



Analiza și Prognoza seriilor de timp

Seminar 7: Cointegrare și modele VECM



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

Structura seminarului:

1. **Test de Recapitulare** – Verificarea cunoștințelor
2. **Întrebări Adevărat/Fals** – Verificări conceptuale
3. **Probleme Practice** – Practică aplicată
4. **Exemple Rezolvate** – Soluții detaliate
5. **Subiecte de Discuție** – Gândire critică
6. **Exerciții cu asistență AI** – Inteligență artificială aplicată

Test 1: Definiția Cointegrării

Întrebare

Două variabile $I(1)$, X_t și Y_t , sunt cointegrate dacă:

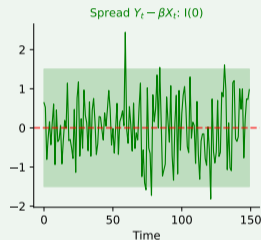
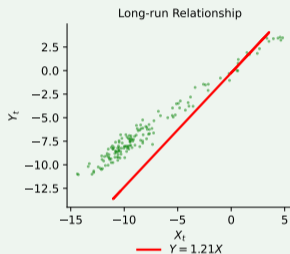
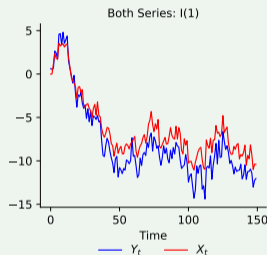
Variante de răspuns

- (A) Ambele sunt staționare
- (B) Suma lor este $I(2)$
- (C) O combinație liniară a lor este $I(0)$
- (D) Au aceeași medie

Răspunsul pe slide-ul următor...

Test 1: Răspuns

Răspuns: C – O combinație liniară este $I(0)$



Cheie: $Y_t - \beta X_t \sim I(0)$ înseamnă că seriile au un trend stochastic comun. Combinația liniară (spread-ul) este staționară chiar dacă ambele serii sunt nestaționare.

Test 2: Regresia falsă

Întrebare

Când regresăm un mers aleator pe alt mers aleator independent, de obicei obținem:

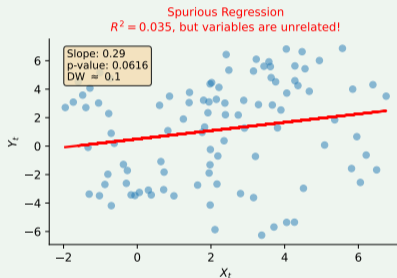
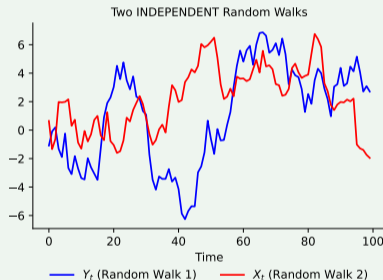
Variante de răspuns

- (A) R^2 mic și coeficienți nesemnificativi
- (B) R^2 mare și coeficienți semnificativi (fals!)
- (C) Coeficienți zero
- (D) Rezultate nedefinite

Răspunsul pe slide-ul următor...

Test 2: Răspuns

Răspuns: B – R^2 mare și coeficienți semnificativi (fals!)



Granger-Newbold (1974): Regresarea seriilor $I(1)$ nerelaționate dă rezultate înșelătoare. Regulă: Dacă $R^2 > DW$, suspectați regresie falsă! 🌐 Exemple reale: tylervigen.com/spurious-correlations

Test 3: Testul Engle-Granger

Întrebare

În metoda Engle-Granger în doi pași, ce testăm în pasul 2?

Variante de răspuns

- (A) Dacă variabilele originale sunt staționare
- (B) Dacă reziduurile regresiei au rădăcină unitară
- (C) Dacă coeficienții sunt semnificativi
- (D) Dacă R^2 este suficient de mare

Răspunsul pe slide-ul următor...

Test 3: Răspuns

Răspuns: B – Dacă reziduurile au rădăcină unitară

Pasul 1: Rulăm OLS: $Y_t = \alpha + \beta X_t + e_t$, salvăm reziduurile \hat{e}_t

Pasul 2: Test ADF pe reziduuri: $\Delta \hat{e}_t = \rho \hat{e}_{t-1} + \dots$

□ $H_0: \rho = 0$ (rădăcină unitară \Rightarrow fără cointegrare)

□ $H_1: \rho < 0$ (staționar \Rightarrow cointegrare!)

