

2_3.

Aplicație Metodologia Box Jenkins

Analiza staționarității, staționarizarea seriei, identificarea unor modele ARMA, estimarea parametrilor, teste de validitate, alegerea celui mai performant model, efectuarea previziunilor.

Aplicație Să se modeleze seria Rata_Somajului utilizând **Metodologia Box Jenkins**

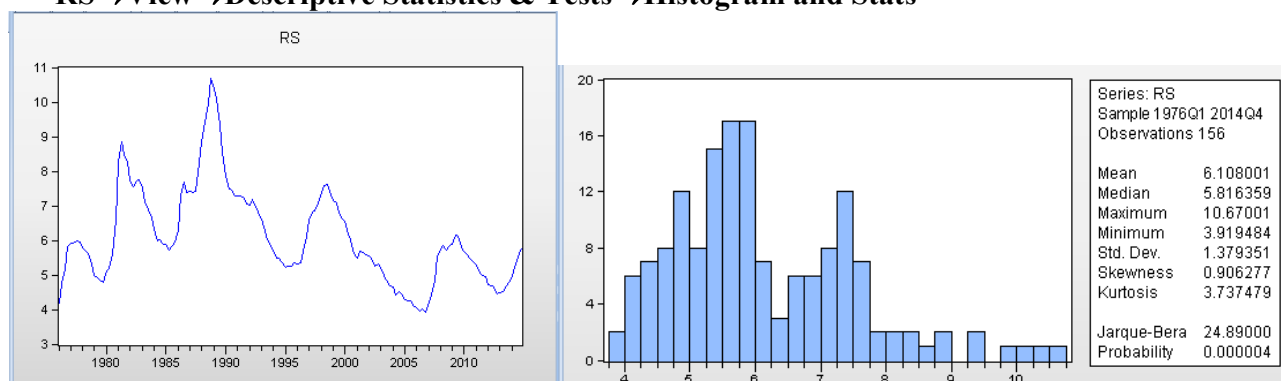
Seria considerată este Rata_Somajului (date trimestriale) pentru perioada 1976:Q1-2014:Q4. (160 obs). Rata șomajului este exprimată procentual(%). Datele sunt în fișierul "RATA_SOMAJ.WF1".

Cerințe:

- Să se analizeze caracteristicile seriei folosind reprezentări grafice și indicatori descriptivi adecvați.
- Să se estimeze, folosind abordarea Box-Jenkins, parametrii unor modele autoregresive adecvate.
- Se se efectueze previziuni pe baza modelului
- Să se interpreteze rezultatele obținute din punct de vedere economic.

Se analizează caracteristicile seriei folosind grafice și indicatori descriptivi.

