Academia de Studii Economice din București

Facultatea de Cibernetică, Statistică și Informatică Economică

Proiect Serii de timp

Prof. Coordonator: Conf. Univ. Dr. Daniel Traian Pele

Studenți:

Negulescu Adrian

Pîrvu Remus Andrei

Pop Vlad George

Anul III, Grupa 1072

București

2017

**Introducere**

**Motivarea alegerii temei**

**Aplicația 1 – Evoluția produsului intern brut al Italiei din 1985 până în 2016**

Țara aleasă pe care a fost efectuată analiza este Italia pe perioada 1985 – 2016 cu date anuale, deoarece Italia este una dintre cele mai dezvoltate țări din Europa, iar în ultima vreme aceasta trece printr-o perioadă economică dificilă așa că am vrut să vedem evoluția acesteia.

**Aplicația 2 – Evolutia vanzarii de bere in Marea Britanie**

Am ales această temă deoarece berea este unul dintre cele mai consumate alimente ce contine alcool. In ultima vreme consumul de bere a crescut, iar promovarea acesteia in media prin intermediul spoturilor publicitare, dar si a reclamelor online a devenit din ce in ce mai intensa.

Datele ce numarul de butoaie totaal vandut (1 butoi = 159 l) sunt inregistrate lunar, incepand cu 01.01.2009 pana in 31.01.2017

Conform lui Platon : “*Este un om inteligent cel care a inventat berea.*”

**Aplicația 3 – PIB si exportul de bunuri si servicii. Cointegrare. Cauzalitate Granger**

Prin Aplicatia 3 am testat daca pentru intervalul de timp ales – anii 1985-2016, intre seria PIB si cea a exportului de bunuri si servicii exista vreo legatura. De aceea am analizat cointegrarea si cauzalitatea.

1. **Aplicația 1 – Evoluția produsului intern brut al Italiei din 1985 până în 2016**
   1. **Descrierea datelor**

Datele inițiale sunt formate dintr-o serie de 32 de observatii reprezentând produsul intern brut per an si exportul de bunuri si servicii, în perioada 1985-2016, iar unitatile de masura sunt miliarde de dolari.

Datele au fost preluate de pe site-ul *data.worldbank.org*.

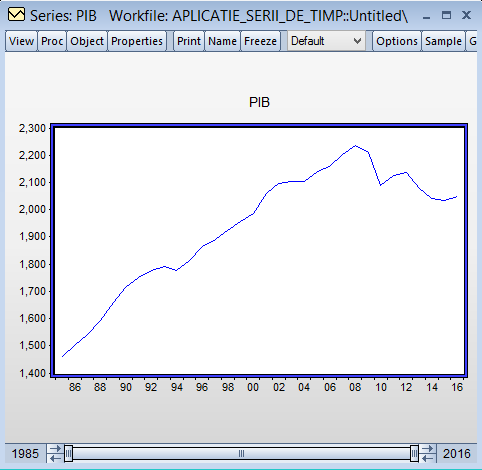


Figura 1 - Graficul produsului intern brut in perioada 1985-2016

Așa cum se poate observa în *Figura 1,* seria are un trend ascendent în prima parte, aproximativ până în anul 2008 cand a intervenit Criza Mondiala, valoarea produsului internet brut crescând destul de repede, iar apoi se observă că are un trend descendent pana prin anul 2014 cand incepe din nou sa creasca.

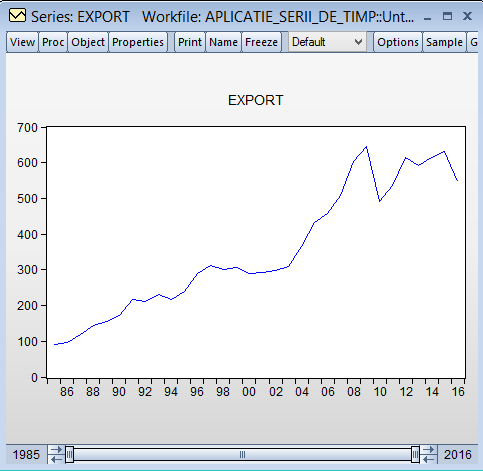


Figura 2 - Graficul exportului de bunuri si servicii in perioada 1985-2016

In *Figura 2,* seria are un trend ascendent în prima parte, aproximativ până în anul 2008 cand a intervenit Criza Mondiala, valoarea exportului de bunuri si servicii crescând constant, dupa care s-a inregistrat o scadere bruta in anul 2008, ca mai apoi dupa criza sa inceapa sa creasca din nou, iar din anul 2015 aceasta scade din nou.

* 1. **Testarea staționarității seriei de timp, pe baza corelogramei**

Un test simplu al staționarității seriei este bazat pe funcția de autocorelație (ACF).

Graficul funcției de autocorelație în raport cu decalajul k, se numeşte corelogramă.

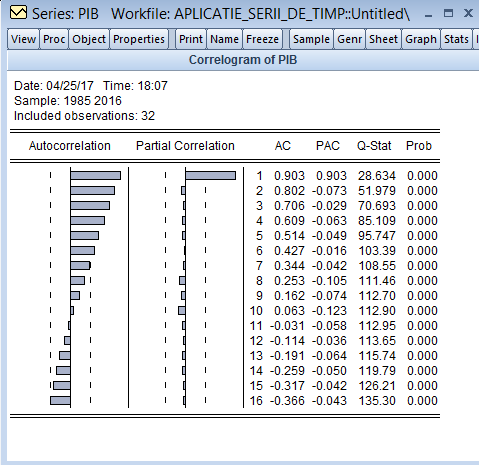


Figura 3 - Corelograma

Așa cum se poate observa în *Figura 3* aceasta începe cu valori mari atât pentru coeficientul de autocorelație, cât și pentru cel de autocorelație parțială (0.903 pentru lag-ul 1) și scade treptat. Se observă că după primele 7 lag-uri valoarea coeficientului de autocorelatie scade spre 0, dar dupa lagul 12 aceasta incepe sa creasca din nou, ceea ce ne indică faptul că seria este nestaționară. Acest lucru poate fi dedus și din faptul că valorile coeficientului AC converg încet către 0, ceea ce este caracteristic seriilor staționare. Se remarcă faptul că structura corelațiilor dintre observații (ACF, respectiv PACF) depășește liniile punctate, ceea ce înseamnă că seria nu este staționară.