Important: Folosiți valorile critice Engle-Granger, nu ADF standard!

 TSA_ch7_engle_granger

Test 4: Avantajul Testului Johansen

Întrebare

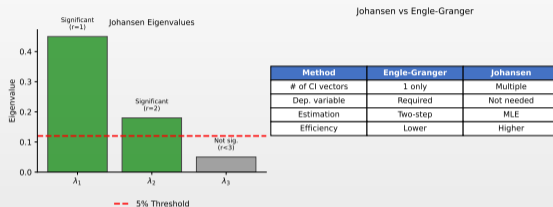
Principalul avantaj al testului Johansen față de Engle-Granger este:

Variante de răspuns

- (A) Este mai simplu de calculat
- (B) Poate detecta relații de cointegrare multiple
- (C) Nu necesită date
- (D) Găsește întotdeauna cointegrare

Răspunsul pe slide-ul următor...

Test 4: Răspuns



Observații

- ▣ Testează pentru $r = 0, 1, 2, \dots, k - 1$ vectori de cointegrare
- ▣ Verosimilitate maximă (mai eficient)
- ▣ Nu necesită alegerea variabilei dependente

Test 5: Rangul Matricei Π

Întrebare

Într-un VECM cu $k = 3$ variabile, dacă $\text{rang}(\Pi) = 2$, aceasta înseamnă:

Variante de răspuns

- (A) Fără cointegrare
- (B) O relație de cointegrare
- (C) Două relații de cointegrare
- (D) Toate variabilele sunt staționare

Răspunsul pe slide-ul următor...

Test 5: Răspuns

Răspuns: C – Două relații de cointegrare

Interpretarea rangului pentru k variabile:

- ☐ $\text{rang}(\Pi) = 0$: Fără cointegrare (folosiți VAR în diferențe)
- ☐ $0 < \text{rang}(\Pi) = r < k$: r vectori de cointegrare (folosiți VECM)
- ☐ $\text{rang}(\Pi) = k$: Toate variabilele sunt $I(0)$ (folosiți VAR în niveluri)

Cu $k = 3$ și $r = 2$:

- ☐ Două relații de echilibru
- ☐ Doar $k - r = 1$ trend stochastic comun

Test 6: Structura VECM

Întrebare

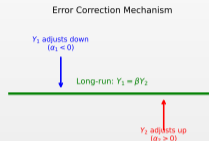
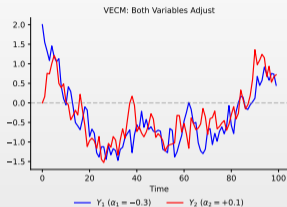
În ecuația VECM $\Delta Y_t = c + \alpha\beta'Y_{t-1} + \dots$, ce reprezintă α ?

Variante de răspuns

- (A) Vectorii de cointegrare
- (B) Coeficienții de ajustare (încărcare)
- (C) Dinamica pe termen scurt
- (D) Varianța erorilor

Răspunsul pe slide-ul următor...

Test 6: Răspuns



Observații

- ▣ β = vectorii de cointegrare (definesc echilibrul)
- ▣ α = vitezele de ajustare (cât de repede corectează fiecare variabilă)

Test 7: Termenul de Corecție a Erorilor

Întrebare

Dacă $Y_t - \beta X_t$ este relația de cointegrare și acest termen este pozitiv, ce se întâmplă?

Variante de răspuns

- (A) Y este deasupra echilibrului; Y ar trebui să scadă (dacă $\alpha < 0$)
- (B) Y este sub echilibru; Y ar trebui să crească
- (C) Nimic, corecția erorilor nu afectează nivelurile
- (D) Ambele variabile cresc

Răspunsul pe slide-ul următor...