RS→View→Descriptive Statistics & Tests→Histogram and Stats

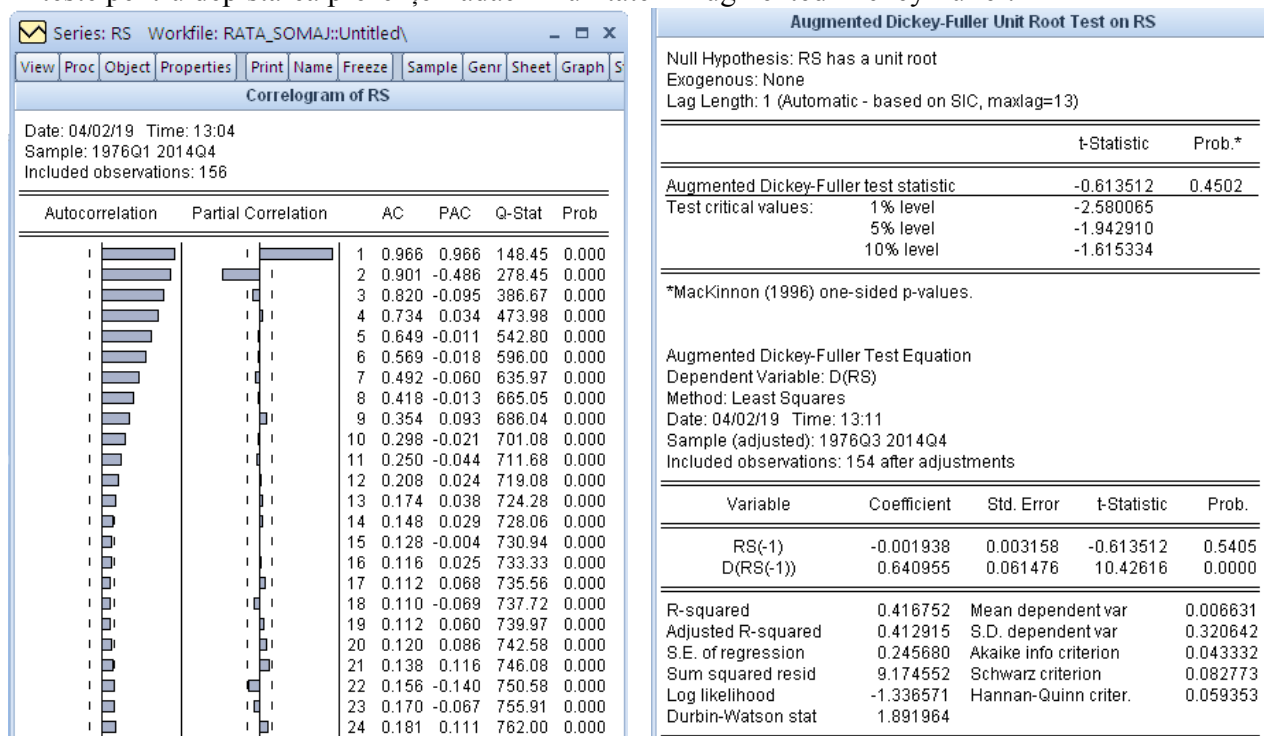


b)

Etapa0. Verificarea staționarității seriei și staționarizare.

Analiza staționarității seriei de date se realizează în Eviews prin:

- analiza corelogramei seriei de date
- teste pentru depistarea prezenței rădăcinii unitate – Augmented Dickey Fuller.



Ne uităm la coloanele AC, PAC, Q-stat și Prob. Din analiza corelogramei deducem că seria **RS** este nestaționară, deoarece funcția de autocorelație descrește foarte lent. Limitele din grafic aproximează două erori standard. Valoarea la lag-ul 5 este foarte mare (0,649).

Testul ADF (Augmented Dickey-Fuller Unit Root Test)

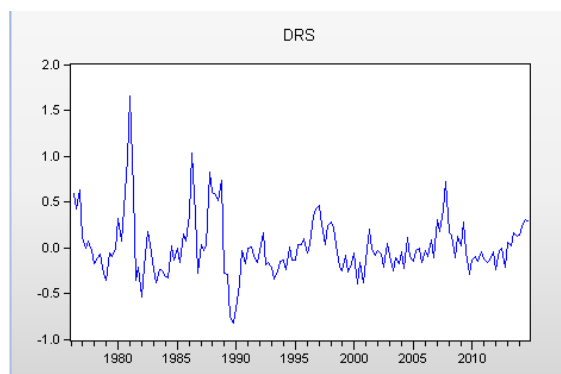
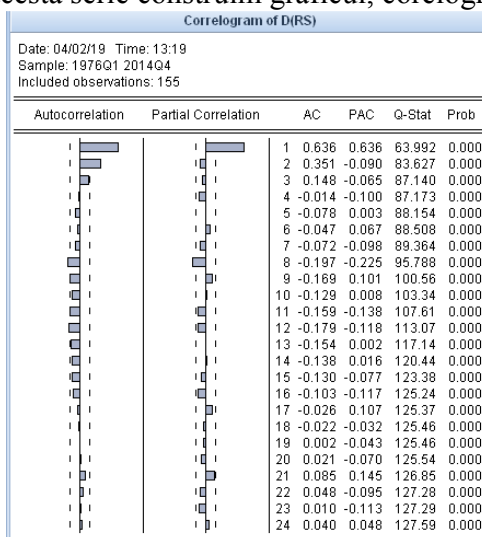
H_0 : seria RS are rădăcină unitară și este nestaționară

H_1 : seria RS este staționară

Valoarea calculată pentru t-Statistic este $-0,613512$, care este mai mare decât valorile critice ($-1,94291$ pentru 5%). Prob=0,4502 > 0,05.

Rezultă $t_{calc} \neq R_C$. Acceptăm H_0 , aceea că există o rădăcină unitară, deci **seria RS este nestaționară**.

Diferențiem seria inițială RS. Seria diferențelor de ordinul 1 este: $D(RS) = \Delta RS_t = RS_t - RS_{t-1}$
Pentru această serie construim graficul, corelograma și aplicăm Testul ADF.



Din corelograma seriei D(RS) deducem că seria **D(RS)** este staționară, deoarece funcția de autocorelație descrește rapid spre zero.

