Analiza și Prognoza seriilor de timp

Seminar 3: Modele ARIMA

Daniel Traian PELE

Academia de Studii Economice din București
IDA Institute Digital Assets
Blockchain Research Center
AI4EFin Artificial Intelligence for Energy Finance
Academia Română, Institutul de Prognoză Economică
MSCA Digital Finance

Cuprins Seminar

Activitățile de Astăzi:

1. **Test de Recapitulare** — Verificarea înțelegerii conceptelor ARIMA
2. **Întrebări Adevărat/Fals** — Verificări conceptuale
3. **Probleme Practice** — Calcule cu ARIMA
4. **Exemple Rezolvate** — Aplicații din lumea reală
5. **Analiză pe Date Reale** — Studiu de caz PIB
6. **Exerciții AI** — Modelare om vs. AI

Test 1: Ordinul de Integrare

Întrebare

O serie de timp Y_t necesită două diferențe pentru a deveni staționară. Care este ordinul ei de integrare?

Variante de răspuns

(A) $I(0)$ indent **(B)** $I(1)$ indent **(C)** $I(2)$ indent **(D)** Nu poate fi determinat

Test 1: Ordinul de Integrare

Întrebare

O serie de timp Y_t necesită două diferențe pentru a deveni staționară. Care este ordinul ei de integrare?

Răspuns: C – $I(2)$

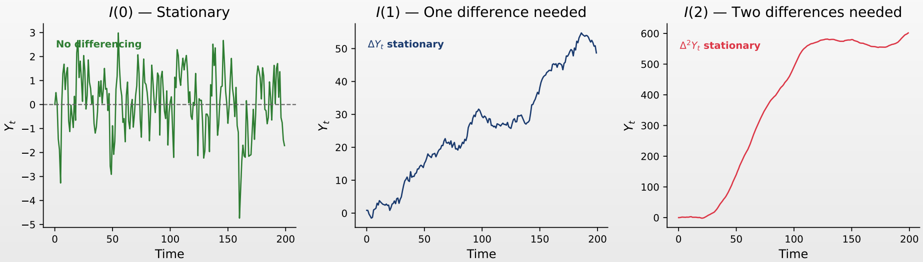
Definiție: $Y_t \sim I(d)$ dacă $\Delta^d Y_t$ este staționară dar $\Delta^{d-1} Y_t$ nu este.

Exemplu: Dacă Y_t urmează $\Delta^2 Y_t = \varepsilon_t$, atunci:

- $\Delta Y_t = \Delta Y_{t-1} + \varepsilon_t$ (încă are rădăcină unitară)
- $\Delta^2 Y_t = \varepsilon_t$ (zgomot alb, staționară)

Lumea reală: Indicii de prețuri pot fi $I(2)$ când inflația însăși este nestaționară.

Vizual: Procese Integrate



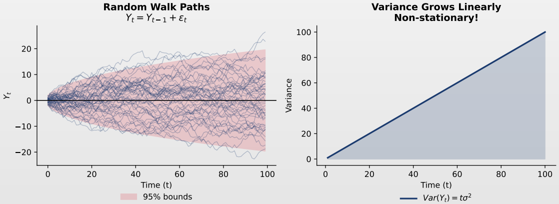
$I(0)$: staționară. $I(1)$: o diferență necesară. $I(2)$: două diferențe necesare pentru a deveni staționară.

TSA_ch3_def_integrated

Test 2: Proprietățile Mersului Aleatoriu

Întrebare

Pentru un mers aleator $Y_t = Y_{t-1} + \varepsilon_t$ cu $\text{Var}(\varepsilon_t) = \sigma^2$, care este $\text{Var}(Y_t)$?



TSA_ch3_rw_variance

Test 3: Specificarea Testului ADF

Întrebare

Când aplicați testul ADF pe datele PIB (care prezintă un trend ascendent clar), ce specificare trebuie folosită?

Variante de răspuns

(A) Fără constantă, fără trend

(B) Cu constantă, fără trend

(C) Cu constantă și trend

(D) Nu contează specificarea

Test 3: Specificarea Testului ADF

Întrebare

Când aplicați testul ADF pe datele PIB (care prezintă un trend ascendent clar), ce specificare trebuie folosită?

Răspuns: C – Cu constantă și trend

Regresia ADF cu trend: $\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \sum \delta_j \Delta Y_{t-j} + \varepsilon_t$

Regula practică:

- **Fără constantă**: serii cu medie zero (rar utilizat)
- **Cu constantă**: serii cu medie nenulă dar fără trend vizibil
- **Cu constantă + trend**: serii cu trend determinist vizibil (PIB, prețuri)

Atenție: Specificarea greșită reduce puterea testului!

Test 4: Notația ARIMA

Întrebare

Ce înseamnă ARIMA(2,1,1)?