Test 7: Răspuns

Răspuns: $A - Y$ deasupra echilibrului; scade dacă $\alpha < 0$

Mecanismul de corecție a erorilor:

$$\Delta Y_t = \alpha(Y_{t-1} - \beta X_{t-1}) + \dots$$

- Dacă $Y_{t-1} - \beta X_{t-1} > 0$: Y este “prea sus”
- Cu $\alpha < 0$: $\Delta Y_t < 0$ (Y scade spre echilibru)
- Aceasta este “corecția erorilor” care trage Y înapoi

Convenție de semn: α ar trebui să fie negativ pentru ca variabila dependentă să se miște înapoi spre echilibru.

Test 8: Exogenitate Slabă

Întrebare

Dacă $\alpha_2 = 0$ într-un VECM bivariat, aceasta înseamnă:

Variante de răspuns

- (A) Nu există cointegrare
- (B) Variabila 2 nu se ajustează la dezechilibru (slab exogenă)
- (C) Variabila 1 nu se ajustează
- (D) Ambele variabile sunt staționare

Răspunsul pe slide-ul următor...

Test 8: Răspuns

Răspuns: B – Variabila 2 este slab exogenă

Exogenitate slabă: Variabila nu răspunde la dezechilibru.

Exemplu: Ratele dobânzii

- ▣ Rata pe termen lung (R_t) adesea slab exogenă ($\alpha_R \approx 0$)
- ▣ Rata pe termen scurt (r_t) se ajustează la spread ($\alpha_r < 0$)
- ▣ Interpretare: Banca centrală ajustează rata scurtă pentru a menține structura termenelor

Implicație: Putem estima o singură ecuație pentru variabila care se ajustează.

Test 9: Testul Trace

Întrebare

Testul trace Johansen cu $H_0 : r \leq 1$ vs $H_1 : r > 1$ testează dacă:

Variante de răspuns

- (A) Există exact un vector de cointegrare
- (B) Există cel mult un vector de cointegrare
- (C) Există mai mult de un vector de cointegrare
- (D) Toate valorile proprii sunt zero

Răspunsul pe slide-ul următor...

Test 9: Răspuns

Răspuns: B/C – H_0 : cel mult 1; H_1 : mai mult de 1

Procedura de testare secvențială:

1. Testăm $H_0 : r = 0$ vs $H_1 : r > 0$
2. Dacă respingem, testăm $H_0 : r \leq 1$ vs $H_1 : r > 1$
3. Continuăm până nu mai respingem...

Statistica trace:

$$\lambda_{\text{trace}}(r) = -T \sum_{i=r+1}^k \ln(1 - \hat{\lambda}_i)$$

Respingem H_0 dacă statistica trace $>$ valoarea critică.

Test 10: VECM vs VAR în Diferențe

Întrebare

Dacă variabilele sunt cointegrate, folosirea VAR în diferențe prime în loc de VECM:

Variante de răspuns

- (A) Dă rezultate identice
- (B) Este mai eficientă
- (C) Pierde informația pe termen lung (model greșit specificat)
- (D) Este abordarea preferată

Răspunsul pe slide-ul următor...

Test 10: Răspuns

Răspuns: C – Pierde informația pe termen lung

Teorema Reprezentării Granger: Dacă există cointegrare, reprezentarea VECM există și trebuie folosită.

	VAR(Δ)	VECM
Echilibru pe termen lung	Pierdut	Păstrat
Corecția erorilor	Nu	Da
Prognoze (termen lung)	Slabe	Mai bune

Concluzie: Diferențierea elimină relația pe termen lung pe care o reprezintă cointegrarea!

Întrebări Adevărat/Fals

Determinați dacă fiecare afirmație este Adevărată sau Falsă:

1. Cointegrarea necesită ca toate variabilele să fie $I(1)$.
2. Vectorul de cointegrare este unic.
3. Regresia falsă are statistică Durbin-Watson mică.
4. În VECM, ambii coeficienți α trebuie să fie nenuli.
5. Testul Johansen necesită alegerea unei variabile dependente.
6. Numărul de trenduri comune $= k - r$.

Răspunsurile pe slide-ul următor...

