
-PROIECT - -ECONOMETRIE -

**Analiza econometrica a calitatii vietii
in Europa, in anul 2018**



**Marin Ramona – Alexandra,
Grupa 1061,
Seria F**

Introducere

Tema aleasa este analiza econometrica a calității vieții în Europa, în anul 2018. Motivul alegerii acestei teme este dorința de documentare în legătură cu acest subiect, fiind același subiect pe care l-am ales și pentru lucrarea de licență. Analizând o serie de indicatori economici, putem concluziona aspect referitoare la calitatea vieții în țările din Europa.

Calitatea vieții se referă la bunăstarea oamenilor în societate și indică măsura în care aceștia sunt mulțumiți de condițiile de trai din țara în care trăiesc. Calitatea vieții este un concept multidimensional care include o serie de domenii: condiții materiale de viață, sănătate, munca, locuirea, viața de familie, siguranța. Deși, abordarea calității vieții este raportată la nivel individual, aceasta surprinde și relația pe care o are individul cu societatea în care trăiește. Astfel, țările dezvoltate sunt caracterizate de niveluri înalte ale satisfacției față de viață, populația fiind mulțumită de modul în care trăiește. Analiza econometrică va fi efectuată prin corelația care există între produsul intern brut, venitul național brut și consumul final, pe un cap de locuitor.

Literatura de specialitate

Fundatia Europeana pentru Imbunatatirea Conditiei de Viata si de Munca a lansat în anul 2003, Studiul referitor la calitatea vieții în Europa (European Quality of Life Survey-EQLS), în cele 28 de state -UE25 și cele trei state candidate (Bulgaria, România, Turcia) care cuprinde mai multe analize realizate cu ajutorul unor indicatori economici, calculați în urma datelor obținute de la cetățenii țărilor analizate prin intermediul chestionarelor, interviurilor și sondajelor statistice. Aceeași fundație realizează studii anuale în ceea ce privește calitatea vieții în țările din Europa și are ca scop descrierea și explicarea calității vieții pe toate dimensiunile sale.

Studii teoretice care analizează problema măsurării și comparării calității vieții în Europa au fost realizate de **Nussbaum și Sen** (1993), **Diener** (2000), **Diener și Suh** (1997) și **Robeyns** (2005).

De asemenea, **Frisch**, în anul 1998) a calculat scorul satisfacerii vieții ca o sumă de ranguri bazate pe satisfacerea în diferite domenii ale vieții. Acesta a susținut necesitatea evaluării a calității vieții în mai multe domenii: medicina generală, psihiatrie, psihologie, medicina comportamentală.

Modelul econometric

Analiza calitatii vietii populatiilor din Europa este efectuată utilizând modele econometrice, în special a regresiei liniare simple și multiple.

Ca variabile ale modelului econometric au fost alese:

- variabila dependentă (explicată) → PIB/locuitor, milioane euro (Y).
- variabile independente (explicative):
 - venitul national brut/ locuitor, milioane euro (X_1);
 - consumul final/locuitor, milioane euro (X_2).

Modelul econometric este:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

Sau $y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \varepsilon_i$,

oricare ar fi i - unitate statistica, cu valori de la 1 la n , unde n reprezinta volumul esantionului.

Pentru acest model econometric, consideram o selectie de volum $n=35$ tari din Europa pentru care s-au inregistrat valorile variabilelor alese.

Nr_crt	Tara	PIB/per capita (milioane euro) _ Y	venit national brut/ per capita (milioane euro) _ X1	Consum final/ per capita (milioane dolari US)	Consum final/ per capita (milioane euro)_ X2
1	Belgia	133.3	36,811	24258.49	20377.14
2	Bulgaria	26.4	15,361	5644.92	4741.74
3	Republica Ceha	65.7	26,864	11378.04	9557.55
4	Danemarca	172.6	41,005	29712.18	24958.23
5	Germania	133.9	39,049	25476.18	21399.99
6	Estonia	65.0	24,708	10614.01	8915.77
7	Irlanda	222.5	46,331	25207.30	21174.13
8	Grecia	55.4	20,301	16023.31	13459.58
9	Spania	85.2	28,279	18203.81	15291.20
10	Franta	116.1	33,028	23455.20	19702.37
11	Croatia	42.0	19,639	9191.26	7720.66
12	Italia	96.9	30,147	21339.44	17925.13
13	Cipru	81.5	26,945	20254.39	17013.69
14	Letonia	50.0	20,979	9885.22	8303.58
15	Lituania	53.7	24,378	11480.76	9643.84
16	Luxembourg	326.3	51,360	34721.42	29165.99
17	Ungaria	46.0	21,236	8618.79	7239.78
18	Malta	85.8	28,220		0.00
19	Olanda	148.6	40,534	23626.01	19845.85
20	Austria	144.2	39,142	25457.98	21384.70
21	Polonia	42.9	20,976	9729.70	8172.95
22	Portugalia	66.0	23,588	15451.77	12979.49
23	Romania	34.7	19,856	8037.47	6751.48
24	Slovenia	73.2	26,530	14120.92	11861.57
25	Slovacia	54.4	21,546	10894.14	9151.08
26	Finlanda	140.1	34,539	26296.62	22089.16
27	Suedia	153.0	37,725	27315.98	22945.42
28	Islanda	208.3	57,028	27577.20	23164.84
29	Norvegia	230.6	72777.6	40750.27	34230.23
30	Elvetia	242.0	71173.2	41728.02	35051.53
31	Regatul Unit	120.5	32220	28296.66	23769.19
32	Muntenegru	24.8	7081.2	6435.05	5405.44
33	Albania	14.8	4082.4	3793.91	3186.88
34	Serbia	20.3	5384.4	4722.78	3967.13
35	Turcia	26.8	8836.8	8835.95	7422.20

Descrierea variabilelor

- **Produsul intern brut (PIB)** este un indicator macroeconomic care reflectă suma valorii de piață a tuturor mărfurilor și serviciilor destinate consumului final, produse în toate ramurile economiei în interiorul unei țări în decurs de un an.

Formula de calcul:

$$\text{PIB} = \text{Consum} + \text{Investitii} + \text{Cheltuieli Guvernamentale} + \text{Exporturi} - \text{Importuri}$$

Sursa date:

[https://ec.europa.eu/eurostat/databrowser/view/NAMA_10_PC\\$DEFAULTVIEW/default/table?!ang=en](https://ec.europa.eu/eurostat/databrowser/view/NAMA_10_PC$DEFAULTVIEW/default/table?!ang=en)

- **Venitul national brut (VNB)** reprezintă ansamblul veniturilor primare de primit de către unitățile instituționale rezidente: remunerarea angajaților, impozitele pe producție și importuri minus subvențiile, veniturile din proprietate (cele de primit minus cele de plătit), excedentul brut de exploatare și venitul mixt brut.

Formula de calcul:

$$\text{VNB} = \text{PIB} - \text{venituri primare de platit pentru unitatile institutionale} + \text{venituri primare primite de la restul lumii}$$

Sursa date:

https://ec.europa.eu/eurostat/databrowser/view/NAMA_10_PP/default/table?lang=en

- **Consumul final (CF)** este suma consumului privat, valoarea bunurilor și serviciilor achiziționate de sectorul privat strict pentru consum și consumul public, valoarea bunurilor și serviciilor produse și achiziționate de instituțiile sectorului public pentru a își desfășura activitățile.

Formula de calcul: $\text{CF} = C_{pv} + C_{pb}$

Sursa date: <https://data.worldbank.org/indicator/NE.CON.PRVT.PC.KD>

Partea I

MODELUL DE REGRSIE LINIARA UNIFACTORIALA (SIMPLA)

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon \text{ sau } y_i = \beta_0 + \beta_1 x_{i1} + \varepsilon_i, \text{ pentru unitatea statistica } i$$

- variabila dependenta (explicata) este PIB/locuitor, milioane euro (Y);
- variabile independenta (explicativa) este venitul national brut/ locuitor, milioane euro (X_1);

Output-ul estimarii modelului de regresie liniara unifactoriala in Excel:

SUMMARY OUTPUT

Regression Statistics					
Multiple R					0.9126
R Square					0.8329
Adjusted R Square					0.8278
Standard Error					30.9726
Observations					35

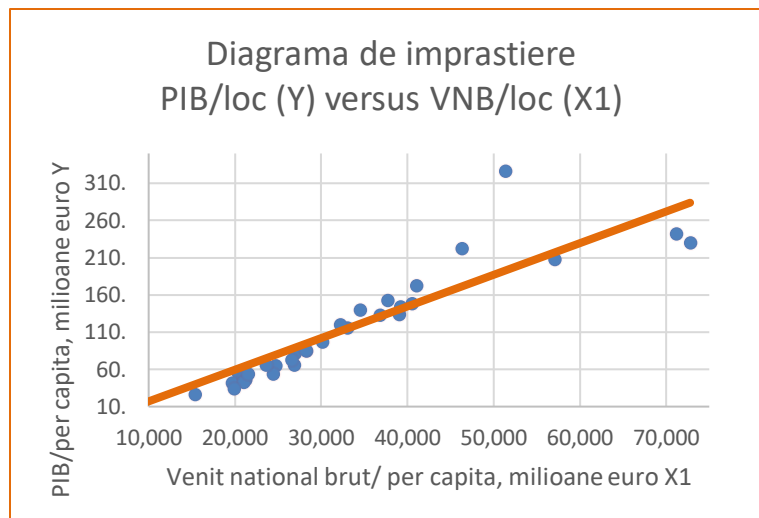
ANOVA					
	df	SS	MS	F	Significance F
Regression	1	157751.4552	157751.4552	164.4441	2.27668E-14
Residual	33	31656.9505	959.3015		
Total	34	189408.4057			

	Coefficients	Standard Error	t Stat	P-value
Intercept	-25.4717	11.3009	-2.2540	0.0310
venit national brut/ per capita (milioane euro)_X1	0.0042	0.0003	12.8236	2.27668E-14

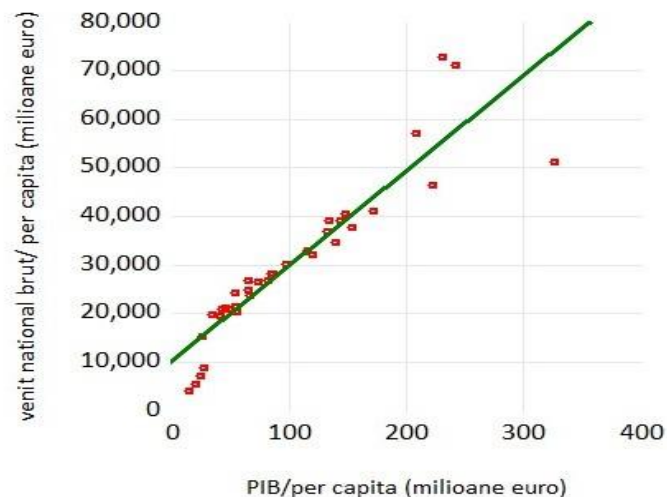
Interpretare SUMMARY OUTPUT:

- **R=0.9126** Valoarea coeficientului de corelatie este foarte apropiata de 1, ceea ce arata ca intre PIB/ loc si venitul national brut/ loc exista o legatura foarte puternica, legatura pusa in evidenta prin modelul de regresie.
- **R² = 0.8329** (coeficientul de determinatie) La nivelul esantionului de 35 de tari, arata ca 83,29% din variatia totala a variabilei dependente (PIB-ului/ loc) este explicata de variatia totala a variabilei independente(VNB/ loc), iar restul de 16,71% din variatia totala este explicata de anumiți factori reziduali.