Variante de răspuns

(A) AR(2) pe date diferențiate cu erori MA(1)

(B) AR(1) cu 2 diferențe și MA(1)

(C) MA(2) cu 1 diferență și AR(1)

(D) 2 lag-uri, 1 trend, 1 componentă sezonieră

Test 4: Notația ARIMA

Întrebare

Ce înseamnă ARIMA(2,1,1)?

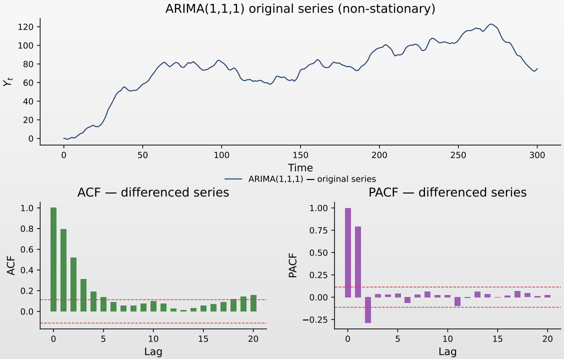
Răspuns: A – AR(2) pe date diferențiate cu erori MA(1)

ARIMA(p, d, q): $\phi(L)(1-L)^d Y_t = \theta(L)\varepsilon_t$
ARIMA(2,1,1) se expandează la:

$$(1-\phi_1 L-\phi_2 L^2)(1-L)Y_t = (1+\theta_1 L)\varepsilon_t$$

Sau echivalent: $(1-\phi_1 L-\phi_2 L^2)\Delta Y_t = (1+\theta_1 L)\varepsilon_t$
Interpretare: Mai întâi diferențiem seria, apoi ajustăm ARMA(2,1) pe ΔY_t.

Vizual: Procesul ARIMA



Sus: seria ARIMA originală. Jos: după diferențiere, folosim ACF/PACF pentru a identifica ordinele AR și MA.

TSA_ch3_def_arima

Test 5: Echivalența ARIMA

Întrebare

Modelul ARIMA(0,1,1) fără constantă, $(1 - L)Y_t = (1 + \theta L)\varepsilon_t$, este echivalent cu:

Variante de răspuns

(A) O netezire exponențială simplă (SES)

(B) Un model AR(1) staționar

(C) Un mers aleatoriu pur

(D) Un model MA(1) staționar

Test 5: Echivalența ARIMA

Întrebare

Modelul ARIMA(0,1,1) fără constantă, $(1 - L)Y_t = (1 + \theta L)\varepsilon_t$, este echivalent cu:

Răspuns: A – O netezire exponențială simplă (SES)

ARIMA(0,1,1): $Y_t = Y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$

SES: $\hat{Y}_{t+1} = \alpha Y_t + (1 - \alpha)\hat{Y}_t$ cu $\alpha = 1 + \theta$

- Când $\theta = 0$: mers aleatoriu pur (naive)
- Când $-1 < \theta < 0$: netezire $(0 < \alpha < 1)$
- Legătura fundamentală între abordarea stochastică și cea deterministă

Concluzie: SES este cazul optim al unui ARIMA(0,1,1)!

Test 6: Matricea de Decizie ADF + KPSS

Întrebare

ADF nu respinge H_0 (p = 0.15) și KPSS nu respinge H_0 (p = 0.08). Ce concluzie rezultă?

Variante de răspuns

(A) Seria este staționară

(B) Seria are o rădăcină unitară

(C) Rezultatele sunt neconcludente — putere statistică insuficientă

(D) Ambele teste sunt greșite

TSA_ch3_adf_kpss

Test 6: Matricea de Decizie ADF + KPSS

Întrebare

ADF nu respinge H_0 ($p = 0.15$) și KPSS nu respinge H_0 ($p = 0.08$). Ce concluzie rezultă?

Răspuns: C – Rezultate neconcludente

	ADF nu resp.	**ADF resp.**
KPSS nu resp.	**Neconcludent**	Staționară
KPSS resp.	Rădăcină unitară	**Neconcludent**

Soluții: Eșantion mai mare, teste PP sau ERS, sau procedura secvențială — diferențiați și retestați.

TSA_ch3_adf_kpss

Test 7: Supradiferențierea

Întrebare

Dacă $Y_t \sim I(1)$ și calculăm $\Delta^2 Y_t$, ce se întâmplă?

Variante de răspuns

(A) Obținem o serie staționară mai bună

(B) Introducem autocorelație negativă artificială

(C) Varianța scade

(D) Nu se schimbă nimic

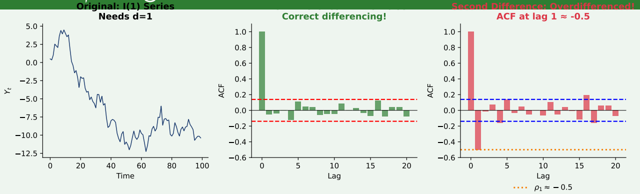
TSA_ch3_overdifferencing

Test 7: Supradiferențierea

Întrebare

Dacă $Y_t \sim I(1)$ și calculăm $\Delta^2 Y_t$, ce se întâmplă?

