoluția importului României în perioada 1991-2009

Indicator	M_{t-1}	$ICHc_{t-1}$	$IGFCFc_t$
$(TOL)^{(1/2)}$	0.75730	0.75975	0.99308
RV_{3stud}	3.482811	3.482811	3.482811
Cddc	2.571643	2.571643	2.571643
q_k	1	1.003227	1.311336
T_{3k}	0,925638	0.072944	0.761988
TFST	8.290522	0.44715	4.224456

Din datele prezentate în tabelul nr. 4 se pot trage următoarele concluzii:

a) valoarea de referință generală a a formei transformate a valorilor calculate ale testului Student este de 3,482811

b) coeficientul de determinare al regresiei liniare cu trei variabile explicative domină gradul maxim de colinearitate indus în cadrul regresiei de non-ortogonalitatea variabilelor explicative, valoarea coeficientului $Cddc_{max}$ fiind de 2,571643.

c) coeficienții de aliniere la hazardul colinearității sunt pozitivi pentru toate variabilele explicative, dar se diferențiază puternic, mai cu seamă între variabila explicativă primordială şi variabila de rangul al doilea, valorile celor doi indicatori fiind de 0,925638 şi 0,072944. Drept urmare, se poate vorbi de o colinearitate degradantă în cazul variabilei explicative de rangul 2. De asemenea, se mai poate observa că valoarea foarte scăzută a coeficientului de aliniere este factorul esențial care determină probabilitatea ridicată de apariție a ipotezei nule în cazul variabilei explicative menționate anterior.

IV. Concluzii. Sporul de cunoaştere adus de regresia liniară cu trei variabile explicative în interpretarea valorilor calculate ale Testului Student

Examinarea atentă a factorilor modelatori ai formei transformate a valorilor calculate ale testului Student aplicat unei regresii liniare cu trei variabile explicative relevă unele proprietăți şi corelații interesante, care pot fi grupate în două mari categorii şi anume:

a) preluarea unor proprietăți care apar încă de la regresia liniară cu două variabile explicative şi b) apariția unor noi proprietăți care se manifestă dacă numărul de variabile explicative este de cel puțin trei.

În cazul regresiei liniare cu trei variabile explicative noile proprietăți pot fi puse mai uşor în evidență datorită unei complexități relativ mai reduse a calculelor comparativ cu situațiile în care numărul de variabile explicative este de cel puțin patru.

Astfel, sunt preluate o serie de proprietăți care apar încă de la regresia liniară cu două variabile explicative şi care se referă la:

a) valorile calculate ale testului Student sunt influențate de mărimea coeficientului de corelație Pearson dintre variabilele explicative cu valoarea absolută cea mai mare (R_{3max}) în sensul polarizării acestora, prin tendința de a creşte gradul de colinearitate indus de adăugarea unei noi variabile explicative.
b) variabilele explicative pot fi ierarhizate, li se pot atribui ranguri, în funcție de mărimea absolută a coeficientului de corelație Pearson dintre variabilele explicative şi variabila rezultativă.
c) diferențierea puternică a valorilor absolute ale coeficienților de corelație Pearson dintre variabilele explicative şi variabila rezultativă, respectiv valori absolute scăzute ale coeficienților de corelație mediată de variabila rezultativă între variabilele explicative, pot favoriza apariția colinearității toxice. În fapt, colinearitatea toxică apare dacă valoarea absolută a coeficientului de corelație mediată este mai mică decât valoarea absolută coeficientul de corelație Pearson, iar ambele tipuri de coeficienți au acelaşi semn.

Proprietățile noi care apar pentru valorile calculate ale testului Student pentru regresia liniară cu cel puțin trei variabile explicative sunt:

a) a) valorile calculate ale testului Student sunt dependente atât de ritmul cât şi de traiectoria de diferențiere a coeficienților de corelație Pearson dintre variabilele explicative şi variabila rezultativă.
b) b) gradul de colinearitate indus de adăugarea unei noi variabile explicative este dependent de mărimile absolute cele mai ridicate ale coeficienților de corelație Pearson dintre variabilele explicative. Se poate demonstra că, gradul de colinearitate indus este egal cu valoarea coeficientului de determinare obținut în urma regresiei liniare dintre variabila explicativă analizată şi celelalte variabile explicative. În consecință, se poate pune în evidență faptul că

valorile calculate ale testului Student sunt dependente de valoarea factorului de inflamare a dispersiei (VIF). De asemenea, variabilele explicative pot fi definite şi ca generatori de colinearitate şi pot fi ierarhizate în cadrul ecuației de regresie liniară şi din respectivul punct de vedere.

c) c) între rangurile variabilelor explicative şi ordinele de colinearitate poate exista sau nu o concordanță perfectă.

d) d) se pot identifica tipuri de colinearitate între două variabile explicative nu numai în raport cu variabila rezultativă, ci şi în raport cu o terță variabilă explicativă. În cazul colinearității în raport cu o terță variabilă explicativă, colinearitatea non-toxică se înscrie în logica metodei de estimare şi apreciem că reprezintă "varianta ordonată" a regresiei liniare cu trei variabile explicative. Situațiile de colinearitate toxică sau de anticolinearitate în raport cu o terță variablă explicativă apreciem că reprezintă "varianta haotică" a respectivei regresii liniare, în sensul intrării în contradicție cu logica metodei de estimare şi adăugare de către fiecare variabilă explicativă a unui grad de colinearitate în cadrul procesului de estimare a parametrilor.