Observation	Predicted PIB/per capita (milioane euro)_Y	Residuals	Standard Residuals
1	130.9734	2.3266	0.0762
2	39.8119	-13.4119	-0.4395
3	88.6991	-22.9991	-0.7537
4	148.7978	23.8022	0.7801
5	140.4848	-6.5848	-0.2158
6	79.5362	-14.5362	-0.4764
7	171.4330	51.0670	1.6736
8	60.8067	-5.4067	-0.1772
9	94.7128	-9.5128	-0.3118
10	114.8959	1.2041	0.0395
11	57.9932	-15.9932	-0.5241
12	102.6517	-5.7517	-0.1885
13	89.0434	-7.5434	-0.2472
14	63.6881	-13.6881	-0.4486
15	78.1337	-24.4337	-0.8007
16	192.8061	133.4939	4.3749
17	64.7804	-18.7804	-0.6155
18	94.4621	-8.6621	-0.2839
19	146.7960	1.8040	0.0591
20	140.8801	3.3199	0.1088
21	63.6754	-20.7754	-0.6809
22	74.7763	-8.7763	-0.2876
23	58.9154	-24.2154	-0.7936
24	87.2796	-14.0796	-0.4614
25	66.0978	-11.6978	-0.3834
26	121.3175	18.7825	0.6155
27	134.8579	18.1421	0.5946
28	216.8931	-8.5931	-0.2816
29	283.8299	-53.2299	-1.7445
30	277.0113	-35.0113	-1.1474
31	111.4619	9.0381	0.2962
32	4.6231	20.1769	0.6612
33	-8.1217	22.9217	0.7512
34	-2.5883	22.8883	0.7501
35	12.0843	14.7157	0.4823



Output-ul estimarii modelului de regresie liniar unifactorial in Eviews:



Dependent Variable: PIB_PER_CAPITA__MILIOANE_EURO__

Method: Least Squares

Date: 04/13/21 Time: 17:42

Sample: 1 35

Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-25.47173	11.30089	-2.253958	0.0310
VENIT_NATIONAL_BRUT__PER_CAPITA...	0.004250	0.000331	12.82358	0.0000
R-squared	0.832864	Mean dependent var		102.9571
Adjusted R-squared	0.827799	S.D. dependent var		74.63803
S.E. of regression	30.97259	Akaike info criterion		9.759528
Sum squared resid	31656.95	Schwarz criterion		9.848405
Log likelihood	-168.7917	Hannan-Quinn criter.		9.790208
F-statistic	164.4441	Durbin-Watson stat		2.123386
Prob(F-statistic)	0.000000			

Estimatiilor parametrilor modelului

Pe baza esantionului de volum $n=35$ tari, am determinat cu ajutorul softurilor (**Eviews**, **Excel**) estimatorii $\hat{\beta}_0$ si $\hat{\beta}_1$ ai parametrilor β_0 - intercept si β_1 – panta ai modelului de regresie liniara unifactoriala.

- Astfel, se obtine: $\hat{\beta}_0 = 25.4717$ si $\hat{\beta}_1 = 0.0042$.
- Dreapta de regresie de selectie(la nivelul esantionului) este de ecuatia $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$,

adica: $\hat{y} = 25.4717 + 0.0042 x$.

- Ecuatia de regresie liniara in esantion este $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$,

adica : $\hat{y} = 25.4717 + 0.0042 x_i$,

unde \hat{y}_i -reprezinta valorile ajustate ale variabilei Y, pentru unitatile statistice i, i ia valori intre 1 si 35,

iar $\hat{\varepsilon}_i = y_i - \hat{y}_i$ sunt reziduurile modelului estimat de regresie(valorile observate ale variabilei eroare ε) .

Interpretarea valorilor coeficientilor:

- ✚ $\hat{\beta}_0$ – intercept, arata nivelul mediu al variabilei dependente la nivelul esantionului daca nivelul variabilei independente ar fi egal cu 0 unitati. Astfel, daca venitul national brut/loc ar fi egal cu 0 milioane de euro, atunci PIB-ul/loc ar fi egal cu 25.4717 milioane de euro.
- ✚ $\hat{\beta}_1$ – panta dreptei de regresie, arata cu cate unitati se modifica, in medie, la nivelul esantionului, nivelul variabilei dependente, daca nivelul variabilei independente creste cu o unitate. Astfel, PIB-ul/loc creste, in medie, cu 0.0042 mil euro daca venitul national brut/loc creste cu o unitate, adica cu 1 mil euro.

Testarea validitatii modelului de regresie

Testarea validitatii modelului de regresie liniara simpla, la un nivel de semnificatie de 5%.

- $Y = \beta_0 + \beta_1 X + \varepsilon$ sau $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, oricare ar fi i- unitate statistica.

Ipotezele

- **H₀**: modelul de regresie nu este valid statistic(variabila Y nu este un predictor al variabilei X₁);
- **H₁**: modelul de regresie este valid statistic.

Se considera selectia de volum n=35 tari din populatia statistica si se presupune ca cele sapte ipoteze fundamentale ale modelului sunt verificate si adevarate.

Testul statistic este $F = \frac{MSE}{MSR} = \frac{\frac{SSE}{k}}{\frac{SSR}{n-k-1}} \sim \text{Fisher}_{(k,n-k-1)} \Rightarrow F_{\text{calc}} = 164.44;$

- k= numarul variabilelor explicative, in cazul nostrum, k=1.

Nivelul de semnificatie al testului este $\alpha = 0.05$ sau 5%, iar nivelul de incredere al testului este $1 - \alpha = 0.95$ sau 95% .

Regiunea critica R_c (regiunea de respingere a ipotezei nule H₀) este definita: $F > F_{\text{critic}}$,

unde $F_{\text{critic}} = F_{\alpha; k; n-k-1}$, in cazul nostru $F_{\text{critic}} = F_{0.05; 1; 33} = \text{F.INV.RT}(0.05, 1, 33)$ (in excel)= 4.14;

Decizia

- + Deoarece $F_{\text{calc}} (164.44) > F_{\text{critic}} (4.14)$, inseamna ca F_{calc} apartine regiunii critice R_c, atunci respingem ipoteza nula H₀ la un nivel de semnificatie de 5%, concluzionand ca datele de selectie sunt in favoarea ipotezei alternative H₁, adica modelul este valid statistic.

Significance F

- + De asemenea, putem compara valoarea “Significance F” cu valoarea lui α pentru a afla daca modelul este valid statistic. Astfel, in cazul nostru, Significance F= 2.27E-14= $2.27 \cdot 10^{-14}$, Significance F < $\alpha(0,05)$, rezulta ca respingem ipoteza nula si acceptam ipoteza alternativa. Asadar, modelul de regresie este valid statistic.

Testarea semnificatiei statistice a parametrului panta $\hat{\beta}_1$

Testarea semnificatiei statistice a parametrului panta $\hat{\beta}_1$ ai modelului de regresie liniara simpla, la un nivel de semnificatie de 5%.

- $Y = \beta_0 + \beta_1 X + \varepsilon$ sau $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, oricare ar fi i- unitate statistica.

Ipotezele:

- $H_0: \beta_1 = 0$ (parametru β_1 nu este semnificativ diferit de 0);
- $H_1: \beta_1 \neq 0$ (parametru β_1 este semnificativ diferit de 0);

Se considera selectia de volum $n=35$ tari din populatia statistica si se presupune ca cele sapte ipoteze fundamentale ale modelului sunt verificate si adevarate.

Testul statistic: $t = \frac{\hat{\beta}_1 - \beta_1}{se(\hat{\beta}_1)} \sim \text{Student}_{(n-k-1)}$, sub presupunerea ca H_0 este adevarata;

❖ **$t_{\text{calc}} = 12.82$;**

Nivelul de semnificatie al testului este $\alpha = 0.05$ sau 5%, iar nivelul de incredere al testului este $1 - \alpha = 0.95$ sau 95%.

Regiunea critica R_c (de respingere a ipotezei nule H_0) este definite: $t < -t_{\text{critic}}$ sau $t > t_{\text{critic}}$, unde $t_{\text{critic}} = t_{\alpha/2; n-k-1}$, reprezinta cuantila superioara de ordinul $\alpha/2$ a distributiei Student $_{(n-k-1)}$.

Valoarea lui t_{critic} este returnata in Excel cu ajutorul functiei T.INV.2T($\alpha, n-k-1$).

In cazul nostru, $t_{\text{critic}} = \text{T.INV.2T}(0.05, 33) = 2.034$.

Decizia

- ✚ Cum $t_{\text{calc}} (12.82) > t_{\text{critic}} (2.034)$, t_{calc} apartine regiunii critice R_c , deci respingem ipoteza nula H_0 la un nivel de semnificatie de 5%, concluzionand ca datele din esantion sunt in favoarea ipotezei alternative H_1 , adica parametrul β_1 este semnificativ diferit de 0.

Intervalul de incredere de 95% pentru parametrul β_1

Intervalul de incredere de 95% pentru parametrul β_1 al modelului de regresie liniara unifactoriala

$$Y = \beta_0 + \beta_1 X + \varepsilon ,$$

sau $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, oricare ar fi i- unitate statistica.

$$CI_{100 \cdot (1-\alpha)\%}(\beta_1) = [\hat{\beta}_1 - t_{\frac{\alpha}{2}; n-k-1} \cdot se(\hat{\beta}_1) ; \hat{\beta}_1 + t_{\frac{\alpha}{2}; n-k-1} \cdot se(\hat{\beta}_1)]$$

In cazul nostru, $\alpha = 0.05$; $n-k-1 = 33$; $\hat{\beta}_1 = 0.0042$; $se(\hat{\beta}_1) = 0.0003$; $t_{\alpha/2; n-k-1} = 2.034$.

- $CI_{95\%}(\beta) = [0.0042 - 2.034 \cdot 0.0003; 0.0042 + 2.034 \cdot 0.0003]$;
- $CI_{95\%}(\beta) = [0.0035; 0.0048] \text{ mil euro}$

✚ Cum intervalul de incredere de 95% determinat pentru parametrul β_1 nu contine valoarea 0, atunci putem spune ca acest parametru este semnificativ diferit de 0.

✚ Intervalul $[0.0035; 0.0048] \text{ mil euro}$ acopera, pentru un nivel de incredere de 95%, valoarea adevarata a parametrului β_1 . Adica, daca venitul national brut/ loc creste cu o unitate(1 mil euro) , atunci PIB-ul/ loc creste, in medie, cu o valoarea acoperita de intervalul $[0.0035; 0.0048] \text{ mil euro}$, la un nivel de incredere de 95%.