Răspuns: B – Autocorelație negativă artificială



Diagnostic: ACF la lag 1 ≈ -0.5 semnalează supradiferențiere. Reduceți d cu 1!

TSA_ch3_overdifferencing

Test 8: Varianța Prognozei

Întrebare

Pentru un model ARIMA(0,1,0) (mers aleator), cum se comportă varianța prognozei când orizontul h crește?

Variante de răspuns

(A) Rămâne constantă indent **(B)** Scade la zero indent **(C)** Crește liniar cu h indent **(D)** Convergă la o limită finită

Test 8: Varianța Prognozei

Întrebare

Pentru un model ARIMA(0,1,0) (mers aleator), cum se comportă varianța prognozei când orizontul h crește?

Răspuns: C – Crește liniar cu h

Prognoza mersului aleatoriu: $\hat{Y}_{T+h|T} = Y_T$ (cea mai bună prognoză este valoarea curentă)
Eroarea de prognoză: $Y_{T+h} - \hat{Y}_{T+h|T} = \sum_{i=1}^{h} \varepsilon_{T+i}$
Varianță:

$$\text{Var}(Y_{T+h} - \hat{Y}_{T+h|T}) = h\sigma^2$$

IC 95%: $Y_T \pm 1.96\sqrt{h}\sigma$ (se lărgește cu $\sqrt{h}$)

Test 9: Puterea Testului de Rădăcină Unitară

Întrebare

Testul ADF are putere scăzută când:

Variante de răspuns

(A) Dimensiunea eșantionului este foarte mare

(B) Rădăcina adevărată este aproape dar nu egală cu 1

(C) Seria nu are trend

(D) Seria este clar staționară

Test 9: Puterea Testului de Rădăcină Unitară

Întrebare

Testul ADF are putere scăzută când:

Răspuns: B – Rădăcina aproape dar nu egală cu 1

Exemplu: AR(1) cu $\phi = 0.95$ vs mers aleator ($\phi = 1$)
Problemă: Ambele au modele ACF similare (descreștere lentă), dar una este staționară!
Putere scăzută înseamnă: Probabilitate mare de eroare de tip II (eșec în respingerea lui H_0 fals)
Soluții:

- Dimensiuni mai mari ale eșantionului
- Testul Phillips-Perron (robust la heteroscedasticitate)
- Teste de rădăcină unitară pe paneluri (serii multiple)

Test 10: Selecția modelului ARIMA

Întrebare

După o diferențiere, ACF arată un vârf doar la lag 1, și PACF descrește. Modelul potrivit este:

Variante de răspuns

(A) ARIMA(1,1,0) indent **(B)** ARIMA(0,1,1) indent **(C)** ARIMA(1,1,1) indent **(D)** ARIMA(0,2,1)

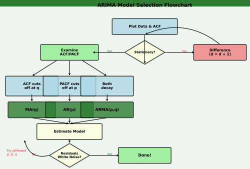
TSA_ch3_arima_flowchart

Test 10: Selecția modelului ARIMA

Întrebare

După o diferențiere, ACF arată un vârf doar la lag 1, și PACF descrește. Modelul potrivit este:

Răspuns: B – ARIMA(0,1,1)



Model: ACF se anulează la lag 1, PACF descrește $\Rightarrow$ MA(1) pentru seria diferențiată. Model complet: ARIMA(0,1,1) = IMA(1,1)

TSA_ch3_arima_flowchart

Test 11: Staționaritate în Trend vs Staționaritate în Diferențe

Întrebare

Un proces staționar în trend devine staționar prin:

Variante de răspuns

(A) Luarea diferențelor de ordinul întâi

(B) Eliminarea trendului determinist prin regresie

(C) Luarea diferențelor de ordinul doi

(D) Aplicarea ajustării sezoniere

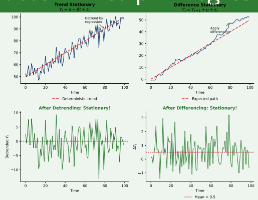
TSA_ch3_trend_vs_diff

Test 11: Staționaritate în Trend vs Staționaritate în Diferențe

Întrebare

Un proces staționar în trend devine staționar prin:

Răspuns: B – Eliminarea trendului determinist prin regresie



Staționar în trend: Eliminarea trendului prin regresie (șocurile sunt temporare). **Staționar în diferențe**: Diferențiere (șocurile sunt permanente). Tratamentul greșit afectează modelul!