* 1. **Testul Dickey-Fuller (fără intercept și fără trend)**

Prima parte a testului prezintă informaţii cu privire la tipul testului (AFD, variabilele exogene introduse) şi cuprinde rezultatul testului, valorile critice pentru fiecare nivel de relevanţă (1, 5 şi 10 la sută), şi probabilitatea, p, asociată rezultatului testului.

Datele de tip serii cronologice, de foarte multe ori, tind să se modifice în aceeaşi directie, din cauza trendului care este comun tuturor.

Dacă o serie de timp are o rădăcină unitară, atunci seria prezintă un trend stochastic şi este de tip DS (Difference-Stationarity). Trendul de tip stochastic se elimină prin calculul diferențelor de ordinul 1 sau de ordin mai mare sau egal cu 2 .

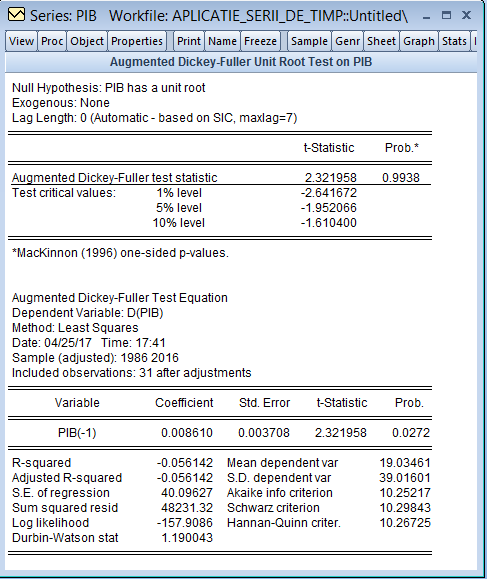


Figura 4 - Testul Dickey-Fuller pentru seria PIB (fără intercept și fără trend)

Formularea ipotezelor

H0: *seria este nestaționară (are rădăcină unitară)*

H1: *seria este staționară*

Așa cum se poate observa în *Figura 4,* valoarea calculată pentru t-statistic este 2,321 și probabilitatea asociată acesteia este de 0.9938.

Dacă valoarea testului este mai mare decât valoarea critică, nu este respinsă ipoteza nulă, deci seria are o rădăcină unitară (este nestaţionară).

Valoarea testului (2,321) este mai mare decât valorile critice ale nivelurilor de semnificație: -2.64167 pentru un nivel de semnificație de 1%, -1.95 pentru nivelul de 5%, -1.61 pentru nivelul de 10%. Astfel, acceptăm impoteza nulă, deci putem trage conluzia că seria de timp nu este staționară, deci are o rădăcină unitară.

Faptul că seria nu este staționară se poate deduce și din analiza probabilității, care atunci când are o valoare mai mică decât 0.05, înseamnă că seria este staționară, iar când are o valoare mai mare înseamnă că seria este nestaționară. În cazul acesta Prob = 0.9938 >> 0.05, ceea ce afirmă, de asemenea, faptul că seria este nestaționară.

Seria prezintă un trend stochastic şi este de tip DS(Difference-Stationarity).

* 1. **Testul Dickey-Fuller pentru seria PIB (cu intercept și trend)**

Prima parte a testului prezintă informaţii cu privire la tipul testului (AFD, variabilele exogene introduse – constantă, trend) şi cuprinde rezultatul testului, valorile critice pentru fiecare nivel de relevanţă (1, 5 şi 10 la sută), şi probabilitatea, p, asociată rezultatului testului.

Formularea ipotezelor

H0: *seria este nestaționară (are rădăcină unitară)*

H1: *seria este staționară*

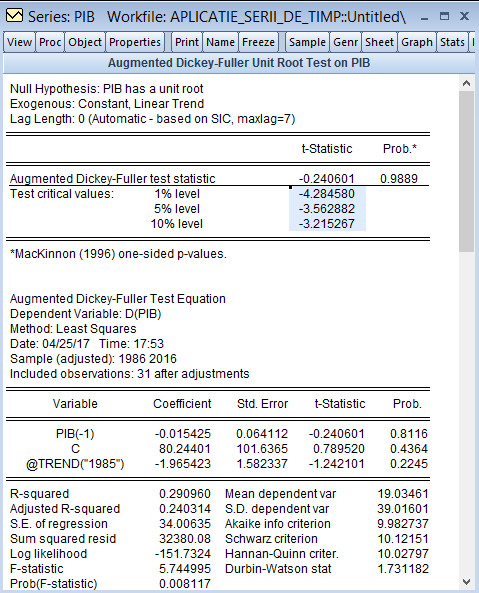


Figura 5 - Testul Dickey-Fuller (cu intercept și trend)

Așa cum se poate observa în *Figura 5,* valoarea calculată pentru t-statistic este –0.24 și probabilitatea asociată acesteia este de 0.9889.

Dacă valoarea testului este mai mare decât valoarea critică, nu este respinsă ipoteza nulă, deci seria are o rădăcină unitară (este nestaţionară).

Valoarea testului (-0.24) este mai mare decât valorile critice ale nivelurilor de semnificație: -4.28 pentru un nivel de semnificație de 1%, -3.56 pentru nivelul de 5%, -3.21 pentru nivelul de 10%. Astfel, acceptăm impoteza nulă, deci putem trage conluzia că seria de timp nu este staționară, deci are o rădăcină unitară.

Analizând și probabilitatea, a carei valori este 0.9889 >> 0.05 observăm de asemenea că seria este nestaționară.