Verificare ipotezelor fundamentale ale modelului clasic de regresie

1) Homoscedasticitatea erorilor aleatoare

-prin metoda grafica

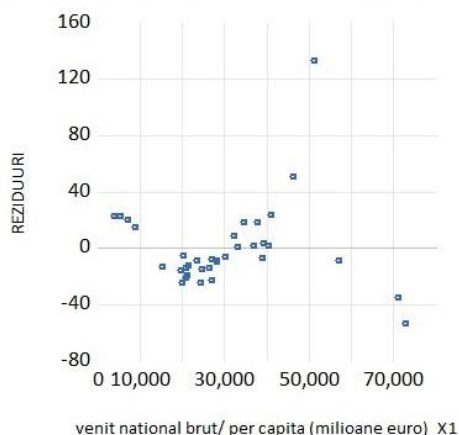


Figura 1. Reziduri vs VNB/ per capita

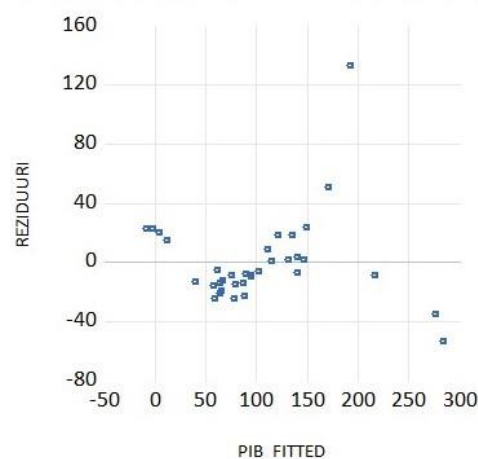


Figura 2. Reziduri vs valorile estimate prim modelul de regresie ale PIB-ului/per capita

✚ Din cele două grafice, Figurile 1 și 2, se observă că valoarea absolută a reziduurilor nu se modifica semnificativ, aproape deloc, pe măsură ce valorile venitului national brut/ per capital cresc, respectiv pe măsură ce valorile ajustate ale PIB-ului/per capita cresc, ceea ce sugerează că ipoteza de homoscedasticitate este îndeplinită, nu există heteroscedasticitate

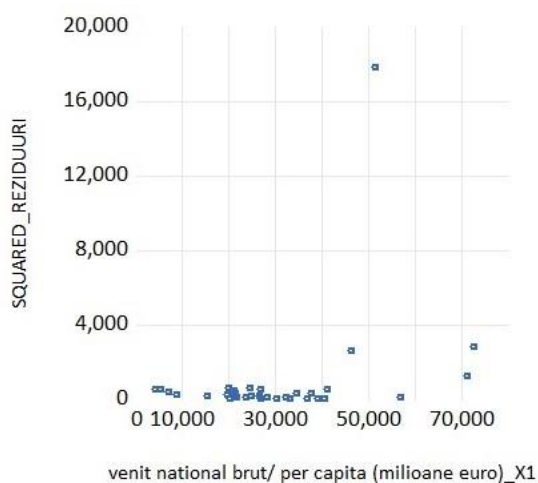


Figura 3. Pătratele reziduurilor versus VNB/per capita

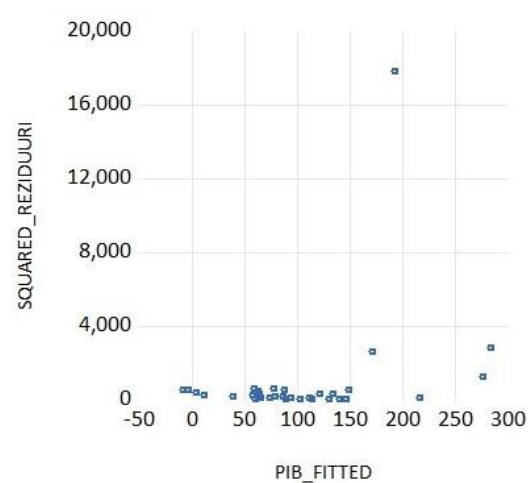


Figura 4. Pătratele reziduurilor versus valorile estimate prin modelul de regresie ale PIB-ului/per capita

✚ Din cele două grafice, Figurile 3 și 4, se observă că valorile pătratelor reziduurilor nu cresc pe măsură ce valorile venitului national brut/ per capita cresc, respectiv pe măsură ce valorile ajustate ale PIB-ului/per capita cresc, ceea ce sugerează că ipoteza de homoscedasticitate a erorilor este îndeplinită, nu există heteroscedasticitate.

-prin teste statistice

Testul White

Testul White solicită ca, după determinarea reziduurilor din regresia originală, să se estimeze o regresie auxiliară a pătratelor reziduurilor în raport cu o constantă, variabilele explicative ale modelului original, pătratele lor și produsele lor încrucișate.

În cazul nostru, modelul de regresie inițial (original) este un model de regresie liniară unifactorială:

$$PIB_i = \beta_0 + \beta_1 \times VNB_i + \varepsilon_i, \text{ unde } Y = \text{PIB/per capita si } X = \text{VNB/per capita},$$

Pentru care am retinut seria reziduurilor $\hat{\varepsilon}_i$, atunci regresia auxiliara este:

- $\hat{\varepsilon}_i^2 = \alpha_0 + \alpha_1 \cdot x_i + \alpha_2 \cdot x_i^2 + \eta_i$, unde x_i este VNB/per capita si η_i -este o variabila de perturbatie (eroare) care verifica ipotezele asociate modelului clasic de regresie liniara.

Ipotezele testului sunt:

- $H_0: \alpha_1 = \alpha_2 = 0$ (nu exista heteroscedasticitate sau erorile aleatoare sunt homoscedastice);
- $H_1: \text{exista } \alpha_i \neq 0$ (exista heteroscedasticitate sau erorile aleatoare sunt heteroscedastice);

Testul statistic este: $W = nR_a^2$, urmeaza asimptotic o distributie χ^2 cu gradele de libertate date de numarul de regresori din ecuatiia auxiliara (in cazul nostru sunt 2 regresori);

Output-ul din Eviews, testul White:

Heteroskedasticity Test: White				
Null hypothesis: Homoskedasticity				
F-statistic	1.944249	Prob. F(2,32)	0.1596	
Obs*R-squared	3.792230	Prob. Chi-Square(2)	0.1502	
Scaled explained SS	18.14080	Prob. Chi-Square(2)	0.0001	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 04/27/21 Time: 16:58				
Sample: 1 35				
Included observations: 35				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-427.3379	1810.024	-0.236095	0.8149
VENIT_NATIONAL_BRUT__PER_CAPITA...	4.74E-07	1.36E-06	0.347396	0.7306
VENIT_NATIONAL_BRUT__PER_CAPITA...	0.025837	0.105585	0.244705	0.8082
R-squared	0.108349	Mean dependent var	904.4843	
Adjusted R-squared	0.052621	S.D. dependent var	3010.548	
S.E. of regression	2930.268	Akaike info criterion	18.88539	
Sum squared resid	2.75E+08	Schwarz criterion	19.01871	
Log likelihood	-327.4944	Hannan-Quinn criter.	18.93141	
F-statistic	1.944249	Durbin-Watson stat	1.860995	
Prob(F-statistic)	0.159630			

Dependent Variable: ABSOLUTE_REZIDUURI

Method: Least Squares

Date: 05/09/21 Time: 12:41

Sample: 1 35

Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.033852	7.990982	0.504801	0.6171
VENIT_NATIONAL_BRUT__PER_CAPITA...	0.000516	0.000234	2.203573	0.0346
R-squared	0.128269	Mean dependent var	19.63903	
Adjusted R-squared	0.101853	S.D. dependent var	23.10955	
S.E. of regression	21.90106	Akaike info criterion	9.066393	
Sum squared resid	15828.67	Schwarz criterion	9.155270	
Log likelihood	-156.6619	Hannan-Quinn criter.	9.097073	
F-statistic	4.855735	Durbin-Watson stat	1.713717	
Prob(F-statistic)	0.034642			

Ecuatia de regresie auxiliara estimata este:

$$\hat{\varepsilon}_i^2 = -427.3379 + \cdot x_i + 4.74 \cdot 10^{-7} \cdot x_i^2;$$

$$W_{\text{calc}} = n \cdot R_a^2 = \text{obs} * R - \text{squared} = 3.79, \text{ cu probabilitatea asociata Prob. Chi-Square}(2) = \alpha = 0.1502.$$

$$\chi_{\text{critic}}^2 = \chi_{\alpha, df}^2 = \chi_{0.05, 2}^2 = 5.99, \text{ returnata de functia } = \text{CHISQ.INV.RT}(0.05, 2) \text{ in Excel};$$

Decizia

Cum $W_{\text{calc}} < \chi_{\text{critic}}^2$ si $\text{Prob Chi-Square}(2) = 0.1502 > 0.05 (\alpha)$, rezulta ca nu avem motive sa respingem ipoteza nula H_0 , deci erorile aleatoare sunt homoscedastice, nu exista heteroscedasticitate.

Testul Glejser

Dintre formele functionale liniare in parametrii propuse de Glejser, incepem cu estimarea modelului:

$$\hat{\varepsilon}_i = \lambda_0 + \lambda_1 x_i + v_i,$$

Ipotezele testului sunt:

- $H_0: \lambda_1 = \lambda_2 = 0$ (nu exista heteroscedasticitate, adica exista homoscedasticitate);
- H_1 : exista $\lambda_1 \neq 0$ (exista heteroscedasticitate);
- **Ecuatia estimata de regresie** este: $|\hat{\varepsilon}_i| = 4.033852 + 0.000516 \text{ VNB/per cap}_i$
- $\text{se} = (7.990982) (0.000234)$
- $R^2 = 0.128269$

Coeficientul panta λ_1 este semnificativ statistic deoarece $t\text{-Statistic} = 2.203$ si $\text{Prob-value} = 0.0346 (< 0.05)$, deci erorile aleatoare ale modelului original de regresie sunt heteroscedastice.

Output-ul din Eviews, testul Glejer:

Heteroskedasticity Test: Glejser
Null hypothesis: Homoskedasticity

F-statistic	4.855735	Prob. F(1,33)	0.0346
Obs*R-squared	4.489431	Prob. Chi-Square(1)	0.0341
Scaled explained SS	6.681422	Prob. Chi-Square(1)	0.0097

Test Equation:
Dependent Variable: ARESID
Method: Least Squares
Date: 05/09/21 Time: 12:54
Sample: 1 35
Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.033852	7.990982	0.504801	0.6171
VENIT_NATIONAL_BRUT__PER_CAPITA...	0.000516	0.000234	2.203573	0.0346
R-squared	0.128269	Mean dependent var	19.63903	
Adjusted R-squared	0.101853	S.D. dependent var	23.10955	
S.E. of regression	21.90106	Akaike info criterion	9.068393	
Sum squared resid	15828.67	Schwarz criterion	9.155270	
Log likelihood	-156.6619	Hannan-Quinn criter.	9.097073	
F-statistic	4.855735	Durbin-Watson stat	1.713717	
Prob(F-statistic)	0.034642			

Rezultatul aplicarii testului Glejer este ca exista heteroscedasticitate, desi testul White a demonstrat contrariul, vom considera ca exista o urma de heteroscedasticitate care trebuie sa fie corectata.

Non-autocorelarea erorilor aleatoare

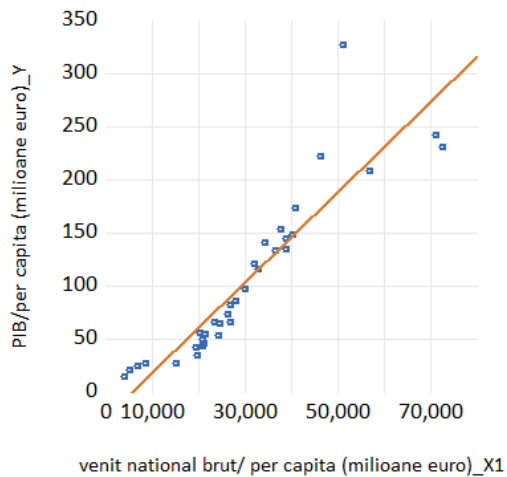


Figura 1. Diagrama de imprastiere: PIB/per cap versus Venit national brut/per cap

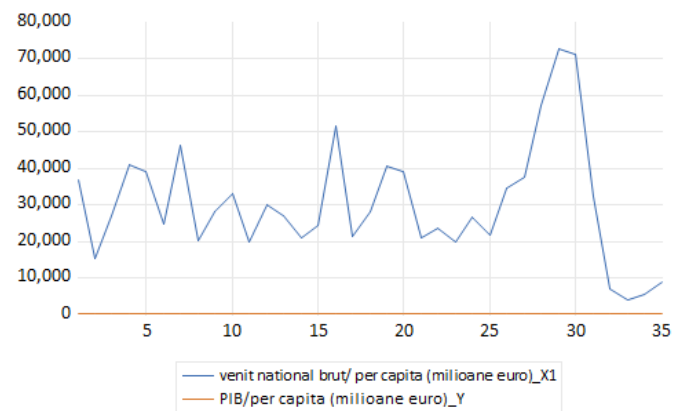


Figura 2. Evolutia celor doua serii de timp, PIB-ul si VNB, in perioada analizata

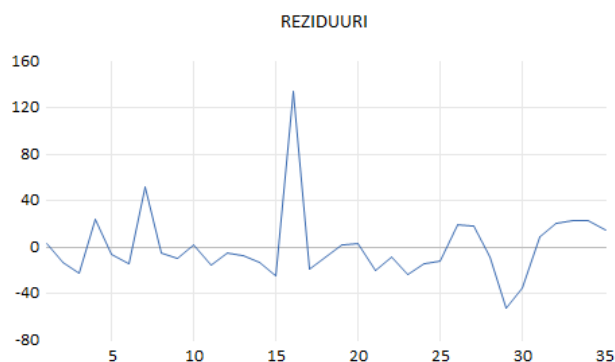


Figura 3. Evolutia seriei reziduurilor $\hat{\varepsilon}_t$.