TSA_ch3_trend_vs_diff

Test 12: Invertibilitatea ARIMA

Întrebare

ARIMA(0,1,1) cu $\theta_1 = 1.2$ este:

Variante de răspuns

(A) Staționară și invertibilă indent **(B)** Nestaționară dar invertibilă indent **(C)** Nestaționară și neinvertibilă indent **(D)** Staționară dar neinvertibilă

Test 12: Invertibilitatea ARIMA

Întrebare

ARIMA(0,1,1) cu $\theta_1 = 1.2$ este:

Răspuns: C – Nestaționară și neinvertibilă

Verificare staționaritate: $d = 1$ înseamnă o rădăcină unitară $\Rightarrow$ Nestaționară
Verificare invertibilitate: Polinomul MA este $\theta(z) = 1 + 1.2z$

- Rădăcină: $z = -1/1.2 = -0.833$ (în interiorul cercului unitate)
- Invertibilitatea necesită rădăcina în afara cercului unitate
- $|\theta_1| = 1.2 > 1 \Rightarrow$ Neinvertibilă

Corecție: Rescrieți cu $\theta^* = 1/1.2 = 0.833$ și ajustați varianța.

Test 13: Regresia falsă

Întrebare

Regresând un mers aleator pe un alt mers aleator independent, de obicei se obține:

Variante de răspuns

(A) Nicio relație semnificativă

(B) R^2 ridicat și statistici t semnificative (fals)

(C) Corelație negativă

(D) Multicolinearitate perfectă

Test 13: Regresia falsă

Întrebare

Regresând un mers aleator pe un alt mers aleator independent, de obicei se obține:

Răspuns: B – R^2 ridicat și statistici t semnificative (fals)

Granger & Newbold (1974): Fenomenul regresiei false

Simptome:

- R^2 ridicat (adesea > 0.9) între serii neînrudite
- Statistici t semnificative
- Statistică Durbin-Watson foarte scăzută ($\ll 2$)
- Reziduuri nestaționare

Soluții: (1) Diferențiați ambele serii, sau (2) Testați pentru cointegrare

Test 14: Prognoza pe Termen Lung

Întrebare

Prognoza pe termen lung din ARIMA(1,1,0) cu $\phi_1 = 0.7$ convergă la:

Variante de răspuns

(A) Zero

(B) Media necondiționată

(C) O extrapolare liniară a trendului

(D) Ultima valoare observată

Test 14: Prognoza pe Termen Lung

Întrebare

Prognoza pe termen lung din ARIMA(1,1,0) cu $\phi_1 = 0.7$ convergă la:

Răspuns: C – O extrapolare liniară a trendului

Model: $(1 - \phi_1 L)(1 - L)Y_t = c + \varepsilon_t$
Prognoza pe termen lung: Pentru modelele I(1) cu derivă c:

$$\hat{Y}_{T+h} \approx Y_T + h \cdot \frac{c}{1 - \phi_1}$$

Diferențe cheie:

- ARMA staționară: Prognozele $\rightarrow$ media necondiționată
- I(1) fără derivă: Prognozele $\rightarrow$ ultima valoare (plată)
- I(1) cu derivă: Prognozele $\rightarrow$ extrapolare liniară

Întrebări Adevărat/Fals

Întrebare

Determinați dacă fiecare afirmație este Adevărată sau Falsă:

1. Un proces I(2) necesită două diferențe pentru a deveni staționar.
2. Testul ADF include întotdeauna un termen constant.
3. ARIMA(0,1,0) este un alt nume pentru un mers aleator.
4. Diferențierea unei serii staționare o face "mai staționară."
5. Testul KPSS are staționaritatea ca ipoteză nulă.
6. Modelele ARIMA pot captura doar dinamici liniare.

Răspunsul pe slide-ul următor...

Adevărat/Fals: Soluții

Răspunsuri

1. I(2) necesită două diferențe. **ADEVĂRAT**
   d diferențe pt. I(d). I(2) = două rădăcini unitare.
2. ADF include întotdeauna un termen constant. **FALS**
   Alegeți: fără constantă, doar constantă, sau constantă + trend.
3. ARIMA(0,1,0) = mers aleator. **ADEVĂRAT**
   $(1-L)Y_t = \varepsilon_t \Rightarrow Y_t = Y_{t-1} + \varepsilon_t$.
4. Diferențierea unei serii staționare → "mai staționară." **FALS**
   Supradiferențierea creează MA neinvertibil.
5. KPSS: H_0 = staționară. **ADEVĂRAT**
   Opus testului ADF (H_0 = rădăcină unitară).
6. ARIMA captează doar modele liniare. **ADEVĂRAT**
   Liniar în parametri. Neliniare → GARCH, rețele neuronale.

Problema 1: Testarea Rădăcinii Unitare

Exercițiu

Aveți date trimestriale de PIB pentru 80 de trimestre. Testul ADF (cu constantă și trend) dă o statistică de test de -2.85. Valoarea critică la 5% este -3.41.