Seria prezintă un trend stochastic şi este de tip DS(Difference-Stationarity).

* 1. **Testul Dickey-Fuller pentru seria PIB (intercept)**

Prima parte a testului prezintă informaţii cu privire la tipul testului (AFD, variabilele exogene introduse – constantă) şi cuprinde rezultatul testului, valorile critice pentru fiecare nivel de relevanţă (1, 5 şi 10 la sută), şi probabilitatea, p, asociată rezultatului testului.

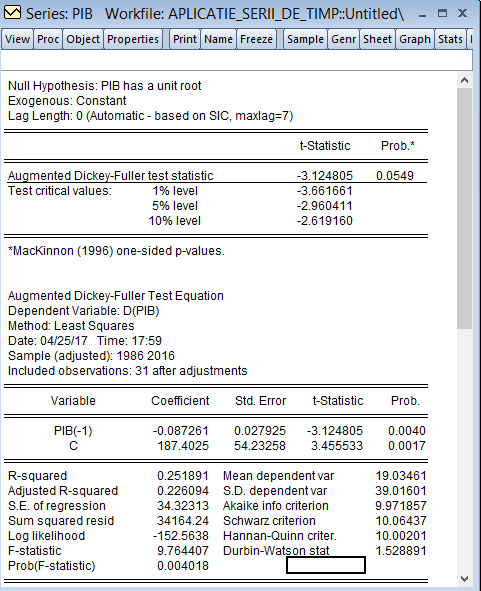


Figura 6 - Testul Dickey-Fuller (intercept)

Formularea ipotezelor

H0: *seria este nestaționară (are rădăcină unitară)*

H1: *seria este staționară*

Așa cum se poate observa în *Figura 6,* valoarea calculată pentru t-statistic este -3.12 și probabilitatea asociată acesteia este de 0.035.

Dacă valoarea testului este mai mare decât valoarea critică, nu este respinsă ipoteza nulă, deci seria are o rădăcină unitară (este nestaţionară).

Valoarea testului (-3.12) este mai mare decât valoarea critica a nivelului de semnificatie: -3.66 pentru un nivel de semnificație de 1%. Astfel, acceptam ipoteza nulă, deci putem trage conluzia că seria de timp nu este staționară, deci are o rădăcină unitară.

Observam faptul că seria nu este staționară și din analiza probabilității asociate testului, care are o valoare de 0.0549 > 0.05.

Seria prezintă un trend stochastic şi este de tip DS(Difference-Stationarity).

* 1. **Logaritmarea seriei PIB-ului**

Având în vedere faptul că s-a observat că seria nu este staționară o vom logaritma.

Logaritmarea seriei se face cu ajutorul functiei series logpib = log(pib)

Vom analiza și pentru această serie corelograma și testul Dickey-Fuller pentru a vedea dacă este staționară.

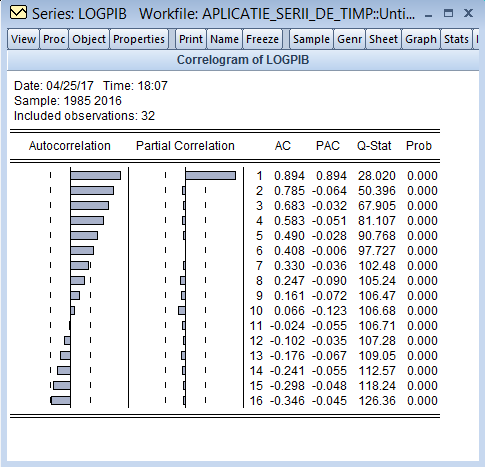


Figura 7 - Corelograma pentru seria PIB-ului logaritmată

Analizând corelograma putem spune că și în acest caz seria este nestaționară. Se vede faptul că valorile coeficientului de autocorelație și ale coeficientului de autocorelație parțială sunt apropiate de cele obținute în cazul seriei inițiale, valorile coeficientului de autocorelatie convergand din nou pentru primele laguri incet catre 0.

Efectuam și testul Dickey-Fuller pentru a testa staționaritatea.

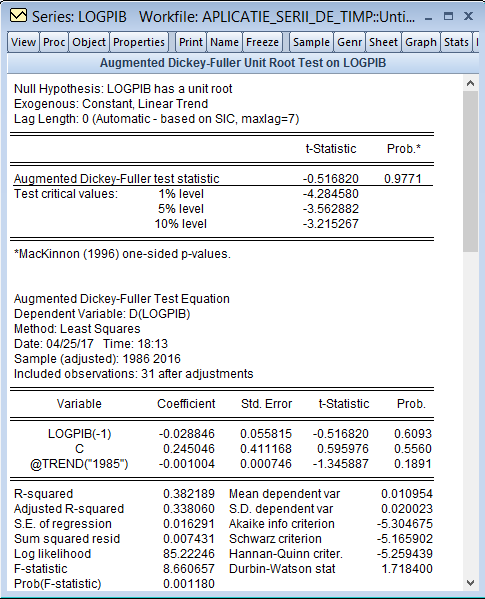


Figura 8 - Testul Dickey-Fuller pentru seria PIB după logaritmare

Așa cum se poate observa în *Figura 8,* valoarea calculată pentru t-statistic este –0.51 și probabilitatea asociată acesteia este de 0.9771.

Dacă valoarea testului este mai mare decât valoarea critică, nu este respinsă ipoteza nulă, deci seria are o rădăcină unitară (este nestaţionară).

Observam ca valoarea testului (-0.51) este mai mare decât valorile critice ale nivelurilor de semnificație: -4.284 pentru un nivel de semnificație de 1%, -3.56 pentru nivelul de 5% și -3.21 pentru nivelul de semnificație 10%. Astfel, acceptăm ipoteza nulă, deci putem trage conluzia că seria de timp nu este staționară, deci are o rădăcină unitară.