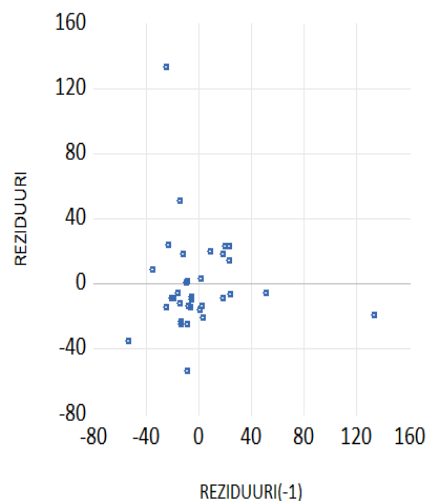


Figura 4. Diagrama de imprastiere a seriei reziduurilor curente $\hat{\varepsilon}_t$ versus reziduurile decalate $\hat{\varepsilon}_{t-1}$

Deoarece valorile reziduurilor sunt distribuite in mod aleator de o parte si de alta a axei orizontale, fara sa apara un model anume, erorile aleatoare nu sunt autocorelate. (figura 3)

Diagrama de imprastiere sugereaza sa consideram modelul de regresie liniara simpla:

- $y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$

Outuput-ul de estimare pentru modelul de regresie:

- $\hat{y}_t = -25.47173 + 0.004250x_t$

Dependent Variable: PIB_PER_CAPITA__MILIOANE_EURO__Y
Method: Least Squares
Date: 05/15/21 Time: 13:56
Sample: 1 35
Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-25.47173	11.30089	-2.253958	0.0310
VENIT_NATIONAL_BRUT__PER_CAPITA...	0.004250	0.000331	12.82358	0.0000
R-squared	0.832864	Mean dependent var		102.9571
Adjusted R-squared	0.827799	S.D. dependent var		74.63803
S.E. of regression	30.97259	Akaike info criterion		9.759528
Sum squared resid	31656.95	Schwarz criterion		9.848405
Log likelihood	-168.7917	Hannan-Quinn criter.		9.790208
F-statistic	164.4441	Durbin-Watson stat		2.123386
Prob(F-statistic)	0.000000			

Testul Durbin- Watson

-verifica daca exista autocorelare de ordinul intai in seria reziduurilor.

-se estimeaza parametrii modelului de regresie prin MCMMP si se obtin reziduurile $\hat{\varepsilon}_t$.

Ipotezele testului sunt:

- $H_0: \rho = 0$ (nu exista autocorelarea erorilor);
- $H_1: \rho \neq 0$ (exista autocorelarea erorilor);

Statistica testului DW:

$$DW = \frac{\sum_{t=2}^n (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^n \hat{\varepsilon}_t^2}.$$

$$\hat{\rho} = \frac{\sum_{t=2}^n \hat{\varepsilon}_t \cdot \hat{\varepsilon}_{t-1}}{\sum_{t=1}^n \hat{\varepsilon}_t^2}$$

Si , unde p este **coeficientul de selectie**

Decizie:

🚦 Cum $DW \approx 2$ si $p \approx 0$, rezulta ca nu exista autocorelatie.

Observation	Residuals ($\hat{\epsilon}_t$)	$\hat{\epsilon}_t - \hat{\epsilon}_{t-1}$	$(\hat{\epsilon}_t - \hat{\epsilon}_{t-1})^2$	$(\hat{\epsilon}_t)^2$	$\hat{\epsilon}_t * \hat{\epsilon}_{t-1}$
1	2.3266			5.41	
2	-13.4119	15.7384	247.6979	179.88	-31.20
3	-22.9991	9.5873	91.9155	528.96	308.46
4	23.8022	-46.8014	2190.3671	566.55	-547.43
5	-6.5848	30.3871	923.3749	43.36	-156.73
6	-14.5362	7.9514	63.2242	211.30	95.72
7	51.0670	-65.6032	4303.7771	2607.84	-742.32
8	-5.4067	56.4736	3189.2696	29.23	-276.10
9	-9.5128	4.1062	16.8605	90.49	51.43
10	1.2041	-10.7170	114.8531	1.45	-11.45
11	-15.9932	17.1973	295.7480	255.78	-19.26
12	-5.7517	-10.2415	104.8873	33.08	91.99
13	-7.5434	1.7916	3.2100	56.90	43.39
14	-13.6881	6.1448	37.7580	187.36	103.25
15	-24.4337	10.7456	115.4680	597.01	334.45
16	133.4939	-157.9277	24941.1456	17820.63	-3261.75
17	-18.7804	152.2743	23187.4608	352.70	-2507.06
18	-8.6621	-10.1183	102.3800	75.03	162.68
19	1.8040	-10.4660	109.5378	3.25	-15.63
20	3.3199	-1.5159	2.2981	11.02	5.99
21	-20.7754	24.0953	580.5825	431.62	-68.97
22	-8.7763	-11.9991	143.9787	77.02	182.33
23	-24.2154	15.4392	238.3677	586.39	212.52
24	-14.0796	-10.1358	102.7342	198.24	340.94
25	-11.6978	-2.3818	5.6729	136.84	164.70
26	18.7825	-30.4803	929.0492	352.78	-219.71
27	18.1421	0.6404	0.4101	329.14	340.75
28	-8.5931	26.7352	714.7720	73.84	-155.90
29	-53.2299	44.6368	1992.4460	2833.43	457.41
30	-35.0113	-18.2186	331.9185	1225.79	1863.65
31	9.0381	-44.0494	1940.3519	81.69	-316.44
32	20.1769	-11.1388	124.0733	407.11	182.36
33	22.9217	-2.7448	7.5338	525.40	462.49
34	22.8883	0.0334	0.0011	523.87	524.64
35	14.7157	8.1726	66.7906	216.55	336.82

SUM($(\hat{\epsilon}_t - \hat{\epsilon}_{t-1})^2$)	67219.9160
SUM($(\hat{\epsilon}_t)^2$)	31656.95
DW	2.123385699

SUM($(\hat{\epsilon}_t * \hat{\epsilon}_{t-1})$)	-2063.99
SUM($(\hat{\epsilon}_t)^2$)	31656.95
p	-0.06

SAU

-se determina, pentru nivelul de semnificatie de 1%, valorile critice $d_1(dL)$ si $d_2(dU)$;

-din tabelul valorilor critice ale statisticii Durbin-Watson, pentru un nivel de semnificatie de 1%, $n=35$ (volumul esantionului) si $k=1$ (o singura variabila explicativa), se gasesc $d_1(dL)=1,195$ si $d_2(dU)=1,307$.

Regiunea 1	Regiunea 2	Regiunea 3	Regiunea 4	Regiunea 5
0	d_1	d_2	$4 - d_2$	$4 - d_1$
4				
Respingem H_0 și acceptăm H_1 , $\rho > 0$.	Indecizie; se recomandă aplicarea altui test.	Nu respingem H_0 , erorile sunt necorelate.	Indecizie; se recomandă aplicarea altui test.	Respingem H_0 și acceptăm H_1 , $\rho < 0$.

📌 Regiunea in care se incadreaza $DW(\text{calc})=2,12$ este **Regiunea 3**, deoarece

$1,307 < DW < 4 - 1,307$, rezulta ca nu respingem ipoteza nula H_0 , deci nu exista autocorelare de ordinul 1, erorile sunt necorelate.

Testul Breusch-Godfrey

-acest test se foloseste pentru detectarea autocorelarii de ordin superior lui 1, in seria reziduurilor.

Ipotezele testului sunt:

- $H_0: \rho_1 = \rho_2 = \dots = \rho_r$ (erorile nu sunt autocorelate, nu exista autocorelare de ordin r a erorilor aleatoare)
- $H_1: \rho_i \neq 0, i \in \{1, 2, \dots, r\}$ (exista autocorelare de ordin r a erorilor aleatoare)

-vom aplica testul Breusch-Godfrey pentru autocorelarea de ordin $r = 2$.

-se estimeaza parametrii modelului de regresie initial prin MCMMP si se obtin reziduurile $\hat{\varepsilon}_t$.

-se considera un model de regresie auxiliar:

$\hat{\varepsilon}_t = \alpha_0 + \alpha_1 x_{t,1} + \alpha_2 x_{t,2} + \dots + \alpha_k x_{t,k} + \rho_1 \hat{\varepsilon}_{t-1} + \rho_2 \hat{\varepsilon}_{t-2} + \dots + \rho_r \hat{\varepsilon}_{t-r} + v_t$, unde v_t este o variabila de perturbatie care verifica ipotezele asociate modelului clasic de regresie liniara.

In cazul nostru, $\hat{\varepsilon}_t = \alpha_0 + \alpha_1 x_t + \rho_1 \hat{\varepsilon}_{t-1} + \rho_2 \hat{\varepsilon}_{t-2} + v$

Statistica testului este: $LM = n \cdot R_{aux}^2$, care urmeaza o distributie χ^2

Breusch-Godfrey Serial Correlation LM Test:
Null hypothesis: No serial correlation at up to 2 lags

F-statistic	0.367278	Prob. F(2,31)	0.6956
Obs*R-squared	0.810141	Prob. Chi-Square(2)	0.6669

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 05/15/21 Time: 18:27
Sample: 1 35
Included observations: 35
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.681727	12.11060	0.138864	0.8905
VENIT_NATIONAL_BRUT__PER_CAPITA...	-6.18E-05	0.000361	-0.171131	0.8652
RESID(-1)	-0.087557	0.190772	-0.458964	0.6495
RESID(-2)	-0.137441	0.180473	-0.761559	0.4521
R-squared	0.023147	Mean dependent var	-1.83E-14	
Adjusted R-squared	-0.071387	S.D. dependent var	30.51371	
S.E. of regression	31.58409	Akaike info criterion	9.850394	
Sum squared resid	30924.19	Schwarz criterion	10.02815	
Log likelihood	-168.3819	Hannan-Quinn criter.	9.911755	
F-statistic	0.244852	Durbin-Watson stat	2.024679	
Prob(F-statistic)	0.864345			

-modelul auxiliar de regresie estimat este:

- $\hat{\varepsilon}_t = 1.68 - (6.18E - 05)x_t - 0.087\hat{\varepsilon}_{t-1} - 0.137\hat{\varepsilon}_{t-2}$

Pe baza rezultatelor output-ului din EViews, putem afirma ca parametrii ρ_1 si ρ_2 nu sunt semnificativ statistic deoarece avem $p\text{-value} > 0.05$, pentru un nivel de semnificatie de 5 %.

- Valoarea calculata a statisticii testului: $LM = n \cdot R_{aux}^2 = \text{obs} \cdot \text{R-squared} = 0.810141$;
- Valoarea critica a testului: $\chi_{critic}^2 = \chi_{\alpha;df}^2 = \chi_{0,05;2}^2 = 5.991464547$;

(valoarea critica este calculata in Excel cu ajutorul functiei =CHUSQ.INV.RT(0,05;2).