1. Care este concluzia dumneavoastră despre staționaritate?
2. Ce ați face în continuare?

Problema 1: Testarea Rădăcinii Unitare

Exercițiu

Aveți date trimestriale de PIB pentru 80 de trimestre. Testul ADF (cu constantă și trend) dă o statistică de test de -2.85. Valoarea critică la 5% este -3.41.

1. Care este concluzia dumneavoastră despre staționaritate?
2. Ce ați face în continuare?

Soluție

1. Deoarece $-2.85 > -3.41$, **nu respingem** H_0. Datele par să aibă o rădăcină unitară (nestaționare).
2. Luați prima diferență ΔY_t și repetați testul ADF pe seria diferențiată pentru a confirma că este acum staționară.

Problema 2: Identificarea modelului

Exercițiu

După diferențierea o dată a unei serii de timp, ACF arată:

- Vârf semnificativ la lag 1 ($\rho_1 = 0.4$)
- Toate celelalte lag-uri nesemnificative

PACF arată descreștere graduală.
Ce model ARIMA este sugerat?

Problema 2: Identificarea modelului

Exercițiu

După diferențierea o dată a unei serii de timp, ACF arată:

- Vârf semnificativ la lag 1 ($\rho_1 = 0.4$)
- Toate celelalte lag-uri nesemnificative

PACF arată descreștere graduală.
Ce model ARIMA este sugerat?

Soluție

- ACF se anulează după lag 1 $\Rightarrow$ componentă MA(1)
- PACF descrește $\Rightarrow$ Confirmă structura MA
- Deoarece am diferențiat o dată: $d = 1$

Model sugerat: ARIMA(0,1,1) sau IMA(1,1)

Problema 3: Ecuația ARIMA

Exercițiu

Scrieți ecuația completă pentru ARIMA(1,1,1):

$$(1 - \phi_1 L)(1 - L)Y_t = c + (1 + \theta_1 L)\varepsilon_t$$

Expandați complet în termenii Y_t, Y_{t-1}, Y_{t-2}, etc.

Problema 3: Ecuația ARIMA

Exercițiu

Scrieți ecuația completă pentru ARIMA(1,1,1):

$$(1-\phi_1 L)(1-L)Y_t = c + (1+\theta_1 L)\varepsilon_t$$

Expandați complet în termenii Y_t, Y_{t-1}, Y_{t-2}, etc.

Soluție

Expandând $(1-\phi_1 L)(1-L) = 1 - L - \phi_1 L + \phi_1 L^2 = 1 - (1+\phi_1)L + \phi_1 L^2$:

$$Y_t - (1+\phi_1)Y_{t-1} + \phi_1 Y_{t-2} = c + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

Sau echivalent:

$$Y_t = c + (1+\phi_1)Y_{t-1} - \phi_1 Y_{t-2} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

Problema 4: Calculul Prognozei

Exercițiu

Dat ARIMA(0,1,1): $\Delta Y_t = \varepsilon_t + 0.3\varepsilon_{t-1}$
La momentul T: $Y_T = 100$, $\hat{\varepsilon}_T = 2$, $\sigma^2 = 4$
Calculați:

1. $\hat{Y}_{T+1|T}$ (prognoza la un pas)
2. $\hat{Y}_{T+2|T}$ (prognoza la doi pași)

Problema 4: Calculul Prognozei

Exercițiu

Dat ARIMA(0,1,1): $\Delta Y_t = \varepsilon_t + 0.3\varepsilon_{t-1}$
La momentul T: $Y_T = 100$, $\hat{\varepsilon}_T = 2$, $\sigma^2 = 4$
Calculați:

1. $\hat{Y}_{T+1|T}$ (prognoza la un pas)
2. $\hat{Y}_{T+2|T}$ (prognoza la doi pași)

Soluție

1. $\hat{Y}_{T+1|T} = Y_T + 0.3\hat{\varepsilon}_T = 100 + 0.3(2) = 100.6$
2. $\hat{Y}_{T+2|T} = \hat{Y}_{T+1|T} + 0.3 \cdot 0 = 100.6 + 0 = 100.6$
   (Șocurile viitoare $\varepsilon_{T+1}, \varepsilon_{T+2}$ se presupun egale cu 0)

Problema 5: Intervale de Încredere

Exercițiu

Continuând de la Problema 4, calculați intervalele de încredere de 95% pentru $\hat{Y}_{T+1|T}$ și $\hat{Y}_{T+2|T}$.
Reamintim: $\sigma^2 = 4$, $\theta_1 = 0.3$

Problema 5: Intervale de Încredere

Exercițiu

Continuând de la Problema 4, calculați intervalele de încredere de 95% pentru $\hat{Y}_{T+1|T}$ și $\hat{Y}_{T+2|T}$.
Reamintim: $\sigma^2 = 4$, $\theta_1 = 0.3$