Se observa faptul că seria nu este staționară și din analiza probabilității asociate testului, care are o valoare de 0.9771 >> 0.05.

Seria prezintă un trend stochastic şi este de tip DS(Difference-Stationarity).

Deci nici dupa logaritmare, seria nu este staționară.

* 1. **Testul Dickey-Fuller pentru seria PIB (prima diferență)**

Prima parte a testului prezintă informaţii cu privire la tipul testului (AFD, variabilele exogene introduse – constantă) şi cuprinde rezultatul testului, valorile critice pentru fiecare nivel de relevanţă (1, 5 şi 10 la sută), şi probabilitatea, p, asociată rezultatului testului.

Vom face prima diferenta pe seria de date logaritmata pentru a vedea daca aceasta va prezenta stationaritate.

Formularea ipotezelor

H0: *seria este nestaționară (are rădăcină unitară)*

H1: *seria este staționară*

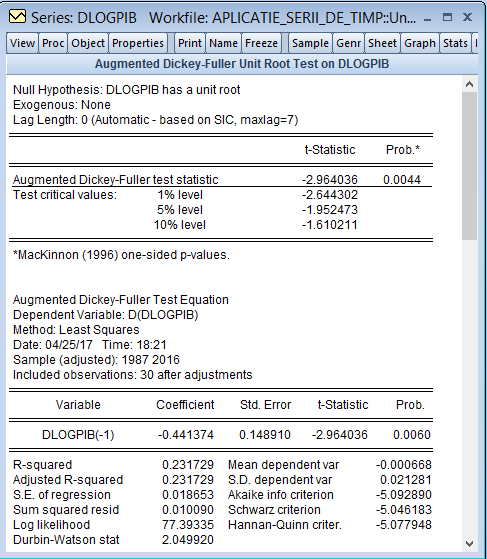


Figura 9 - Testul Dickey-Fuller pentru seria PIB-ului (prima diferență)

Așa cum se poate observa în *Figura 9,* valoarea calculată pentru t-statistic este –2.964 și probabilitatea asociată acesteia este de 0.0044.

Dacă valoarea testului este mai mare decât valoarea critică, nu este respinsă ipoteza nulă, deci seria are o rădăcină unitară (este nestaţionară).

Valoarea testului (-2.964) este mai mica decât valorile critice ale nivelurilor de semnificație: -2.644 pentru un nivel de semnificație de 1%, -1.95 pentru nivelul de 5%, -1.61 pentru nivelul de 10%. Astfel, se respinge ipoteza nulă, deci putem trage conluzia că seria de timp este staționară, deci nu are o rădăcină unitară.

De asemenea valoarea probabilității testului este de 0.0044 < 0.05, ceea ce ne indică încă o dată că seria este staționară.

Deci seria a devenit staționară dupa logaritmarea și aplicarea diferenței de ordinul intai.

* 1. **Corelograma seriei PIB diferențiată**

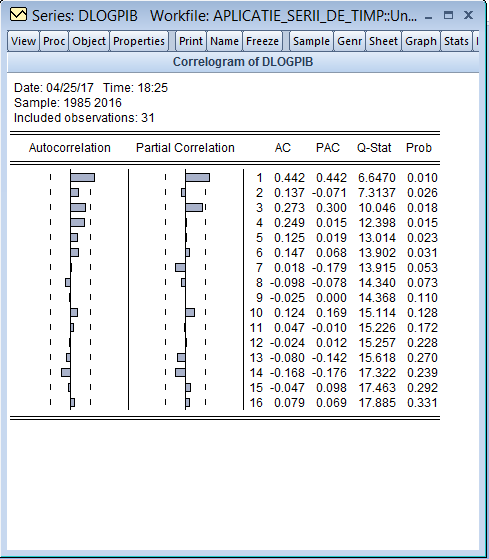


Figura 10 - Corelograma seriei PIB-ului dupa diferențierea de ordinul intai

În această corelogramă putem observa faptul că coeficienții de autocorelare și cei de autocorelare parțială au valori mai mici și se încadrează mult mai bine în intervale, iar si valorile pentru coeficientul de autocorelare scad foarte mult chiar dupa lagul 5 ceea ce ne indică faptul că seria este staționară.

* 1. **Modele ARIMA-ARIMA**

Vom folosi aceste modele pentru a identifica un model pentru seria PIB. Vom folosi seria logaritmată și diferențiată de ordinul intai deoarece aceasta prezintă staționaritate.

Pentru identificarea modelului vom analiza ARMA(p, q), unde p și q îi vom identifica cu ajutorul corelogramei din figura 11. p îl identificam uitându-ne la coeficientul de autocorelație parțială și identificând valorile care depășesc intervalul și sunt mai mari decât restul, iar pe q îl identificăm în același mod uitându-ne la coeficientul de autocorelație.

În final se va alege modelul care este valid, care are coeficienții semnificativi, care maximizează log likelihood și care minimizează valorile Akaike info criterion, Schwarz criterion și Hannan-Quinn criterion.

În cele ce urmează vom analiza ecuațiile aferente.

**1. AR(1)**

Ecuația va fi dlogpib c ar(1).