Decizia:

Cum $LM(\text{calc}) < 5,99$, rezulta ca nu respingem ipoteza nula , deci nu exista autocorelare a erorilor modelului de regresie.

Normalitatea erorilor aleatoare

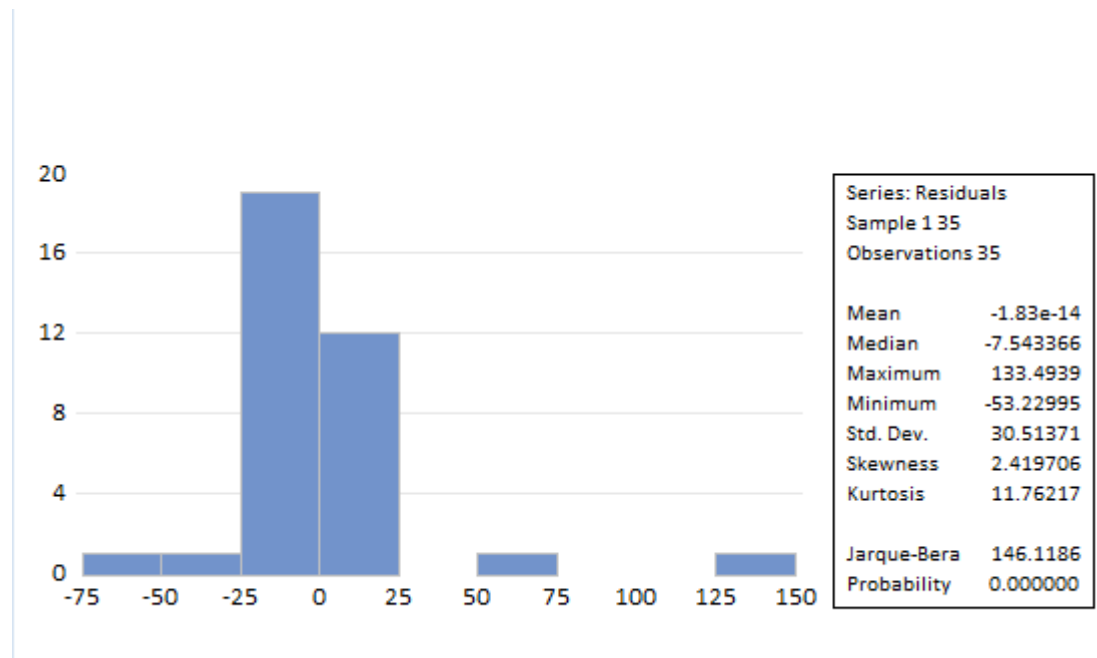
-aplicam **Testul Jarque-Bera(JB)** privind distributia normala a erorilor;

-acest test calculeaza mai intai coeficientul de asimetrie(Skewness) si coeficientul de boltire(Kurtosis) pentru reziduurile obtinute prin MCMMP.

Ipotezele testului sunt:

- H_0 : Reziduurile provin dintr-o distributie normala ($S=0$ si $K=3$), unde S = Skewness; K = Kurtosis;
- H_1 : Reziduurile provin dintr-o distributie normala.

Output din EViews:



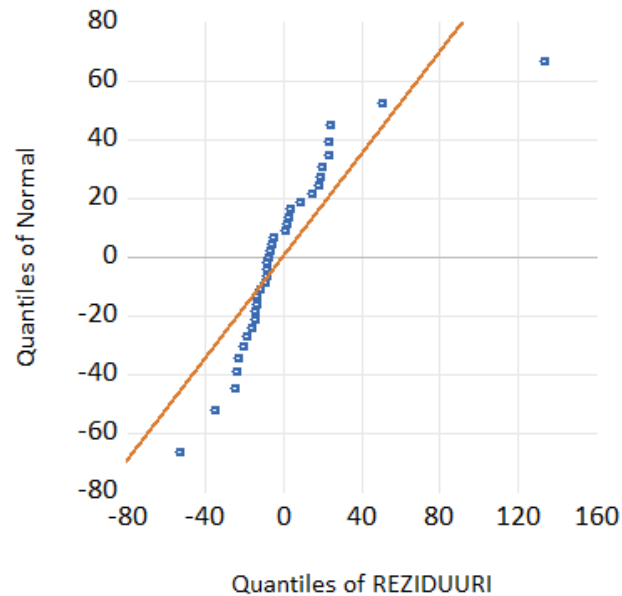
- **Jarque-Bera= 146,11 si Probability=0.000**

Decizia:

Deoarece $JB=146,11 > x_{critic}^2(3,84)$ si de asemenea, probabilitatea asociata statisticii

$JB < 0,05$, respingem H_0 , in favoarea ipotezei alternative, adica reziduurile nu provin dintr-o distributie normala.

Diagrama Q-Q Plot pentru erorile aleatoare



✚ Se observa ca punctele sunt putin deviate de la linia dreapta, ceea ce indica o distributie putin anormala, dar se apropie de o distributie normala.

Corectarea modelului de regresie

-pentru corectarea heteroscedasticitatii erorilor aleatoare, vom aplica metoda celor mai mici patrate – ponderata;

-consideram ca dispersiile erorilor σ_i^2 sunt necunoscute;

a)-presupunem ca dispersia erorilor este proportionala cu patraturul unei variabile explicative;\

- $\sigma_i^2 = \sigma^2 \cdot x_i^2$,

atunci modelul de regresie $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ se imparte prin x_i ;

-se obtine: $\frac{y_i}{x_i} = \beta_0 \frac{1}{x_i} + \beta_1 + \frac{\varepsilon_i}{x_i}$

Output-ul estimarii modelului de regresie obtinut dupa impartire este:

Dependent Variable: PIB_Y/VNB_X
 Method: Least Squares
 Date: 05/20/21 Time: 17:15
 Sample: 1 35
 Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.003245	0.000227	14.29288	0.0000
1/VNB_X	-0.972974	3.266676	-0.297848	0.7677
R-squared	0.002681	Mean dependent var		0.003196
Adjusted R-squared	-0.027541	S.D. dependent var		0.000915
S.E. of regression	0.000927	Akaike info criterion		-11.07324
Sum squared resid	2.84E-05	Schwarz criterion		-10.98437
Log likelihood	195.7817	Hannan-Quinn criter.		-11.04256
F-statistic	0.088714	Durbin-Watson stat		2.299167
Prob(F-statistic)	0.767686			

- Modelul de regresie nu este valid statistic, deoarece probabilitate Prob(F-statistic) este egala cu $0,76 > 0,05$. Deci, presupunerea ca dispersia erorilor este proportionala cu patrutul unei variabile explicative nu este relevanta.

b) -presupunem ca dispersia erorilor este proportionala cu o variabila explicative, adica

- $\sigma_i^2 = \sigma^2 \cdot x_i$,

atunci modelul de regresie $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ se imparte prin $\sqrt{x_i}$;

-se obtine: $\frac{y_i}{\sqrt{x_i}} = \beta_0 \cdot \frac{1}{\sqrt{x_i}} + \beta_1 \frac{x_i}{\sqrt{x_i}} + \frac{\varepsilon_i}{\sqrt{x_i}} \Rightarrow \frac{y_i}{\sqrt{x_i}} = \beta_0 \cdot \frac{1}{\sqrt{x_i}} + \beta_1 \sqrt{x_i} + \frac{\varepsilon_i}{\sqrt{x_i}}$;

Output-ul estimarii modelului de regresie obtinut prin impartire este:

Dependent Variable: PIB_Y/SQR(VNB_X)
 Method: Least Squares
 Date: 05/20/21 Time: 17:36
 Sample: 1 35
 Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
1/SQR(VNB_X)	-12.26588	6.662386	-1.841064	0.0746
SQR(VNB_X)	0.003813	0.000272	14.02984	0.0000
R-squared	0.658380	Mean dependent var		0.552116
Adjusted R-squared	0.648028	S.D. dependent var		0.275475
S.E. of regression	0.163432	Akaike info criterion		-0.729398
Sum squared resid	0.881427	Schwarz criterion		-0.640521
Log likelihood	14.76447	Hannan-Quinn criter.		-0.698718
Durbin-Watson stat	2.195380			

Testarea validitatii modelului

-deoarece valoarea lui F-statistic nu este afisata, aceasta trebuie calculata.

- SSR=Sum squared resid= 0,881;
- MSR=(S.E. of regression)² = 0,163² = 0,0265;

$$R^2 = \frac{SSE}{SST} \Leftrightarrow R^2 = \frac{SSE}{SSE + SSR} \Rightarrow SSE = \frac{R^2}{1 - R^2} \cdot SSR \Rightarrow SSE = 1,695;$$

- SST= SSE+SSR=> SST=2,576;
- Verificare SST: (S.D. dependent var)²=SST/(n-1) => 0,075=0,075 (A)

-Statistica testului:

$$F_{calc} = \frac{MSE}{MSR} = \frac{\frac{SSE}{k}}{\frac{SSR}{n-k}} \Rightarrow F_{calc} = 67,8.$$

-cu ajutorul functiei =F.INV.RT(0.05,1,34), aplicata in Excel, obtinem **valoarea critica a testului** F_{critic}= 4,14.

Decizia:

Cum F_{calc} > F_{critic}, rezulta ca modelul de regresie, obtinut prin impartirea la $\sqrt{x_i}$ a modelului initial, este valid statistic.

-Verificam daca heteroscedasticitatea a fost eliminata, aplicand **testul White**.

Heteroskedasticity Test: White				
Null hypothesis: Homoskedasticity				
F-statistic	0.751219	Prob. F(2,32)	0.4799	
Obs*R-squared	1.569597	Prob. Chi-Square(2)	0.4562	
Scaled explained SS	4.708364	Prob. Chi-Square(2)	0.0950	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 05/20/21 Time: 18:06				
Sample: 1 35				
Included observations: 35				
Collinear test regressors dropped from specification				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.019835	0.044009	-0.450710	0.6552
1/SQR(VNB_X)^2	193.2357	323.5536	0.597229	0.5546
SQR(VNB_X)^2	1.17E-06	9.83E-07	1.188737	0.2433
R-squared	0.044846	Mean dependent var	0.025184	
Adjusted R-squared	-0.014852	S.D. dependent var	0.066378	
S.E. of regression	0.066869	Akaike info criterion	-2.490349	
Sum squared resid	0.143086	Schwarz criterion	-2.357034	
Log likelihood	46.58111	Hannan-Quinn criter.	-2.444329	
F-statistic	0.751219	Durbin-Watson stat	1.760221	
Prob(F-statistic)	0.479929			

- ✚ Observand ca valoarea calculata a testului statistic este $W_{\text{calc}} = \text{Obs} * R\text{-squared} = 1,59$ cu probabilitatea Prob. Chi-Square(2)=0,4562 > 0,05 rezulta ca nu avem motive sa respingem ipoteza nula H_0 , deci erorile aleatoare sunt homoscedastice, nu exista heteroscedasticitate.
- ✚ Asadar, transformarea propusa prin impartire a corectat heteroscedasticitatea erorilor aleatoare ale modelului original de regresie.

Previzionarea punctuala si prin interval de incredere de 95% a valorii variabilei dependente Y daca variabila explicativa X_1 creste cu 10% fata de ultima valoarea inregistrata, utilizant modelul econometric corectat.

-pentru unitatea statistica: Romania.

- **Valoarea cunoscuta** a variabilei explicative: venit national brut/ per capita este, in 2018, egala cu 19,856 milioane euro. Dupa o crestere de 10% aceasta va fi: $x_0 = 21,841$ milioane euro.