Soluție

Pentru IMA(1,1), ponderile MA(∞) sunt $\psi_0 = 1$, $\psi_j = 1 + \theta_1$ pentru $j \geq 1$.
1 pas: $\text{Var}(e_{T+1}) = \sigma^2\psi_0^2 = 4$, deci $SE = 2$
$100.6 \pm 1.96(2) = [96.68, 104.52]$
2 pași: $\text{Var}(e_{T+2}) = \sigma^2(\psi_0^2 + \psi_1^2) = 4(1 + 1.3^2) = 10.76$, $SE = 3.28$
$100.6 \pm 1.96(3.28) = [94.17, 107.03]$

Exemplu: Testarea Rădăcinii Unitare în Prețurile Acțiunilor

Scenariu

Aveți prețuri de închidere zilnice pentru o acțiune pe parcursul a 500 de zile. Vreți să determinați dacă prețurile urmează un mers aleator.

Abordare Pas cu Pas

1. **Inspecție vizuală**: Reprezentați grafic prețurile – probabil arată trend
2. **Testul ADF pe prețuri**: Așteptați să nu respingeți H_0 (rădăcină unitară)
3. **Luați randamentele logaritmice**: $r_t = \ln(P_t/P_{t-1}) = \Delta \ln(P_t)$
4. **Testul ADF pe randamente**: Ar trebui să respingeți H_0 (staționară)
5. **Concluzie**: Log prețurile sunt $I(1)$, randamentele sunt $I(0)$

Exemplu: Box-Jenkins pentru Date de Inflație

Scenariu

Rate lunare ale inflației pentru 10 ani. Construiți un model ARIMA.

Flux de lucru

1. **Reprezentare grafică și test**: ADF sugerează limită – încercați atât $d = 0$ cât și $d = 1$
2. **Dacă** $d = 0$: Ajustați modele ARMA, comparați AIC
3. **Dacă** $d = 1$: Examinați ACF/PACF ale lui ΔY_t
   - ACF: vârf la lag 1, apoi se anulează
   - PACF: descrește
   - $\Rightarrow$ Încercați ARIMA(0,1,1)
4. **Estimare**: Ajustați ARIMA(0,1,1), verificați coeficienții
5. **Diagnostic**: Ljung-Box pe reziduuri (vrem $p > 0.05$)
6. **Comparare**: AIC al ARIMA(0,1,1) vs ARMA(1,1) pe seria originală

Exemplu: interpretarea Rezultatelor Python

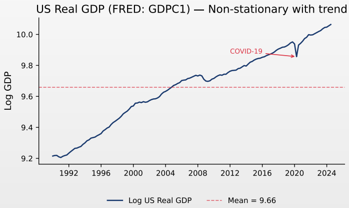
Rezultate ARIMA din statsmodels

```
                    ARIMA Model Results
==============================================================
Dep. Variable:            D.y   No. Observations:    99
Model:             ARIMA(1,1,1)   AIC                  285.32
                                  BIC                  295.63
==============================================================
                coef     std err      z      P>|z|
--------------------------------------------------------------
const          0.0521     0.048     1.085    0.278
ar.L1          0.4532     0.102     4.443    0.000
ma.L1         -0.2891     0.118    -2.450    0.014
sigma2         1.2340     0.176     7.011    0.000
```

Interpretare

- AR (0.45) semnificativ, MA (-0.29) semnificativ
- Constanta (0.052) nesemnificativă – se poate seta $c = 0$
- Verificare: $|\phi_1| < 1$ (staționar), $|\theta_1| < 1$ (invertibil) – OK!

Studiu de Caz: PIB Real SUA (1990–2024)

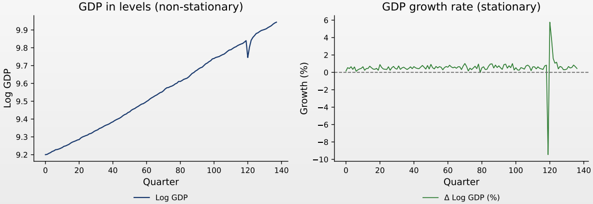


Observații

PIB Real SUA în miliarde $ 2017 (trimestrial). **Trend ascendent** clar. Scăderi în recesiuni (2008–09, 2020). Nestaționară: necesită diferențiere.

TSA_ch3_gdp_levels

Staționaritate Prin Diferențiere

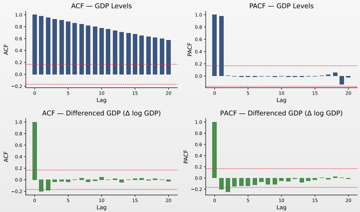


Observații

- **Stânga**: PIB (serie originală) — trend ascendent clar (nestaționară)
- **Dreapta**: Rata de creștere a PIB $= \Delta \log(Y_t) \times 100$ — staționară, fluctuează în jurul mediei ($\approx 0.6\%$/trim.)