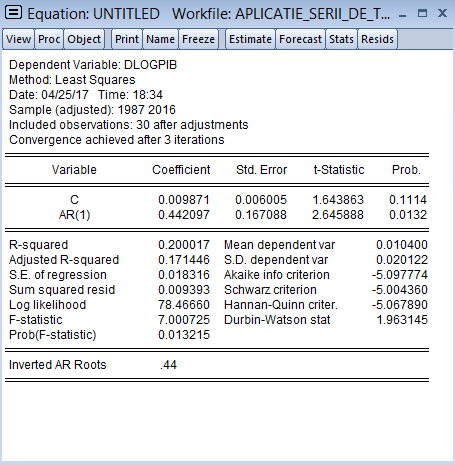


Figura 11 - AR(1)

**2. AR(2)**

Ecuația va fi dlogpib c ar(1) ar(2)

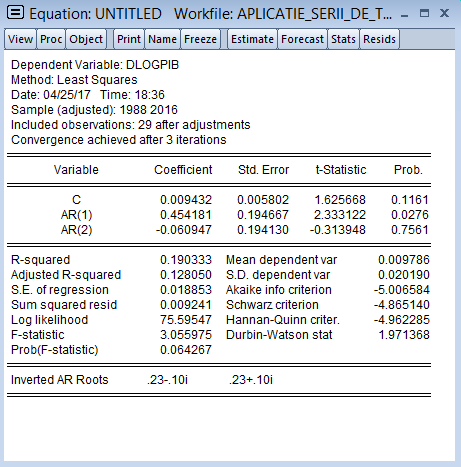


Figura 12 - AR(2)

**3. AR(3)**

Ecuația va fi dlogpib c ar(1) ar(2) ar(3)

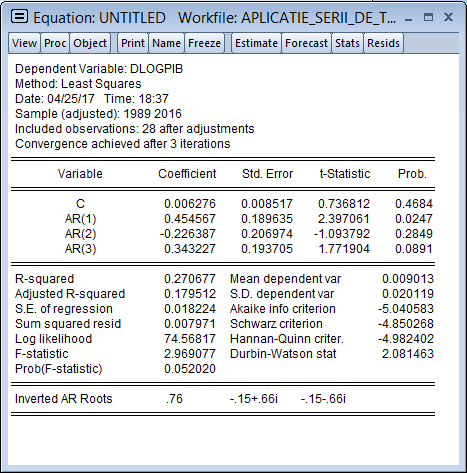


Figura 13 - AR(3)

**4. MA(1)**

Ecuația va fi dlogpib c ma(1)

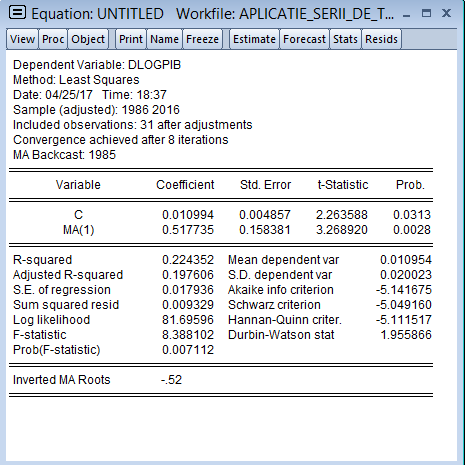


Figura 14 - MA(1)

**5. ARMA(1, 1)**

Ecuația va fi dlogpib c ar(1) ma(1)

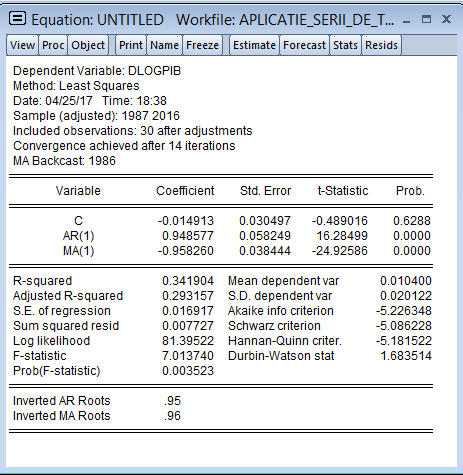


Figura 15 - ARMA(1,1)

**6. ARMA(2, 1)**

Ecuația va fi dlogpib c ar(1) ar(2) ma(1).

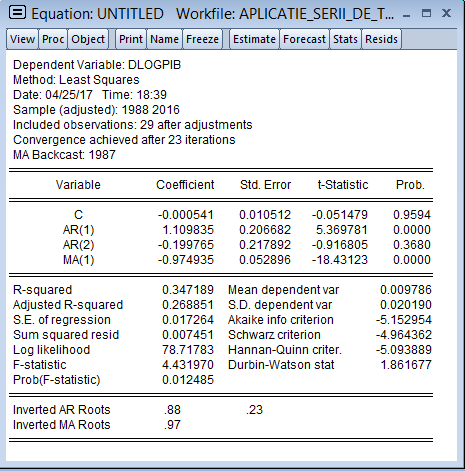


Figura 16 - ARMA(2,1)

**7. ARMA(3, 1)**

Ecuația va fi dlogpib c ar(1) ar(2) ar(3) ma(1).

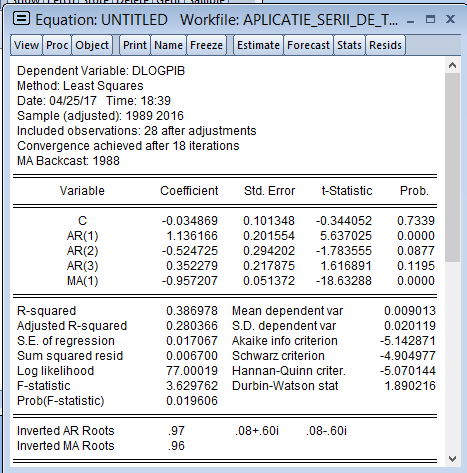


Figura 17 - ARMA(3,1)

Vom pune într-un tabel toate valorile care trebuie reținute pentru a putea identifica un model și anume: validitatea coeficienților, Akaike info criterion, Schwarz criterion, Hannan-Quinn criterion și validitatea modelului.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Coef. | ln L | AIC | SIC | HQIC | Prob. F |
| AR(1) | V | 78.46 | -5.09 | -5.00 | -5.06 | V |
| AR(2) | V | 75.59 | -5.00 | -4.86 | -4.96 | V |
| AR(3) | V | 74.56 | -5.04 | -4.85 | -4.98 | V |
| MA(1) | V | 81.69 | -5.14 | -5.04 | -5.11 | V |
| ARMA(1,1) | V | 81.39 | -5.22 | -5.08 | -5.18 | V |
| ARMA(2,1) | V | 78.71 | -5.15 | -4.96 | -5.09 | V |
| ARMA(3,1) | V | 77.00 | -5.14 | -4.90 | -5.07 | V |

După analiza tabelului de mai sus, luând în considerare criteriile menționate mai sus, am ajuns la concluzia că modelul potrivit pentru estimarea datelor este ARMA(1, 1), folosind drept criteriu minimizarea AIC.