Modelul econometric corectat:

- $\frac{y_i}{\sqrt{x_i}} = \beta_0 \cdot \frac{1}{\sqrt{x_i}} + \beta_1 \sqrt{x_i} + \frac{\varepsilon_i}{\sqrt{x_i}};$
- **Estimatiile parametrilor** sunt: $\hat{\beta}_0 = -12.265$ si $\hat{\beta}_1 = 0.0038;$
- **Dreapta de regresie de selectie**, la nivelul esantionului: $\frac{\hat{y}}{\sqrt{x}} = -12.265 \cdot \frac{1}{\sqrt{x}} + 0.0038 \sqrt{x}$
- **Ecuatia de regresie in esantion:** $\frac{\hat{y}_i}{\sqrt{x_i}} = -12.265 \frac{1}{\sqrt{x_i}} + 0.0038 \sqrt{x_i}$

Previzionarea mediei $M(Y|X=x_0=21,841)$

-estimatia punctuala a mediei

- $\frac{\hat{y}_0}{\sqrt{x_0}} = -12,265 \frac{1}{\sqrt{x_0}} + 0,0038 \sqrt{x_0}$

$$\Rightarrow \hat{y}_0 = -12.265 + 0.0038 * 21841$$

$$\Rightarrow \hat{y}_0 = 70.73 \text{ milioane euro.}$$

-interval de incredere $100(1-\alpha)\% = 95\%$ al mediei

$$CI_{100(1-\alpha)\%}(M(Y|X = x_0)) = \left[\hat{y}_0 - t_{\frac{\alpha}{2}; n-2} \cdot \hat{\sigma}_\varepsilon \cdot \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}; \hat{y}_0 + t_{\frac{\alpha}{2}; n-2} \cdot \hat{\sigma}_\varepsilon \cdot \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} \right]$$

In cazul nostru,

- $t_{\frac{\alpha}{2}; n-2} = 2.03$ (=T.INV.2T(0.05,33));
- $\hat{\sigma}_\varepsilon = \text{MSR} = 95.93$;
- $\bar{x} = 30,218$;
- $\sum_{i=1}^n (x_i - \bar{x})^2 = 8733,821$.
- $CI_{95\%}(M(Y|X = x_0 = 21,841)) =$
- $= [70.73 - 2.03 \cdot 95.93 \cdot \sqrt{\frac{1}{35} + \frac{107,371}{8733,821}}; 70.73 + 2.03 \cdot 95.93 \cdot \sqrt{\frac{1}{35} + \frac{107,371}{8733,821}}]$
- $= [70.73 - 38.94; 70.73 + 38.94]$
- $\Rightarrow CI_{95\%}(M(Y|X = x_0 = 21,841)) = [31.79; 109.67]$ milioane euro.

- Astfel, in selectii repetate, 95 din 100 de intervale ca cel obtinut anterior vor acoperi valoarea adevarata a PIB-ului/per capita mediu al Romaniei, daca VNB/ per capita este egal cu 21,841 milioane euro.

Partea a II-a

MODEL ECONOMETRIC MULTIFACTORIAL

Variabile ale modelului econometric sunt:

- variabila dependentă (explicată) → PIB/locuitor, milioane euro (Y).
- variabile independente (explicative):
 - venitul national brut/ locuitor, milioane euro (X_1);
 - consumul final/locuitor, milioane euro (X_2).

Modelul econometric este:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

Sau $y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \varepsilon_i$,

oricare ar fi i - unitate statistica, cu valori de la 1 la n , unde n reprezinta volumul esantionului.

Output-ul estimarii modelului de regresie liniara multifactoriala in Excel:

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.9266
R Square	0.8586
Adjusted R Square	0.8498
Standard Error	28.9301
Observations	35

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	2	162626.04	81313.02	97.15	2.55384E-14
Residual	32	26782.36	836.95		
Total	34	189408.41			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	-25.4924	10.5556	-2.4151	0.0216
X1	0.0027	0.0007	3.8662	0.0005
X2	0.0031	0.0013	2.4133	0.0217

Residual output:		
<i>Observation</i>	<i>Predicted Y</i>	<i>Residuals</i>
1	137.1029	-3.8029
2	30.8472	-4.4472
3	76.9077	-11.2077
4	162.5420	10.0580
5	146.3272	-12.4272
6	69.0728	-4.0728
7	165.4658	57.0342
8	70.9763	-15.5763
9	98.3061	-13.1061
10	124.7366	-8.6366
11	51.6123	-9.6123
12	111.4529	-14.5529
13	99.9444	-18.4444
14	57.0450	-7.0450
15	70.4021	-16.7021
16	203.6157	122.6843
17	54.4896	-8.4896
18	51.3538	34.4462
19	145.6153	2.9847
20	146.5337	-2.3337
21	56.6371	-13.7371
22	78.4581	-12.4581
23	49.2375	-14.5375
24	83.0486	-9.8486
25	61.1824	-6.7824
26	136.1549	3.9451
27	147.4509	5.5491
28	200.6854	7.6146
29	277.4350	-46.8350
30	275.5792	-33.5792
31	134.9809	-14.4809
32	10.3313	14.4687
33	-4.6236	19.4236
34	1.3095	18.9905
35	21.2834	5.5166

Output-ul estimarii modelului de regresie liniara multifactoriala in Eviews:

Estimation Command:

```
=====
LS PIB_PER_CAPITA__MILIOANE_EURO____Y C VENIT_NATIONAL_BRUT__PER_CAPITA__MILIOANE_EURO____X1
CONSUM_FINAL__PER_CAPITA__MILIOANE_EURO____X2
```

Estimation Equation:

```
=====
PIB_PER_CAPITA__MILIOANE_EURO____Y = C(1) + C(2)
*VENIT_NATIONAL_BRUT__PER_CAPITA__MILIOANE_EURO____X1 + C(3)
*CONSUM_FINAL__PER_CAPITA__MILIOANE_EURO____X2
```

Substituted Coefficients:

```
=====
PIB_PER_CAPITA__MILIOANE_EURO____Y = -25.4924270786 + 0.00272311073514
*VENIT_NATIONAL_BRUT__PER_CAPITA__MILIOANE_EURO____X1 + 0.00306004243125
*CONSUM_FINAL__PER_CAPITA__MILIOANE_EURO____X2
```

Dependent Variable: PIB_PER_CAPITA__MILIOANE_EURO____Y

Method: Least Squares

Date: 04/18/21 Time: 14:29

Sample: 1 35

Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-25.49243	10.55564	-2.415053	0.0216
VENIT_NATIONAL_BRUT__PER_CAPITA...	0.002723	0.000704	3.866178	0.0005
CONSUM_FINAL__PER_CAPITA__MILIO...	0.003060	0.001268	2.413345	0.0217
R-squared	0.858600	Mean dependent var		102.9571
Adjusted R-squared	0.849762	S.D. dependent var		74.63803
S.E. of regression	28.93007	Akaike info criterion		9.649456
Sum squared resid	26782.36	Schwarz criterion		9.782772
Log likelihood	-165.8655	Hannan-Quinn criter.		9.695477
F-statistic	97.15411	Durbin-Watson stat		2.042596
Prob(F-statistic)	0.000000			

Interpretare SUMMARY OUTPUT, din Excel:

- **R=0.9266** -> Valoarea coeficientului de corelatie este foarte apropiata de 1, ceea ce arata ca intre PIB/ loc si venitul national brut/ loc si consumul final/ loc exista o legatura foarte puternica, legatura pusa in evidenta prin modelul de regresie liniara multifactoriala.
- **R² = 0.8586** (coeficientul de determinatie)-> La nivelul esantionului de 35 de tari, arata ca 85,86% din variatia totala a variabilei dependente (PIB-ului/ loc) este explicata de variatia totala a celor doua variabile independente(VNB/ loc si CF/loc), iar restul de 14,14% din variatia totala este data de factori reziduali.

- Pe coloana “**Actual**” sunt valorile observate y_i ale variabilei dependente PIB/ loc (Y), pe urmatoare coloana “**Fitted**” sunt valorile ajustate \hat{y}_i ale variabilei dependente PIB/ loc (Y), ajustate prin modelul de regresie estimata si pe ultima coloana “**Residual**” sunt reziduurile modelului de regresie estimate $\hat{\varepsilon}_i$, pentru toate unitatile statistice i ale esantionului de volum $n= 35$ tari.

obs	Actual	Fitted	Residual	Residual Plot
1	133.300	137.103	-3.80292	
2	26.4000	30.8472	-4.44720	
3	65.7000	76.9077	-11.2077	
4	172.600	162.542	10.0580	
5	133.900	146.327	-12.4272	
6	65.0000	69.0728	-4.07283	
7	222.500	165.466	57.0342	
8	55.4000	70.9763	-15.5763	
9	85.2000	98.3061	-13.1061	
10	116.100	124.737	-8.63656	
11	42.0000	51.6123	-9.61229	
12	96.9000	111.453	-14.5529	
13	81.5000	99.9444	-18.4444	
14	50.0000	57.0450	-7.04502	
15	53.7000	70.4021	-16.7021	
16	326.300	203.616	122.684	
17	46.0000	54.4896	-8.48959	
18	85.8000	51.3538	34.4462	
19	148.600	145.615	2.98471	
20	144.200	146.534	-2.33366	
21	42.9000	56.6371	-13.7371	
22	66.0000	78.4581	-12.4581	
23	34.7000	49.2375	-14.5375	
24	73.2000	83.0486	-9.84861	
25	54.4000	61.1824	-6.78241	
26	140.100	136.155	3.94514	
27	153.000	147.451	5.54912	
28	208.300	200.685	7.61456	
29	230.600	277.435	-46.8350	
30	242.000	275.579	-33.5792	
31	120.500	134.981	-14.4809	
32	24.8000	10.3313	14.4687	
33	14.8000	-4.62361	19.4236	
34	20.3000	1.30948	18.9905	
35	26.8000	21.2834	5.51660	

Estimatiile parametrilor modelului de regresie liniara multifactoriala

Pe baza esantionului de volum $n=35$ tari, am determinat cu ajutorul softurilor (**Eviews**, **Excel**) estimatorii $\hat{\beta}_0$, $\hat{\beta}_1$ si $\hat{\beta}_2$ ai parametrilor β_0 - intercept si β_1 si β_2 – pante ale modelului de regresie liniara multifactoriala.

- Astfel, se obtine: $\hat{\beta}_0 = -25.49$; $\hat{\beta}_1 = 0.0027$; $\hat{\beta}_2 = 0.0031$.

Dreapta de regresie de selectie(la nivelul esantionului) este:

$$\bullet \quad \hat{y} = -25.49 + 0.0027X_1 + 0.0031X_2;$$

Ecuatia de regresie liniara, in esantion, este $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i,1} + \hat{\beta}_2 x_{i,2}$,

adica: $\hat{y} = -25.49 + 0.0027x_{i,1} + 0.0031x_{i,2}$,

unde \hat{y}_i -reprezinta valorile ajustate ale variabilei Y, pentru unitatile statistice i, i ia valori intre 1 si 35,

iar $\hat{\varepsilon}_i = y_i - \hat{y}_i$ sunt reziduurile modelului estimat de regresie(valorile observate ale variabilei eroare ε).

Interpretarea paramterilor ai modelului de regresie estimat liniar multifactorial

✚ $\hat{\beta}_0 = -25.49$, este estimatorul parametrului β_0 , numit intercept sau termen liber, care arata, pe baza datelor din esantion, nivelul mediu al variabilei dependente atunci cand nivelul tuturor variabilelor independente este egal cu 0 unitati. Daca venitul national brut si consumul national brut ar fi egal cu 0 mil euro,atunci PIB-ul/ loc mediu estimat este de - 25.49 mil euro(acest parametru nu are intotdeauna o interpretare economica).