Testul ADF (Augmented Dickey-Fuller Unit Root Test)

H_0 : seria **D(RS)** are rădăcină unitară și este nestaționară

H_1 : seria **D(RS)** este staționară

Augmented Dickey-Fuller Unit Root Test on D(RS)		
Null Hypothesis: D(RS) has a unit root		
Exogenous: None		
Lag Length: 0 (Automatic - based on SIC, maxlag=13)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-5.887228	0.0000
Test critical values: 1% level	-2.580065	
5% level	-1.942910	
10% level	-1.615334	

Valoarea calculată pentru t-Statistic este $-5,887228$, care este mai mică decât valorile critice.

Prob=0,0000 < 0,05.

Rezultă $t_{calc} \in R_C$. Respingem H_0 . Acceptăm H_1 , deci seria diferențiată, **D(RS)**, este o serie staționară.

Deoarece seria nestaționară RS a devenit staționară după diferențierea de ordin 1, putem concluziona că este integrată de ordin d=1. Simbolizăm acest lucru scriind $RS \sim I(1)$.

Drept urmare, vom aplica procedura Box-Jenkins asupra seriei de date staționarizată prin diferențele de ordin 1 (seria fiind integrată de ordin 1) și vom determina procesul ARMA(p,q) corespunzător.

Etapa1. Identificarea.

Corelograma seriei diferențiate ne va permite să alegem valorile p și q potrivite pentru seria de date.

Ordinul p al părții AR este dat de coeficienții de autocorelație parțială (coloana PAC) statistic semnificativi.

Ordinul q al părții MA este dat de coeficienții de autocorelație (coloana AC) statistic semnificativi (coeficienții semnificativi se află în afara intervalului $(\pm 1,96/\sqrt{T})$).

Etapa2. Estimarea (Estimarea parametrilor modelelelor AR, MA, ARMA identificate pe baza ACF/PACF.)

Vom estima mai multe modele de tip ARMA(p,q) pentru seria diferențiată D(RS), care este staționară. La fiecare model estimat notăm dacă parametrii sunt semnificativ diferiți de zero, valorile pentru R-squared și pentru Criteriile informaționale Akaike și Schwarz. Am listat output-ul pentru modelul ARMA(2,1) și corelograma reziduurilor.

În EViews am specificat ecuația de estimat: **D(RS) C AR(1) AR(2) MA(1)**

Vom nota variabila D(RS) cu Y. Modelul de estimat va fi:

$$Y_t = \phi_0 + \phi_1 \cdot Y_{t-1} + \phi_2 \cdot Y_{t-2} + \theta_1 \cdot \varepsilon_{t-1} + \varepsilon_t$$

Dependent Variable: D(RS) Method: Least Squares Date: 04/02/19 Time: 13:28 Sample (adjusted): 1976Q4 2014Q4 Included observations: 153 after adjustments Convergence achieved after 10 iterations MA Backcast: 1976Q3					Correlogram of REZIDUURI2_1						
					Date: 04/03/19 Time: 15:12 Sample: 1976Q1 2014Q4 Included observations: 153						
Variable	Coefficient	Std. Error	t-Statistic	Prob.	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
C	-0.013748	0.008386	-1.639364	0.1032			1 0.017 0.017 0.0462 0.830				
AR(1)	1.601042	0.061841	25.88949	0.0000			2 -0.030 -0.030 0.1865 0.911				
AR(2)	-0.659034	0.060630	-10.86976	0.0000			3 0.043 0.044 0.4819 0.923				
MA(1)	-0.989455	0.008338	-118.6749	0.0000			4 -0.064 -0.066 1.1263 0.890				
							5 -0.084 -0.079 2.2459 0.814				
							6 0.108 0.107 4.1279 0.659				
							7 0.151 0.151 7.8512 0.346				
							8 -0.184 -0.190 13.405 0.099				
							9 -0.025 -0.033 13.505 0.141				
							10 0.090 0.094 14.857 0.137				
							11 -0.019 0.031 14.917 0.186				
							12 -0.096 -0.124 16.458 0.171				
							13 -0.011 -0.084 16.478 0.224				
							14 -0.019 0.022 16.537 0.282				
							15 -0.073 0.017 17.452 0.293				
							16 -0.078 -0.159 18.516 0.295				
							17 0.057 -0.000 19.089 0.323				
							18 -0.065 -0.001 19.831 0.342				
							19 -0.051 -0.000 20.295 0.377				
							20 -0.028 -0.110 20.437 0.431				
							21 0.167 0.175 25.470 0.227				
							22 -0.013 0.046 25.499 0.274				
							23 -0.061 -0.082 26.182 0.292				
							24 0.085 -0.004 27.495 0.282				
R-squared	0.449811	Mean dependent var		0.003914							
Adjusted R-squared	0.438733	S.D. dependent var		0.319911							
S.E. of regression	0.239670	Akaike info criterion		0.006692							
Sum squared resid	8.558846	Schwarz criterion		0.085919							
Log likelihood	3.488059	Hannan-Quinn criter.		0.038875							
F-statistic	40.60528	Durbin-Watson stat		1.954821							
Prob(F-statistic)	0.000000										
Inverted AR Roots	.80-.13i	.80+.13i									
Inverted MA Roots	.99										