* 1. **Predicție**

Vom realiza o predicție pentru seria de date pornind de la modelul identificat anterior(ARMA(1, 1)). Pentru ecuație luand in calcul ar(1) ar(2) ma(1) alegem Forecast și Static Forecast ca și metodă.

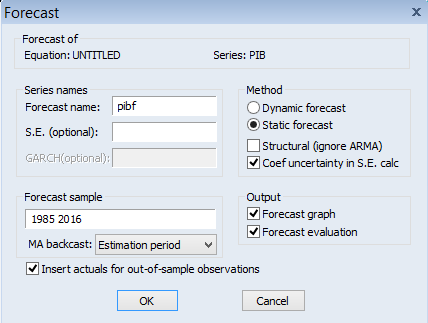


Figura 18 - Forecast

În *Figura 19 - Grafic previzionare* putem observa intervalul în care se va afla în viitor valoarea pib-ului( valoarea logaritmată și diferențiată de ordinul intai).

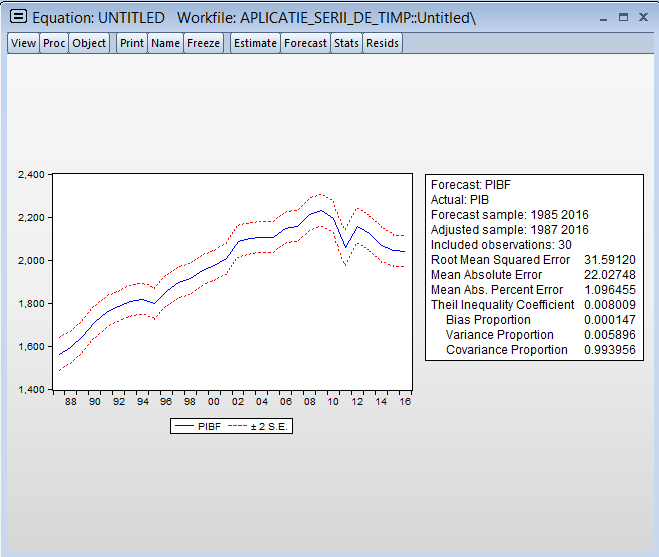


Figura 19 - Grafic previzionare

Coeficientul Bias Proportion are valoarea 0.0063, valoare care este foarte apropiată de 0, ne sugerează faptul că previziunea este una corectă.

De asemenea, valoarea lui Root Mean Squared Error este 0.01, o valoare destul de mică, acest indicator reprezentând abaterea standard a predicției.

Valoarea previzionata pentru anul 2016 a fost de 2041,611 pe cand valoarea actuala a PIB-ului in anul 2016 a fost de 2049,316.

1. **Aplicația 2 – Evolutia vanzarilor de bere in Marea Britanie**
   1. **Descrierea datelor**

Am ales această temă deoarece berea este unul dintre cele mai consumate alimente ce contine alcool. In ultima vreme consumul de bere a crescut, iar promovarea acesteia in media prin intermediul spoturilor publicitare, dar si a reclamelor online a devenit din ce in ce mai intensa.

Datele ce numarul de butoaie totaal vandut (1 butoi = 159 l) sunt inregistrate lunar, incepand cu 01.01.2009 pana in 31.01.2017

Conform lui Platon : “*Este un om inteligent cel care a inventat berea.*”

După cum se poate observa în *Figura 20 - Graficul seriei vanzarilor de bere* seria nu are un trend neted iar seria prezintă sezonalitate și nu este staționară deoarece media, varința și covarianța se schimbă în timp.

Seria este sezonieră, datele fiind prezentate pe luni. Graficul ne arată de asemenea faptul că în fiecare an se produce o creștere treptată, la sfarșitul anului producându-se o scădere bruscă, rapidă. Astfel acest fenomen poate fi determinat de la an la an, el repretându-se.

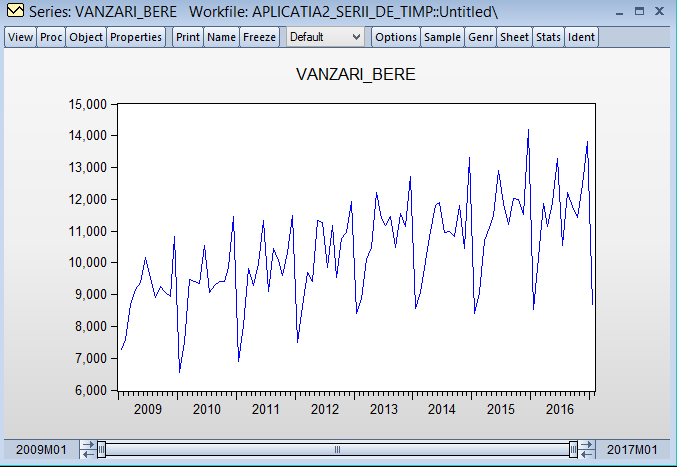


Figura 20- Graficul seriei vanzarilor de bere

* 1. **Ajustarea sezonieră a seriilor de timp**

Vom desezonaliza seria (*Figura 21 - Serie desezonalizată*) prin ajustarea acesteia cu ajutorul Proc – Seasonal Adjustment – Moving Average Methods – Additive (*Figura 22 - Seasonal Adjustment*).