✚ $\hat{\beta}_1 = 0.0027$, este parametrul panta corespunzator variabilei X_1 (VNB/ loc). Daca variabila X_1 creste cu o unitate(1 mil euro) si nivelul consumului final/per capita ramane constant sau nemodificat, PIB-ul/loc va creste, in medie, cu 0.0027 mil euro.

✚ $\hat{\beta}_2 = 0.0031$, este parametrul panta corespunzator variabilei X_2 (CF/ loc). Daca variabila X_2 creste cu o unitate(1 mil euro) si nivelul veniturii national brut/per capita ramane constant sau nemodificat , PIB-ul/loc va creste, in medie, cu 0.0031 mil euro.

Multicoloniaritatea variabilor explicative

Modelul de regresie 1

-este un model multifactorial cu $k=2$ variabile explicative: $y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \varepsilon_i$;

Output-ul de estimare pentru modelul de regresie 1:

Dependent Variable: PIB_PER_CAPITA__MILIOANE_EURO___Y

Method: Least Squares

Date: 05/16/21 Time: 11:06

Sample: 1 35

Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-25.49243	10.55564	-2.415053	0.0216
VENIT_NATIONAL_BRUT__PER_CAPITA...	0.002723	0.000704	3.866178	0.0005
CONSUM_FINAL__PER_CAPITA__MILIO...	0.003060	0.001268	2.413345	0.0217
R-squared	0.858600	Mean dependent var	102.9571	
Adjusted R-squared	0.849762	S.D. dependent var	74.63803	
S.E. of regression	28.93007	Akaike info criterion	9.649456	
Sum squared resid	26782.36	Schwarz criterion	9.782772	
Log likelihood	-165.8655	Hannan-Quinn criter.	9.695477	
F-statistic	97.15411	Durbin-Watson stat	2.042596	
Prob(F-statistic)	0.000000			

- Rezultatele estimarii modelului de regresie 1 arata ca Venitul Nation Brut si Consumul Final, impreuna, explica 85,86% din variatia PIB-ului.
- Deoarece p-value(Prob) pentru cei doi parametrii panta (β_1, β_2): 0,0027 si 0,003 < 0,05 , rezulta ca acestia sunt semnificativ statistic. De asemenea, Prob(F-Statistic) este mai mica decat nivelul de semnificatie 0,05, ceea ce inseamna ca modelul de regresie este valid statistic.

Detectarea multicoliniaritatiei pe baza coeficientilor de corelatie dintre variabilele explicative

Correlation				
	A	B	C	D
1				
2				
3	PIB_PE...	1.000000	0.912614	0.890254
4	VENIT_...	0.912614	1.000000	0.898240
5	CONS...	0.890254	0.898240	1.000000

- Intre variabilele X_1 si X_2 exista o legatura liniara directa puternica deoarece coeficientul de corelatie liniara Pearson este destul de apropiat de valoarea 1.
- Cum $r_{x_1, x_2} = 0,89$, rezulta ca variabilele X_1 si X_2 sunt corelate foarte puternic.

Model de regresie 2

-regresam variabila X_2 in raport cu variabila X_1 ;

Dependent Variable: CONSUM_FINAL__PER_CAPITA__MILIOANE_EURO
__X2

Method: Least Squares

Date: 05/16/21 Time: 11:10

Sample: 1 35

Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.765066	1449.170	0.004668	0.9963
VENIT_NATIONAL_BRUT__PER_CAPITA...	0.498962	0.042499	11.74047	0.0000
R-squared	0.806835	Mean dependent var		15084.84
Adjusted R-squared	0.800982	S.D. dependent var		8903.036
S.E. of regression	3971.774	Akaike info criterion		19.46726
Sum squared resid	5.21E+08	Schwarz criterion		19.55614
Log likelihood	-338.6770	Hannan-Quinn criter.		19.49794
F-statistic	137.8387	Durbin-Watson stat		1.817572
Prob(F-statistic)	0.000000			

- $\hat{x}_{i,2} = 6.765066 + 0.498962 x_{i,2}$
- F-Statistic= 137.8387 cu $\text{Prob(F-Statistic)} < 0.05$
- DW= 1.817572
- $R^2 = 0.806$

🚦 Ecuatia de regresie arata ca exista coliniaritate puternica intre variabilele X_2 si X_1 .

Criteriul factorului de inflatie a variantei(Variance Inflationary Factor)

-coeficientul de determinatie R_j^2 ;

-calcularea factorului de inflatie al variantei: $VIF = \frac{1}{1-R_j^2}$;

Output-ul din EViews pentru VIFX2 al modelului de regresie 2:

	Value
VIFX2	5.176929

🚦 Cum $VIFX2 = 5,177 \in [5,10)$, rezulta ca multicoloniaritatea este moderata.

Output-ul din EViews pentru VIFX2 si VIFX1 al modelului de regresie 2:

Variance Inflation Factors
Date: 05/16/21 Time: 12:19
Sample: 1 35
Included observations: 35

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	111.4215	4.659488	NA
VENIT_NATIONAL_B...	4.96E-07	24.12182	5.176929
CONSUM_FINAL_P...	1.61E-06	20.47606	5.176929

Eliminarea multicoliniarit tii

-pentru a elimina multicoliniaritatea, vom elimina una din variabile coloniare, deoarece nu se pot
logaritma datele care au valori negative.

Modelul de regresie 4:

-regresam variabila Y in raport cu variabila X_1

Dependent Variable: PIB_PER_CAPITA__MILIOANE_EURO____Y
Method: Least Squares
Date: 05/16/21 Time: 12:36
Sample: 1 35
Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-25.47173	11.30089	-2.253958	0.0310
VENIT_NATIONAL_BRUT__PER_CAPITA...	0.004250	0.000331	12.82358	0.0000
R-squared	0.832864	Mean dependent var		102.9571
Adjusted R-squared	0.827799	S.D. dependent var		74.63803
S.E. of regression	30.97259	Akaike info criterion		9.759528
Sum squared resid	31656.95	Schwarz criterion		9.848405
Log likelihood	-168.7917	Hannan-Quinn criter.		9.790208
F-statistic	164.4441	Durbin-Watson stat		2.123386
Prob(F-statistic)	0.000000			

- Comparand Prob(p-value) pentru variabila X_1 din modelul de regresie 4 cu cel din modelul de regresie initial multifactorial(1), se observa ca parametrul panta corespunzator variabilei X_1 este mai semnificativ.

Model de regresie 5:

-regresam variabila Y in raport cu variabila X_2 ;

Dependent Variable: PIB_PER_CAPITA__MILIOANE_EURO____Y
Method: Least Squares
Date: 05/16/21 Time: 12:36
Sample: 1 35
Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-9.626825	11.59981	-0.829912	0.4126
CONSUM_FINAL__PER_CAPITA__MILIO...	0.007463	0.000665	11.22834	0.0000
R-squared	0.792551	Mean dependent var	102.9571	
Adjusted R-squared	0.786265	S.D. dependent var	74.63803	
S.E. of regression	34.50627	Akaike info criterion	9.975604	
Sum squared resid	39292.52	Schwarz criterion	10.06448	
Log likelihood	-172.5731	Hannan-Quinn criter.	10.00628	
F-statistic	126.0755	Durbin-Watson stat	1.937222	
Prob(F-statistic)	0.000000			

- ✚ Comparand Prob(p-value) pentru variabila X_2 din modelul de regresie 5 cu cel din modelul de regresie initial multifactorial(1), se observa ca parametrul panta corespunzator variabilei X_1 este cu mult mai semnificativ.
- ✚ Cele doua modele de regresie 4 si 5 arata ca, in cazul multicolinariatii, eliminarea unei variabilei va face ca cealalta variabila explicativa sa fie mai semnificativa statistic.

Verificarea indeplinirii ipotezelor fundamentale ale modelului clasic de regresie liniara

1. Heteroscedasticitatea erorilor aleatoare(ipoteza I3)

-pentru detectarea heteroscedasticitatii, vom aplica **testul White**, in EViews;

Heteroskedasticity Test: White				
Null hypothesis: Homoskedasticity				
F-statistic	2.159619	Prob. F(5,29)	0.0863	
Obs*R-squared	9.496265	Prob. Chi-Square(5)	0.0908	
Scaled explained SS	43.57612	Prob. Chi-Square(5)	0.0000	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 05/16/21 Time: 13:03				
Sample: 1 35				
Included observations: 35				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	969.0195	1600.379	0.605494	0.5496
X1^2	-1.08E-05	6.94E-06	-1.562293	0.1291
X1*X2	2.31E-05	1.95E-05	1.182839	0.2465
X1	0.307654	0.184245	1.669808	0.1057
X2^2	6.02E-06	1.45E-05	0.416384	0.6802
X2	-0.803507	0.398985	-2.013875	0.0534
R-squared	0.271322	Mean dependent var	765.2104	
Adjusted R-squared	0.145688	S.D. dependent var	2572.506	
S.E. of regression	2377.742	Akaike info criterion	18.54049	
Sum squared resid	1.64E+08	Schwarz criterion	18.80713	
Log likelihood	-318.4587	Hannan-Quinn criter.	18.63254	
F-statistic	2.159619	Durbin-Watson stat	1.884566	
Prob(F-statistic)	0.086348			

Ecuatia de regresie auxiliara estimata este:

$$\hat{\varepsilon}_i^2 = 969.0195 + 0.307 \cdot X_{i1} - (1.08E-05) \cdot X_{i1}^2 + (2.31E-05) \cdot X_{i1}X_{i2} - 0.803 \cdot X_{i2} - (6.02E-06) \cdot X_{i2}^2$$

Valoarea calculata a testului statistic $F=F\text{-Statistic}= 2.15$, cu probabilitatea asociata Prob $F=0.08>0.05$;

Valoarea calculata a testului statistic $W_{\text{calc}} = n \cdot R_a^2 = \text{obs} * R - \text{squared} = 9.49$, cu probabilitatea asociata Prob. Chi-Square(5)= $\alpha=0.0908>0.05$.

$$\chi_{\text{critic}}^2 = \chi_{\alpha;df}^2 = \chi_{0.05;5}^2 = 11.07, \text{ returnata de functia } = \text{CHISQ.INV.RT}(0.05,5) \text{ in Excel};$$

Decizia

- ✚ Pentru ca probabilitatile asociate Prob $F=0.08$ si Prob Chi-Square(5) = 0.0908 sunt mai mari decat nivelul de semnificatie de 0.05, nu respingem ipoteza nula si acceptam H_1 .
- ✚ De asemenea, $W_{\text{calc}} < \chi_{\text{critic}}^2$, rezulta ca nu avem motive sa respingem ipoteza nula H_0 , deci erorile aleatoare sunt homoscedastice, nu exista heteroscedasticitate.

2. Autocorelarea erorilor aleatoare

Testul Breusch-Godfrey

-acest test se foloseste pentru detectarea autocorelarii de ordin superior lui 1, in seria reziduurilor.

Ipotezele testului sunt:

- $H_0: \rho_1 = \rho_2 = \dots = \rho_r$ (erorile nu sunt autocorelate, nu exista autocorelare de ordin r a erorilor aleatoare)
- $H_1: \rho_i \neq 0, i \in \{1, 2, \dots, r\}$ (exista autocorelare de ordin r a erorilor aleatoare)

-vom aplica testul Breusch-Godfrey pentru autocorelarea de ordin r=2.

-se estimeaza parametrii modelului de regresie initial prin MCMMP si se obtin reziduurile $\hat{\varepsilon}_t$.

-se considera un **model de regresie auxiliar**:

$\hat{\varepsilon}_t = \alpha_0 + \alpha_1 x_{t,1} + \alpha_2 x_{t,2} + \dots + \alpha_k x_{t,k} + \rho_1 \hat{\varepsilon}_{t-1} + \rho_2 \hat{\varepsilon}_{t-2} + \dots + \rho_r \hat{\varepsilon}_{t-r} + v_t$, unde v_t este o variabila de perturbatie care verifica ipotezele asociate modelului clasic de regresie liniara.