Modelul estimat este:

$$\hat{Y}_t = -0,013748 + 1,601042 \cdot Y_{t-1} - 0,659034 \cdot Y_{t-2} - 0,989455 \cdot \hat{\varepsilon}_{t-1}$$

Etapa3. Diagnosticarea modelelor și alegerea celui mai bun model.

3.1. Testarea modelelor

a)– Parametrii modelului sunt statistic semnificativi ?

(constanta nu este semnificativă deoarece Prob=0,1032>0,05; ceilalți coeficienți sunt semnificativ diferiți de zero deoarece Prob=0,0000<0,05)

R-squared=0,4498

b)– Reziduurile obținute sunt zgomot alb?

Aplicăm teste asupra reziduurilor obținute din modelul estimat.

Din corelograma reziduurilor: Probabilitățile asociate coeficienților Ljung-Box (coloana Q-Stat) sunt >0,05 și indică independența erorilor. Nu există corelație serială în reziduuri.

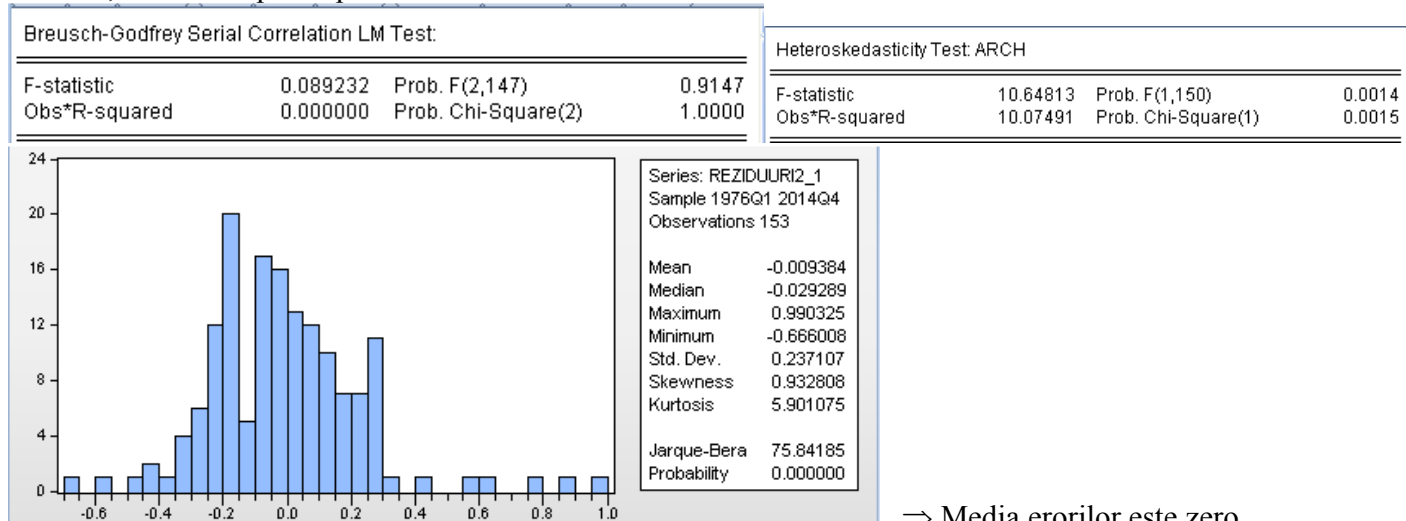
DW=1,9548 arată că nu există Autocorelarea erorilor aleatoare de ordin 1 (testul Durbin-Watson).

Testul Multiplicatorului lui Lagrange (Breuch-Godfrey Serial Correlation Test) ne arată același lucru.