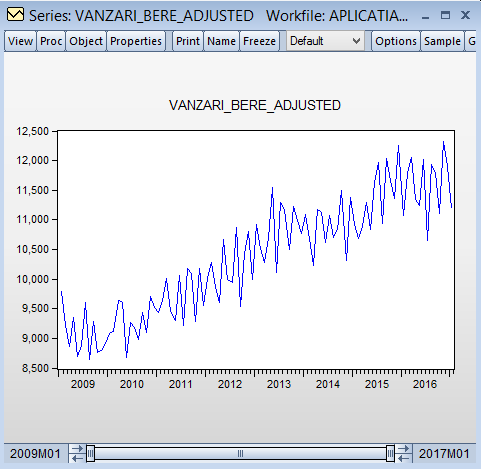


Figura 21 - Serie desezonalizată

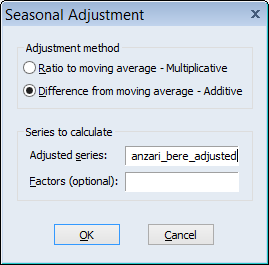


Figura 22 - Seasonal Adjustment

Estimam ecuația desezonalizată:

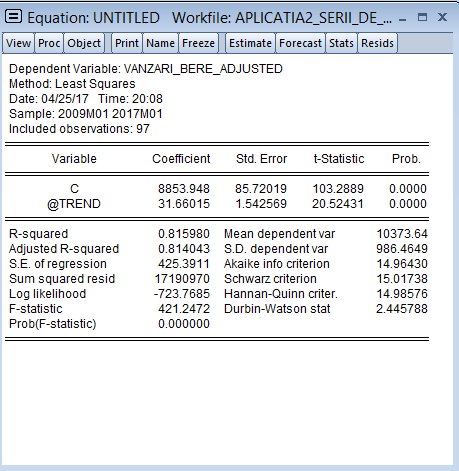


Figura 23 - Estimarea ecuației desezonalizată

Din *Figura 23 - Estimarea ecuației desezonalizată* observăm că numarul de butoaie de bere vandute creste de la o luna la alta cu aproximativ 31 de butoaie.

1. **Aplicația 3 – PIB si exportul de bunuri si servicii. Cointegrare. Cauzalitate Granger**

Cum seria PIB a trecut deja prin etapele necesare ca aceasta sa fie stationara iar trendul a fost eliminat – in cadrul aplicatiei 1, acum vom analiza stationaritatea si trendul si pentru seria exportului de bunuri si servicii.

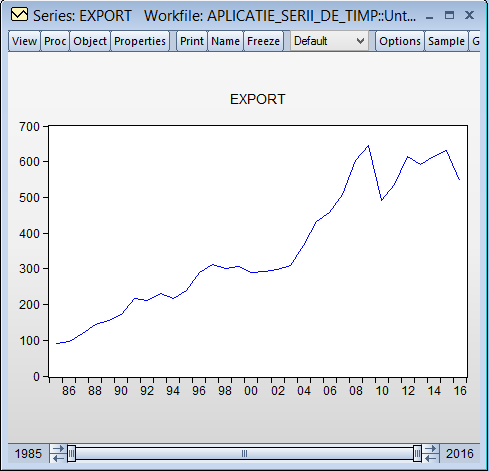
****

Figura 24 - Evolutia exportului de bunuri si servicii

Pentru a determina daca seria este stationara si daca are trend vom analiza mai intai corelograma:

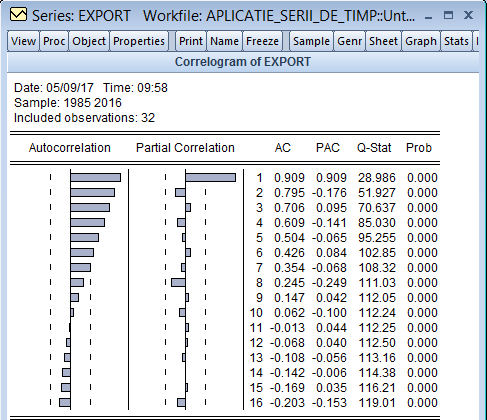
****

Figura 25 - Corelograma exportului de bunuri si servicii

Analiza corelogramei releva faptul ca seria este nestationara, autocorelatiile apropiindu-se rapid de 0 atunci cand k creste.

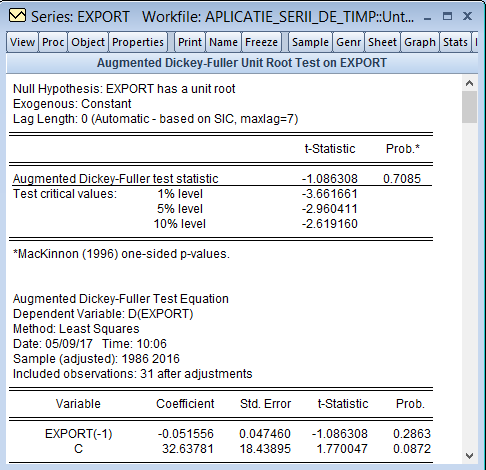
****

Figura 26 - Testul ADF pentru seria exportului

Probabilitatea asociata testului este de 0.70, mai mare decat toate pragurile de semnificatie de 1%, 5%, 10%. De asemenea, valoarea asociata testului t, -1.086 este mai mare decat valorile testului t corespunzatoare fiecarui prag de semnificatie. De aceea rezulta ca **seria este nestationara.**

Vom aplica: series dlogexport = dlog(export,1)

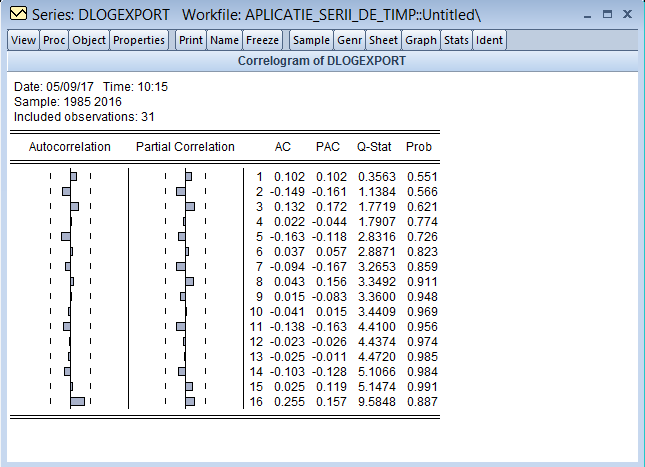


Figura 27 - Corelograma pentru seria exportului logaritmata si diferentiata de ordinul 1