TSA_ch3_differencing

ACF/PACF: Serie originală vs Diferențiată



Observații

- **Rândul de sus**: ACF/PACF ale seriei originale PIB — descreștere lentă ⇒ nestaționaritate
- **Rândul de jos**: ACF/PACF ale creșterii PIB — valori în limitele de încredere
- Un model ARIMA de ordin mic este potrivit

TSA_ch3_acf_pacf

Rezultate Estimare ARIMA: Creșterea PIB SUA

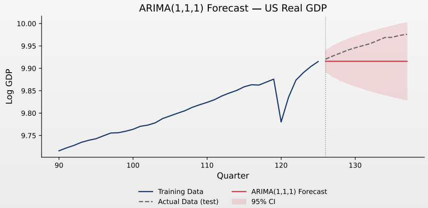
Model: ARIMA$(1, 1, 1)$ pe log(PIB)

Parametru	Estimat	Eroare Std.	z-stat	valoare-p
ϕ_1 (AR.L1)	0.312	0.185	1.69	0.091
θ_1 (MA.L1)	-0.087	0.203	-0.43	0.668
σ^2	0.00012	–	–	–

Interpretare

- ARIMA de ordin mic captează rezonabil dinamica PIB
- Coeficientul AR(1) pozitiv – creșterea PIB arată persistență
- Alternativă: mersul aleatoriu simplu (ARIMA(0,1,0)) adesea competitiv

Prognoză: ARIMA vs Real

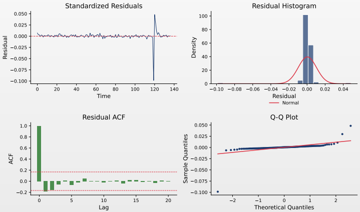


Observații

- **Albastru**: date istorice de antrenare; **Verde**: date reale de test
- **Roșu**: prognoze ARIMA cu IC 95% — IC se lărgesc cu orizontul de prognoză

TSA_ch3_arima_forecast

Diagnostice Model: Analiza Reziduurilor



Observații

- Reziduurile fără tipare sistematice în timp; distribuție aproximativ normală (histogramă, Q-Q)
- ACF reziduuri în limite — fără autocorelare; modelul captează adecvat procesul generator de date

TSA_ch3_diagnostics

Discuție: Trenduri Deterministe vs Stochastice

Întrebare Cheie

De ce este important să distingem între trendurile deterministe și stochastice?

Puncte de Discuție

- **Consecințele tratamentului greșit**:
  - Eliminarea trendului prin regresie când seria are rădăcină unitară $\Rightarrow$ staționaritate falsă
  - Diferențierea unei serii staționare în trend $\Rightarrow$ supradiferențiere
- **Interpretare economică**:
  - Trend determinist: șocurile sunt temporare
  - Trend stochastic: șocurile au efecte permanente
- **Implicații de politică**:
  - O recesiune reduce permanent PIB-ul, sau economia revine la trend?

Discuție: Criterii de Selecție a Modelului

Întrebare Cheie

Când ar trebui să folosiți AIC vs BIC pentru selecția modelului ARIMA?

Considerații

- **AIC**: Minimizează eroarea de predicție, poate supraajusta
  - Mai bun pentru prognoză
  - Tinde să selecteze modele mai mari
- **BIC**: Selecție consistentă a modelului, mai simplu
  - Mai bun pentru identificarea modelului "adevărat"
  - Penalizează complexitatea mai puternic
- **Sfat practic**: Raportați ambele, preferați BIC dacă diferă substanțial

Discuție: Limitările ARIMA

Întrebare Cheie

Care sunt principalele limitări ale modelelor ARIMA?

Puncte de Discuție

- **Liniaritate**: Nu poate captura dinamici neliniare
- **Varianță constantă**: Presupune homoscedasticitate (fără GARCH)
- **Fără rupturi structurale**: Parametrii presupuși constanți
- **Univariat**: Ignoră relațiile cu alte variabile
- **Simetric**: Tratează șocurile pozitive și negative la fel
- **Prognoze pe termen lung**: Incertitudinea crește rapid

Extensii

Aceste limitări motivează GARCH (volatilitate), VAR (multivariat), modele cu schimbări de regim, etc.

Exercițiu AI 1: Critica unei Analize AI

Scenariu

Ați cerut unui AI: „Aplică cel mai bun model ARIMA pe datele PIB-ului României." A returnat:

- A ajustat ARIMA(3,2,3) cu AIC = 1542.7
- Nu a efectuat testul ADF
- Ljung-Box p-value = 0.02 (raportat ca „acceptabil")
- Prognoză pe 30 de ani cu intervale de încredere înguste

Critica voastră:

1. Este ARIMA(3,2,3) supra-parametrizat? Ce ar sugera BIC?
2. De ce Ljung-Box p = 0.02 **nu** este acceptabil la pragul de 5%?
3. Sunt prognozele pe 30 de ani fiabile pentru modele ARIMA? De ce?
4. Ce pași din metodologia Box-Jenkins au fost omiși?