H_0 : Nu există Autocorelarea erorilor aleatoare de ordin 2

H_1 : Există Autocorelarea erorilor aleatoare de ordin 2

Prob >0,05 \Rightarrow Acceptăm ipoteza nulă \Rightarrow Nu există Autocorelarea erorilor aleatoare de ordin 2



Reziduurile nu sunt normal distribuite, deoarece valoarea testului JB este mult mai mare decât valoarea tabelată (5,99) și în plus Prob tinde către zero.

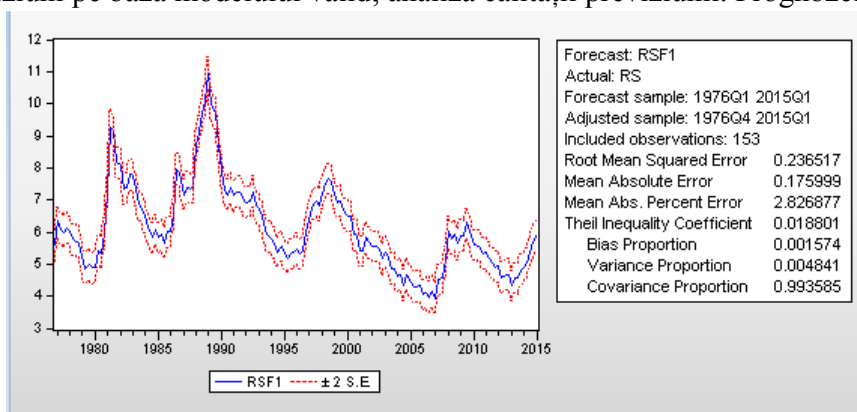
Ipoteza de homoscedasticitate se testează cu ajutorul Corelogramei patratelor reziduurilor și cu ajutorul testului ARCH LM. Din Heteroskedasticity Test ARCH \Rightarrow Erorile sunt heteroscedastice.

3.2. Alegerea celui mai bun model (R^2 , F, valorile la Criteriul Akaike și Criteriul Schwartz să fie minime)

Modele	Criteriul Akaike	Criteriul Schwartz
ARMA(1,1)	0,0518	0,111
ARMA(1,2)	0,0648	0,1437
ARMA(1,3)	0,058	0,1567
ARMA(2,1)	0,0067	0,0859
ARMA(3,1)	0,067	0,166

Etapa4. Efectuarea de prognoze

Previziuni pe baza modelului valid; analiza calității previziunii. Prognozele se vor realiza pentru anul 2015.



Previziunea statică-utilizează în formulele de recurență valorile observate pe perioada de previziune.

Evaluarea bonității previziunii este dată de indicatorii:

- **Mean Abs. Percent Error** Acesta trebuie să aibă o valoare cât mai mică (aici e 2,82);
- **Coeficientul Theil** cu valori în (0,1); valori apropiate de 0 indică o ajustare bună;

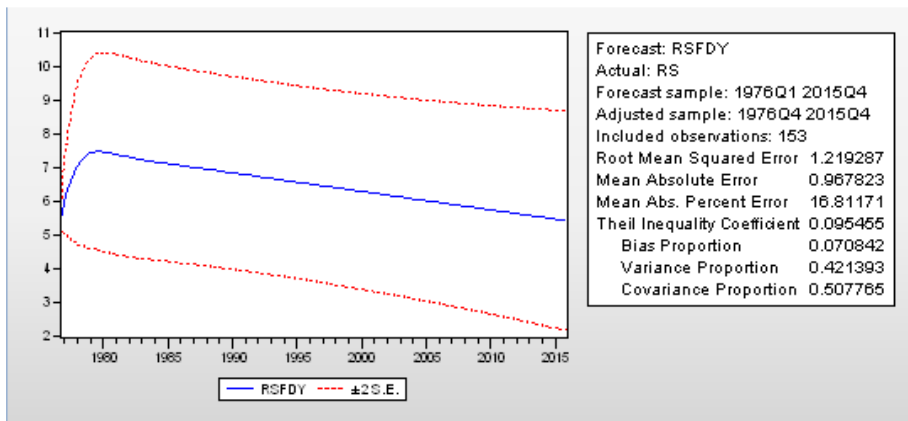
– Grupul de indicatori:

Bias Proportion trebuie să fie cât mai mic ;

Variance Proportion trebuie să fie cât mai mic;

Covariance Proportion trebuie să fie cât mai mare (suma lor este 1).

(Aici pare a indica o legătură foarte bună între valori și previziuni.)



Previziunea dinamică-utilizează în formulele de recurență valorile previzionate anterior