Adevărat/Fals: Soluții

1. **Cointegrarea necesită ca toate variabilele să fie $I(1)$.**
Cazul standard $CI(1,1)$: toate variabilele $I(1)$, combinația liniară $I(0)$.
2. **Vectorul de cointegrare este unic.**
Unic doar până la înmulțirea cu un scalar. De obicei normalizat ($\beta_1 = 1$).
3. **Regresia falsă are statistică Durbin-Watson mică.**
 $DW \approx 0$ indică reziduuri puternic autocorelate (nestaționare).
4. **În VECM, ambii coeficienți α trebuie să fie nenuli.**
Unul poate fi zero (exogenitate slabă). Cel puțin unul trebuie să fie nenul.
5. **Testul Johansen necesită alegerea unei variabile dependente.**
Asta e pentru Engle-Granger. Johansen tratează toate variabilele simetric.
6. **Numărul de trenduri comune $= k - r$.**
 k variabile, r relații de cointegrare $\Rightarrow k - r$ trenduri stochastice comune.

ADEVĂRAT

FALS

ADEVĂRAT

FALS

FALS

ADEVĂRAT

Problema 1: Identificarea Cointegrării

Exercițiu

Aveți date trimestriale pentru consum (C_t) și venit (Y_t). Testele ADF arată că ambele sunt $I(1)$. Regresia $C_t = 0.85Y_t + e_t$ dă reziduuri cu statistica $ADF = -3.92$. Valoarea critică Engle-Granger la 5% pentru 2 variabile este -3.34 .

Sunt C_t și Y_t cointegrate?

Răspunsul pe slide-ul următor...

Problema 1: Soluție

Soluție: Da, sunt cointegrate

Test: H_0 : Fără cointegrare (reziduurile au rădăcină unitară)

Statistica ADF: -3.92

Valoarea critică (5%): -3.34

Deoarece $-3.92 < -3.34$, **respingem** H_0 la nivelul de 5%.

Concluzie: Reziduurile sunt staționare \Rightarrow Există cointegrare!

Interpretare: Consumul și venitul au un trend comun. Vectorul de cointegrare este aproximativ $(1, -0.85)$, consistent cu ipoteză venitului permanent.

Problema 2: interpretarea VECM

Exercițiu

Un VECM pentru rata pe termen scurt (r_t) și rata pe termen lung (R_t) dă:

$$\Delta r_t = 0.01 - 0.25(r_{t-1} - R_{t-1}) + \dots$$

$$\Delta R_t = 0.005 - 0.02(r_{t-1} - R_{t-1}) + \dots$$

Interpretați coeficienții de ajustare.

Răspunsul pe slide-ul următor...

Problema 2: Soluție

TSA ch7 vecm interpretation

Soluție

Termenul de corecție a erorilor: $(r_{t-1} - R_{t-1}) = \text{spread-ul}$

Rata pe termen scurt ($\alpha_r = -0.25$):

- Când spread-ul este pozitiv (scurt > lung), rata scurtă scade
- 25% din dezechilibru este corectat per perioadă

Rata pe termen lung ($\alpha_R = -0.02$):

- Coeficient de ajustare foarte mic — rata lungă este aproape slab exogenă
- Condușă mai mult de așteptări, nu de corecția erorilor

Interpretare economică: Banca centrală (rata scurtă) se ajustează pentru a menține curba randamentelor.

Problema 3: Rezultatele Testului Johansen

Exercițiu

Testul trace Johansen pentru 3 variabile dă:

H_0	Stat. Trace	VC 5%
$r = 0$	45.2	29.8
$r \leq 1$	18.1	15.5
$r \leq 2$	3.2	3.8

Care este rangul de cointegrare?

Răspunsul pe slide-ul următor...

Problema 3: Soluție

Soluție: Rangul = 2

Testare secvențială:

1. $H_0 : r = 0$: $45.2 > 29.8 \Rightarrow$ **Respingem** (cel puțin 1)
2. $H_0 : r \leq 1$: $18.1 > 15.5 \Rightarrow$ **Respingem** (cel puțin 2)
3. $H_0 : r \leq 2$: $3.2 < 3.8 \Rightarrow$ **Nu respingem**

Concluzie: $r = 2$ relații de cointegrare

Implicații:

- Două relații de echilibru între 3 variabile
- Doar $3 - 2 = 1$ trend stochastic comun
- Folosiți VECM cu 2 termeni de corecție a erorilor

Exemplu: Structura la Termen a Ratelor Dobânzii

Teoria Economică

Ipoteza aşteptărilor: $R_t^{(n)} = \frac{1}{n} \sum_{i=0}^{n-1} E_t[r_{t+i}] + \text{primă}$

Dacă prima este constantă \Rightarrow spread-ul ($R_t - r_t$) ar trebui să fie staţionar.