- In cazul nostru, $\hat{\varepsilon}_t = \alpha_0 + \alpha_1 x_{t1} + \alpha_2 x_{t2} + \rho_1 \hat{\varepsilon}_{t-1} + \rho_2 \hat{\varepsilon}_{t-2} + \rho_3 \hat{\varepsilon}_{t-3} + v$

Statistica testului este: $LM = n \cdot R_{aux}^2$, care urmeaza o distributie χ^2 .

Breusch-Godfrey Serial Correlation LM Test:
Null hypothesis: No serial correlation at up to 3 lags

F-statistic	0.394471	Prob. F(3,29)	0.7579
Obs*R-squared	1.372258	Prob. Chi-Square(3)	0.7121

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 05/16/21 Time: 13:27
Sample: 1 35
Included observations: 35
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.432411	11.52741	-0.124261	0.9020
X1	-0.000300	0.000857	-0.350356	0.7286
X2	0.000693	0.001570	0.441113	0.6624
RESID(-1)	0.013499	0.191086	0.070642	0.9442
RESID(-2)	0.167486	0.220173	0.760701	0.4530
RESID(-3)	-0.143609	0.194102	-0.739861	0.4653

R-squared	0.039207	Mean dependent var	8.10E-15
Adjusted R-squared	-0.126447	S.D. dependent var	28.06629
S.E. of regression	29.78793	Akaike info criterion	9.780888
Sum squared resid	25732.30	Schwarz criterion	10.04752
Log likelihood	-165.1655	Hannan-Quinn criter.	9.872929
F-statistic	0.236682	Durbin-Watson stat	2.013420
Prob(F-statistic)	0.943085		

-modelul auxiliar de regresie estimat este:

$$\hat{\varepsilon}_t = -1.432 - 0.0003x_{t1} + 0.00069x_{t2} + 0.0134\hat{\varepsilon}_{t-1} + 0.167\hat{\varepsilon}_{t-2} - 0.143\hat{\varepsilon}_{t-3}$$

✚ Se poate observa, din output-ul din EViews, ca niciunul dintre parametrii modelului auxiliar nu sunt semnificativ statistic deoarece avem p-value(Prob) mult mai mare decat 0,05.

- **Valoarea calculata** a testului este: $LM = n \cdot R_{aux}^2 = \text{obs} \cdot \text{R-squared} = 1,3722$;
- **Valoarea critica** a testului este: $\chi_{critic}^2 = \chi_{\alpha;df}^2 = \chi_{0,05;3}^2 = 7,814$;

(valoarea critica este calculata in Excel cu ajutorul functiei =CHUSQ.INV.RT(0,05;3).

Decizia:

✚ Cum $LM(\text{calc}) < 7,814$, rezulta ca nu respingem ipoteza nula, deci nu exista autocorelare a erorilor modelului de regresie.

3. Normalitatea erorilor aleatoare

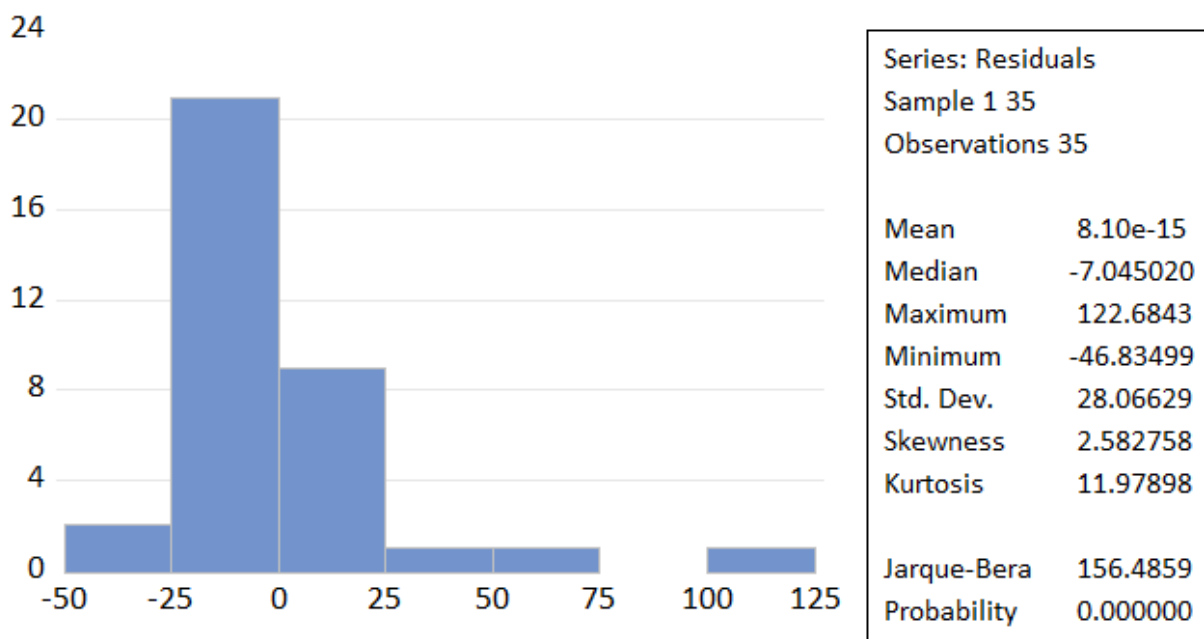
-aplicam **Testul Jarque-Bera(JB)** privind distributia normala a erorilor;

-acest test calculeaza mai intai coeficientul de asimetrie(Skewness) si coeficientul de boltire(Kurtosis) pentru reziduurile obtinute prin MCMMP.

Ipotezele testului sunt:

- H_0 : Reziduurile provin dintr-o distributie normala ($S=0$ si $K=3$), unde S = Skewness; K = Kurtosis;
- H_1 : Reziduurile provin dintr-o distributie normala.

Output din EViews:



Jarque-Bera= 156,48 si Probability=0.000

✚ Deoarece $JB=156,48 > \chi^2_{critic}$ si in plus, probabilitatea asociata statisticii $JB < 0,05$, respingem H_0 , in favoarea ipotezei alternative, adica reziduurile nu provin dintr-o distributie normala.

Variabila DUMMY

Vom introduce o variabila de tip Dummy, in functie de populatia pe care o are fiecare tara, astfel:

- **D=1**, daca tara respectiva are populatia peste cea medie din Europa;
- **D=0**, daca tara respectiva are populatia sub cea medie din Europa;

Nr_crt	Tara	PIB/per capita (milioane euro) _ Y	Populatia(numar locuitori)	Variabila Dummy
1	Belgia	133.3	11,398,589	0
2	Bulgaria	26.4	7,050,034	0
3	Republica Ceha	65.7	10,610,055	0
4	Danemarca	172.6	5,781,190	0
5	Germania	133.9	82,792,351	1
6	Estonia	65.0	1,319,133	0
7	Irlanda	222.5	4,830,392	0
8	Grecia	55.4	10,741,165	0
9	Spania	85.2	46,658,447	1
10	Franta	116.1	67,026,224	1
11	Croatia	42.0	4,105,493	0
12	Italia	96.9	60,483,973	1
13	Cipru	81.5	864,236	0
14	Letonia	50.0	1,934,379	0
15	Lituania	53.7	2,808,901	0
16	Luxembourg	326.3	602,005	0
17	Ungaria	46.0	9,778,371	0
18	Malta	85.8	475,701	0
19	Olanda	148.6	17,181,084	0
20	Austria	144.2	8,822,267	0
21	Polonia	42.9	37,976,687	1
22	Portugalia	66.0	10,291,027	0
23	Romania	34.7	19,533,481	1
24	Slovenia	73.2	2,066,880	0
25	Slovacia	54.4	5,443,120	0
26	Finlanda	140.1	5,513,130	0
27	Suedia	153.0	10,120,242	0
28	Islanda	208.3	348,450	0
29	Norvegia	230.6	5,295,619	0
30	Elvetia	242.0	8,484,130	0
31	Regatul Unit	120.5	66,273,576	1
32	Muntenegru	24.8	622,359	0
33	Albania	14.8	2,870,324	0
34	Serbia	20.3	7,001,444	0
35	Turcia	26.8	80,810,525	1

Exista diferente intre tarile din Europa in functie de numarul de locuitori(peste medie, sub medie), in ceea ce priveste calitatea vietii, indicata de PIB/per capita(milioane euro)?

-vom considera regresia $PIB_i = \beta_0 + \beta_1 D_i + \varepsilon_i$;

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.154145714
R Square	0.023760901
Adjusted R Square	-0.005822102
Standard Error	74.85498937
Observations	35

ANOVA

	df	SS	MS	F	Significance F
Regression	1	4500.514418	4500.514418	0.803194362	0.376631425
Residual	33	184907.8913	5603.269433		
Total	34	189408.4057			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	109.1296296	14.40584942	7.575369313	1.02555E-08	79.82070861	138.4385506	79.82070861	138.4385506
Variabila Dummy	-27.00462963	30.1319959	-0.896211115	0.376631425	-88.30863622	34.29937696	-88.30863622	34.29937696

✚ $\hat{\beta}_0 = 109,12$, rezulta ca nivelul mediu al PIB-ului/per capita al tarilor cu o populatie sub medie este 109,12 milioane euro.

✚ $\hat{\beta}_1 = 109,12 - 27,0046 = 82.1154$, ceea ce inseamna ca nivelul mediu al PIB-ului/per capita al tarilor cu o populatie peste medie este 82,12 milioane euro.

Concluzii

- ✚ In acest proiect, s-a urmarit caracterizarea calitatii vietii in tarile din Europa, analizand econometric legatura dintre trei indicatori economici: produsului intern brut, venitul national brut si consumul final, pe un cap de locuitor.
 - ✚ Dupa verificare ipotezelor fundamentale, in ceea ce priveste homoscedasticitatea, autocorelarea si normalitatea erorilor aleatoare , reiese ca datele constituie un model multifactorial destul de omogen.
 - ✚ In urma analizei, se poate afirma ca exista o legatura intre cei trei indicatori macroeconomici, adica Produsul Intern Brut este influentat direct de ceilalti doi indicatori macroeconomici: Venitul National Brut si Consumul Final.
 - ✚ De asemenea, s-a demonstrat, cu ajutorul variabilei Dummy, ca PIB-ul/per capita al unei tari nu este direct proportional cu numarul locuitorilor acesteia. Din contra, tarile cu o populatie mai mica au avut nivelul mediu al PIB-ului mai mare decat restul, mai exact cu 27 milioane de euro.
 - ✚ Concluzionam ca satsifacerea populatiei fata de nivelul de trai este influentata direct de venitul de care acestia beneficiaza in tara lor si ca Produsul Intern Brut poate fi un factor care sa descrie bunastarea unei societati.
-

Bibliografia

- https://ro.wikipedia.org/wiki/Produs_intern_brut
- http://www.ipe.ro/RePEc/WorkingPapers/cs18_1.pdf
- <https://www.macrotrends.net/countries/SRB/serbia/gni-per-capita>
- https://scholar.google.ro/scholar?q=European+Quality+of+Life+Survey&hl=ro&as_sdt=0&as_vis=1&oi=scholart
- <http://www.iccv.ro/wp-content/uploads/2018/03/Raport-Calitatea-Vietii.pdf>
- <https://link.springer.com/article/10.1007/s11205-018-1854-y>
- <https://ec.europa.eu/eurostat/data/database>
- *Cursuri si seminarii Econometrie anul II, Conf. univ. dr. Mihaela Covrig, Departamentul de Statistica si Econometrie.*