Exercițiu AI 2: Rafinarea Prompturilor pentru ARIMA

Sarcină

Îmbunătățiți iterativ prompturile pentru ajustarea unui model ARIMA pe date PIB.

Runda 1 (vag): *„Ajustează un model de serie de timp pe PIB"*

- Ce a produs AI-ul? Ce lipsește?

Runda 2 (mai bun): *„Testează staționaritatea cu ADF și KPSS, diferențiază dacă e necesar, examinează ACF/PACF, ajustează ARIMA(p,d,q) folosind BIC, verifică reziduurile cu Ljung-Box"*

- A urmat AI-ul metodologia Box-Jenkins?

Runda 3 (expert): *„Urmează Box-Jenkins: (1) grafic & test staționaritate ADF+KPSS, (2) diferențiere, (3) identificare ordine din ACF/PACF, (4) estimare ARIMA(1,1,1), (5) Ljung-Box pe reziduuri, (6) prognoză 8 trimestre cu IC 95%"*

- Comparați rezultatele din cele trei runde

Exercițiu AI 3: Competiție de Selecție a Modelului

Sarcină

Descărcați date trimestriale PIB real SUA de pe FRED (seria GDPC1).

Abordarea voastră (manuală):

- Test ADF + KPSS → diferențiere
- ACF/PACF → modele candidat
- AIC/BIC: ARIMA(0,1,0), (1,1,0), (0,1,1), (1,1,1)
- Diagnostice reziduuri + prognoză rolling 1-pas

Abordarea AI:

- Cereți AI-ului: „găsește cel mai bun ARIMA și fă prognoze"

Comparați:

- Ce model a selectat fiecare? Comparați RMSE
- Prognoze rolling vs multi-pas?
- **Predați:** reflexie 1 pagină despre AI

Rezumat Formule cheie

Concept	Formula
Mers aleatoriu	$Y_t = Y_{t-1} + \varepsilon_t$
Varianța mersului aleatoriu	$\text{Var}(Y_t) = t\sigma^2$
ARIMA(p, d, q)	$\phi(L)(1-L)^d Y_t = \theta(L)\varepsilon_t$
Prima diferență	$\Delta Y_t = Y_t - Y_{t-1} = (1-L)Y_t$
A doua diferență	$\Delta^2 Y_t = Y_t - 2Y_{t-1} + Y_{t-2}$
Regresia ADF	$\Delta Y_t = \alpha + \gamma Y_{t-1} + \sum \delta_j \Delta Y_{t-j} + \varepsilon_t$
Ipoteza nulă ADF	$H_0 : \gamma = 0$ (rădăcină unitară)
Prognoză mers aleator	$\hat{Y}_{T+h\|T} = Y_T$
IC prognoză mers aleator	$Y_T \pm z_{\alpha/2}\sqrt{h}\,\sigma$
AIC	$-2\ln(\hat{L}) + 2k$
BIC	$-2\ln(\hat{L}) + k\ln(n)$

Notații: $\hat{L}$ = maximul funcției de verosimilitate, k = nr. parametri, n = dimensiunea eșantionului, σ^2 = varianța zgomotului alb

Vă mulțumim!

Întrebări?

Materialele seminarului sunt disponibile la:
https://danpele.github.io/Time-Series-Analysis/

Quantlet Quantinar

Bibliografie I

Manuale fundamentale

- Hyndman, R.J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice*, 3rd ed., OTexts.
- Shumway, R.H., & Stoffer, D.S. (2017). *Time Series Analysis and Its Applications*, 4th ed., Springer.
- Brockwell, P.J., & Davis, R.A. (2016). *Introduction to Time Series and Forecasting*, 3rd ed., Springer.

Serii de timp financiare

- Tsay, R.S. (2010). *Analysis of Financial Time Series*, 3rd ed., Wiley.
- Franke, J., Härdle, W.K., & Hafner, C.M. (2019). *Statistics of Financial Markets*, 4th ed., Springer.

Bibliografie II

Abordări moderne și Machine Learning

- Nielsen, A. (2019). *Practical Time Series Analysis*, O'Reilly Media.
- Petropoulos, F., et al. (2022). *Forecasting: Theory and Practice*, International Journal of Forecasting.
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020). The M4 Competition, International Journal of Forecasting.

Resurse online și cod

- **Quantlet**: `https://quantlet.com` — Platformă de cod pentru statistică
- **Quantinar**: `https://quantinar.com` — Platformă pentru metode cantitative
- **GitHub TSA**: `https://github.com/QuantLet/TSA/tree/main/TSA_ch3` — Cod Python pentru fiecare capitol