Constatări Tipice

- Ambele rate sunt $I(1)$ (confirmat de ADF)
- Testul Johansen: $r = 1$ vector de cointegrare
- Vector de cointegrare $\approx (1, -1)$: spread-ul este staţionar
- Rata scurtă se ajustează ($\alpha_r < 0$), rata lungă slab exogenă

Implicaţie de Politică

Banca centrală controlează rata scurtă; rata lungă este condusă de aşteptări.

Exemplu: Paritatea Puterii de Cumpărare (PPP)

Teoria PPP

$e_t = p_t - p_t^*$ (log curs de schimb = diferențialul de prețuri)

Cursul real de schimb: $q_t = e_t - p_t + p_t^*$ ar trebui să fie staționar (PPP pe termen lung)

Provocări Empirice

- Teste rădăcină unitară: e_t, p_t, p_t^* toate $I(1)$
- Teste de cointegrare: Rezultate mixte în funcție de eșantion
- Timp de înjumătățire al deviațiilor PPP: 3-5 ani (ajustare lentă)
- Exogenitate slabă: Cursul de schimb adesea nu se ajustează

Puzzle-ul PPP

Cursul real de schimb este foarte persistent—revenirea lentă la medie este greu de explicat cu modelele standard.

Exemplu: Strategia Pairs Trading

Ideea

Găsiți acțiuni cointegrate \Rightarrow tranzacționați spread-ul staționar

Pași de Implementare

1. **Identificați perechile:** Testați cointegrarea (ex., Coca-Cola & Pepsi)
2. **Estimați spread-ul:** $z_t = P_A - \beta P_B$
3. **Reguli de tranzacționare:**
 - ▶ $z_t > \mu + 2\sigma$: Vindeți A, Cumpărați B (spread prea larg)
 - ▶ $z_t < \mu - 2\sigma$: Cumpărați A, Vindeți B (spread prea îngust)
 - ▶ Ieșiți când $z_t \approx \mu$

Riscuri

Cointegrarea se poate rupe; spread-ul poate să nu revină; costuri de tranzacție.

Analiza Cointegrării în Python: Funcții Cheie

Biblioteci Esențiale

```
from statsmodels.tsa.stattools import coint, adfuller  
from statsmodels.tsa.vector_ar.vecm import coint_johansen, VECM
```

flux de lucru

1. Teste rădăcină unitară: `adfuller(serie)`
2. Engle-Granger: `coint(y, x)` returnează stat. test & p-value
3. Johansen: `coint_johansen(data, det_order, k_ar_diff)`
4. Estimare VECM: `model = VECM(data, k_ar_diff=2, coint_rank=1)`
5. Rezultate: `results = model.fit()`

Notă

Exemple complete de lucru sunt furnizate în notebook-urile Jupyter.

Discuție: Cointegrare vs Corelație

Întrebare Cheie

Două serii sunt puternic corelate. Sunt ele cointegrate?

Răspuns: Nu neapărat!

- ▣ **Corelație:** Măsoară co-mișcarea (poate fi falsă pentru $I(1)$)
- ▣ **Cointegrare:** Necesită combinație liniară staționară

Exemplu

Două mersuri aleatoare independente pot avea corelație > 0.9 pur întâmplător (corelație falsă). Dar NU sunt cointegrate—spread-ul lor este tot $I(1)$.

Cointegrarea implică o relație de echilibru pe termen lung semnificativă.

Discuție: Alegerea Componentelor Deterministe

Întrebare Cheie

Testul Johansen are 5 cazuri pentru componentele deterministe. Pe care să îl alegem?

Ghid

1. **Fără constantă, fără trend:** Rar folosit (necesită date cu medie zero)
2. **Constantă doar în EC:** Serii în nivel, fără drift
3. **Constantă nerestricționată:** Cel mai comun pentru date economice
4. **Trend în EC:** Seriile au trenduri deterministe
5. **Trend nerestricționat:** Diferențe cu trend (necomun)

Sfat Practic

Începeți cu Cazul 3 (constantă nerestricționată). Verificați sensibilitatea la specificație. Folosiți raționament economic: au nivelurile trenduri?