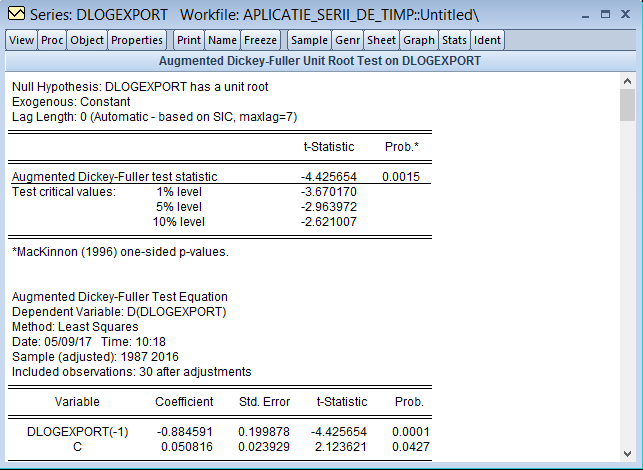


Figura 28 - Testul ADF

Conform corelogramei, dar si testului ADF de unde reiese ca probabilitatea asociata testului este de 0.0015 care este mult mai mica chiar si decat pragul de 1% reiese ca se respinge ipoteza nula conform careia seria este stationara si se accepta ipoteza alternativa, adica **seria este stationara.**

Asadar, pentru analiza cauzalitatii Granger si a cointegrarii celor doua serii mai intai vom determina numarul optim de laguri:

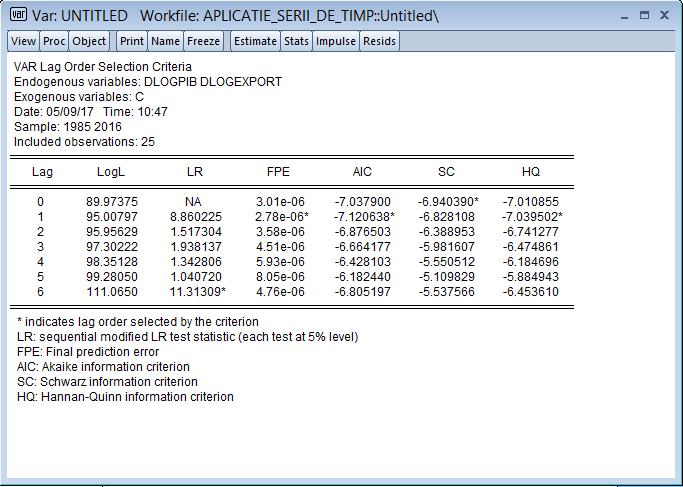


Figura 29 - Determinarea numarului optim de laguri

Conform rezultatelor testului, reiese ca lagul optim este lagul 1.

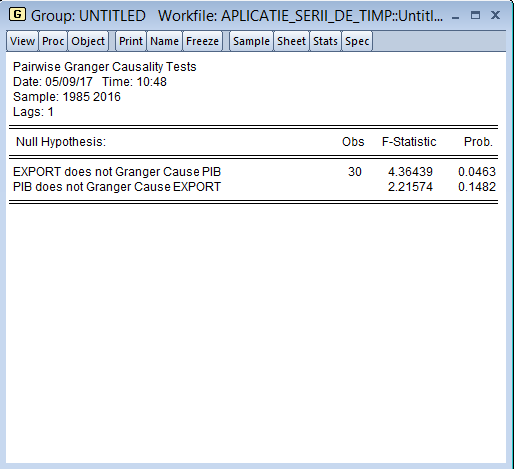


Figura 30 - Cauzalitatea Granger

Conform testului reiese ca exista cauzalitate intre cele doua serii. Deoarece probabiltatea de 0.046 < 0.05 se respinge ipoteza nula conform careia exportul de bunuri si servicii nu cauzeaza granger PIB.

Asadar, modificari in valoarea exportului de bunuri si servicii au impact asupra evolutiei PIB.

**Cointegrare – Testul Johansen**

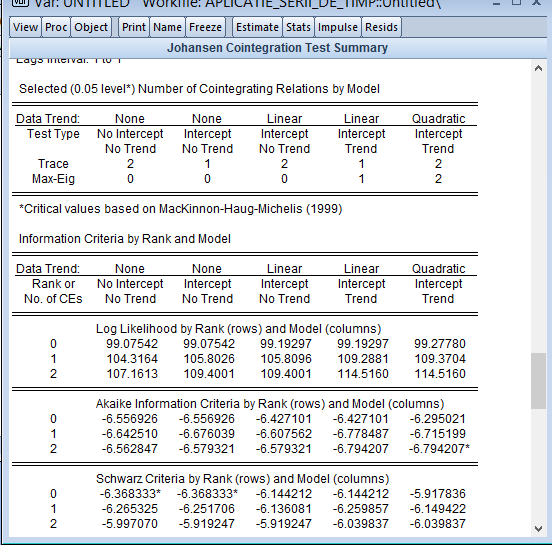
****

Figura 31 - Testul Johansen pentru cointegrare

Conform outputului testului Johansen reiese ca seriile sunt cointegrate.

**Bibliografie**

1. [http://data.worldbank.org](http://data.worldbank.org/indicator/SP.URB.TOTL?page=6) / 24.04.2017
2. [http://www.beerandpub.com/statistics / 26.04.2017](http://www.beerandpub.com/statistics%20/%2026.04.2017)
3. <https://www.statista.com> / 24.04.2017
4. <https://data.oecd.org> / 25.04.2017