Dacă se au în vedere toți factorii care modelează valorile calculate ale testului Student, precum şi totalitatea situațiilor posibile, rezultă că în cazul unei regresii liniare cu trei variabile explicative se pot identifica un număr de 54 cazuri. La numărul de cazuri menționat anterior se ajunge prin înmulțirea numărului de traiectorii de diferențiere a coeficienților de corelație dintre variabilele explicative şi variabila rezultativă (3), numărul de tipuri de colinearitate între generatorul principal şi generatorul de colinearitate de ordinul 2 (3) şi numărul de situații posibile dintre rangurile şi ordinele de colinearitate ale variabilelor explicative (6).

Cele mai favorabile premise pentru evitarea apariției colinearității toxice în cadrul ecuației de regresie liniară cu trei variabile explicative sunt:

o diferențiere cât mai mică şi pe o traiectorie subexponențială a coeficienților de corelație Pearson dintre variabilele explicative şi variabila rezultativă.

lipsa unei colinearități toxice dintre două variabile explicative în raport cu variabila rezultativă.

existența unei concordanțe între rangul şi ordinul de colinearitate al variabilelor explicative.

Condițiile prezentate anterior sunt valabile mai ales în contextul "variantei ordonate" a respectivei regresii liniare.

În ceea ce priveşte "varianta haotică", experiențele practice relevă faptul că de multe ori, manifestarea colinearității toxice în raport cu o terță variabilă explicativă provoacă, o contorsionare a relațiilor între variabilele explicative, ceea ce determină o creştere a polarizării mărimii coeficienților de aliniere la hazardul colinearității, mergând până la apariția unor valori negative.

În cazul existenței anticolinearității în raport cu o terță variabilă explicativă, apar o serie de efecte pozitive în sensul relaxării tensiunilor dintre variabilele explicative, care se concretizează prin sporirea valorilor coeficienților de aliniere la hazardul colinearității, diminuarea posibilităților apariției colinearității toxice sau degradante.

Bibliografie

Alin, A., 2010. Multicollinearity. Focus Article. Computational Statistics, 2(3).

Belsley, D., 1976. Multicollinearity: Diagnosing its presence and assessing the potential damage it causes least square estimation. *NBER Working Paper no. 15.*

Belsley, D., 1991. Conditional diagnostics: collinearity and weak data in regression. Wiley Series in Probability John Wiley, New York.

Douglas, D. et. al., 2003. Test about harmful collinearity among predictor variables used in modeling global temperature. *Climate research, 24.*

Conrad, C., 2006. Applied Regression Analysis. Lectures Notes, Department of Economics, Pomona College, Claremont, California, Spring.

Dougherty, C., 2002. Introduction to econometrics-Second Edition. Oxford University Press.

Farrar, D. Glauber, R., 1967. Multicollinearity in regression analysis: The problem revisited. *Review of Economics and Statistics.*

Greene, W., 1993. Econometric Analysis. New York: Mc Millan Publishing Company.

Isaic - Maniu, A. Mitruț, C. and Voineagu, V., 1996. Statistica pentru managementul afacerilor, Bucuresti: Editura Economica.

Jula, D., 2003. Introducere în Econometrie. Ed Professional Consulting, Bucureşti.

Klein, L., 1962. An introduction to econometrics. Prentice Hall.

Pavelescu, F.M., 1986. Some considerations regarding the Cobb-Douglas production function estimated parameters. A new approach. *Revue Roumaine des Sciences Sociales, 3(1-2).*

Pavelescu, F.M., 2004a. Features of Ordinary Least Square (OLS) method. Implications for estimation methodology. *Romanian Journal of Economic Forecast, 5(2).*

Pavelescu, 2004b. Considerații privind regresia linară trifactorială. *Oeconomica nr. 3.*

Pavelescu, F.M., 2004c. O reinterpretare a metodologiei standard de estimare a parametrilor modelului de regresie liniară. *Caiet de Studii nr.4 al Seminarului de Modelare Macroeconomică al INCE.*

Pavelescu, F.M., 2005. Impact of collinearity on the estimated parameters and values of classical Statistical tests in conditions of multifactorial linear OLS regressions. *Romanian Journal of Economic Forecast, 6(2).*

Pavelescu, F.M., 2009. A review of Student Test properties in condition of multifactorial linear regression. *Romanian Journal of Economic Forecast, 10(1).*

Pavelescu, F.M., 2010. An Analysis Model for the Disturbances Generated by Collinearity in the context of the OLS Method, *Romanian Journal of Economic Forecast no.2.*

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, 2(4).*

Pavelescu, F.M., 2012. Interdependența dintre tipul de colinearitate şi valorile calculate ale testului Student în cazul unei rgresii lineare cu două variabile rezultative. *Caietul de Studiu nr. 25 al Seminarului de Modelare Macroeconomică al INCE*

Wiliams, R., 2008. Sociology Graduate Statistics. University of Notre Dame, Indiana.