Exerciții de Făcut Acasă

1. **Teoretic:** Arătați că dacă Y_t și X_t sunt ambele mersuri aleatoare cu aceeași inovație, ele sunt cointegrate.

2. **Calcul:** Având estimările VECM:

$$\Delta Y_t = 0.5 - 0.3(Y_{t-1} - 2X_{t-1}) + 0.2\Delta Y_{t-1}$$

$$\Delta X_t = 0.1 + 0.1(Y_{t-1} - 2X_{t-1}) + 0.4\Delta X_{t-1}$$

- ▶ Care este vectorul de cointegrare?
- ▶ Care variabilă se ajustează mai rapid?
- ▶ Care este relația de echilibru pe termen lung?

3. **Aplicat:** Descărcați ratele trezoreriei la 10 ani și 3 luni:

- ▶ Testați pentru rădăcini unitare; Testați pentru cointegrare
- ▶ Estimați VECM; Interpretați coeficienții de ajustare

4. **Gândire Critică:** De ce ar putea PPP să fie valabilă pe termen lung dar nu pe termen scurt?

Indicii pentru Soluțiile Exercițiilor

Indicii

1. Dacă $Y_t = Y_{t-1} + \varepsilon_t$ și $X_t = X_{t-1} + \varepsilon_t$ (același șoc), atunci $Y_t - X_t = Y_0 - X_0$ este constantă (staționară).
2. Din VECM:
 - ▶ Vector de cointegrare: $(1, -2)$ (normalizat pe Y)
 - ▶ Y se ajustează mai rapid: $|\alpha_Y| = 0.3 > |\alpha_X| = 0.1$
 - ▶ Termen lung: $Y = 2X$ (când termenul EC = 0)
3. Pentru ratele dobânzii:
 - ▶ Ambele sunt de obicei $I(1)$; spread-ul de obicei staționar
 - ▶ Așteptați un vector de cointegrare cu $(1, -1)$
 - ▶ Rata scurtă se ajustează de obicei; rata lungă adesea slab exogenă
4. Deviații PPP: Costurile de transport, bunurile netranzactionate, prețurile rigide, tarifele, segmentarea pieței toate încetinesc ajustarea dar nu previn convergența pe termen lung.

Concluzii cheie din Acest Seminar

Puncte Principale

1. **Cointegrarea:** Variabile $I(1)$ cu combinație liniară staționară
2. **Regresia falsă:** R^2 mare fără cointegrare este lipsit de sens
3. **Engle-Granger:** Simplu, dar doar un vector de cointegrare
4. **Johansen:** Vectori multipli, MLE, mai puternic

Perspective VECM

- ▣ β definește echilibrul; α determină viteza de ajustare
- ▣ Exogenitate slabă ($\alpha = 0$): Variabila nu răspunde la dezechilibru
- ▣ Folosiți întotdeauna VECM (nu VAR în diferențe) când există cointegrare

Amintiți-vă

Cointegrarea este despre **echilibrul pe termen lung**, nu doar corelație!

Exercițiu AI: Gândire critică

Prompt de testat în ChatGPT / Claude / Copilot

“Descarcă prețurile spot și futures pe petrol (CL=F, BZ=F). Testează cointegrarea și estimează un model VECM. Pot folosi spread-ul pentru o strategie de pairs trading?”

Exercițiu:

1. Rulați prompt-ul într-un LLM la alegere și analizați critic răspunsul.
2. AI testează ordinul de integrare al fiecărei serii înainte de a testa cointegrarea?
3. Testul Johansen sau Engle-Granger? AI justifică alegerea?
4. Rangul de cointegrare (r) este interpretat corect?
5. Strategia de pairs trading ține cont de costurile de tranzacționare și riscul de model?

Atenție: Codul generat de AI poate rula fără erori și arăta profesional. *Asta nu înseamnă că e corect.*


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

Abordari moderne si 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 si cod

- ▣ **Quantlet**: <https://quantlet.com> — Repository de cod pentru statistica
- ▣ **Quantinar**: <https://quantinar.com> — Platforma de invatare metode cantitative
- ▣ **GitHub TSA**: <https://github.com/QuantLet/TSA> — Cod Python pentru fiecare capitol