



# Analiza și Prognoza Seriilor de Timp

Seminar 1: Introducere în Analiza Seriilor de Timp



Daniel Traian PELE

Academia de Studii Economice din Bucureşti

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. Recapitulare Rapidă** – Rezumatul conceptelor cheie
- 2. Test Grilă** – Verificarea cunoștințelor
- 3. Întrebări Adevărat/Fals** – Verificări conceptuale
- 4. Exerciții de Calcul** – Practică aplicată
- 5. Exerciții Python** – Practică de programare
- 6. Întrebări de Discuție** – Gândire critică
- 7. Exerciții cu asistență AI** – Analiză om vs. AI



## Formule Esențiale

### Descompunere:

- Aditivă:  $X_t = T_t + S_t + \varepsilon_t$
- Multiplicativă:  $X_t = T_t \times S_t \times \varepsilon_t$

### Netezire Exponențială:

- SES:  $\hat{X}_{t+1} = \alpha X_t + (1 - \alpha) \hat{X}_t$
- Holt: adaugă trend  $b_t$
- HW: adaugă sezonalitate  $S_t$

### Staționaritate:

- $\mathbb{E}[X_t] = \mu$  (constantă)
- $\text{Var}(X_t) = \sigma^2$  (constantă)
- $\text{Cov}(X_t, X_{t+h}) = \gamma(h)$

### Mers aleatoriu:

- $X_t = X_{t-1} + \varepsilon_t$
- $\text{Var}(X_t) = t\sigma^2$  (crește cu timpul)



## Sinteză: Concepte și Metode

Concept	Idee principală	Când se aplică
Descompunere aditivă	Amplitudine sezonieră constantă	Varianță stabilă
Descompunere multiplicativă	Sezonalitatea crește cu nivelul	Varianță în creștere
SES	Doar nivel ( $\alpha$ )	Fără trend, fără sezonalitate
Holt	Nivel + Trend ( $\alpha, \beta$ )	Trend, fără sezonalitate
Holt-Winters	Nivel + Trend + Sezonalitate	Trend și sezonalitate
Testul ADF	$H_0$ : rădăcină unitară	Test pentru nestaționaritate
Testul KPSS	$H_0$ : staționară	Confirmă staționaritatea
Diferențiere	Elimină trendul stochastic	Mers aleatoriu, rădăcină unitară
Regresie	Elimină trendul determinist	Trend liniar/polinomial



## Test 1: Bazele Seriilor de Timp

### Întrebare

Care dintre următoarele NU este o caracteristică a datelor de tip serie de timp?

- A. Observațiile sunt ordonate în timp
- B. Observațiile consecutive sunt de obicei corelate
- C. Observațiile sunt independente și identic distribuite
- D. Datele au o ordonare temporală naturală

*Răspunsul pe slide-ul următor...*



## Test 1: Răspuns

Răspuns: C – Observațiile sunt independente și identic distribuite

Întrebare: Care NU este o caracteristică a datelor de tip serie de timp?

- A. Observațiile sunt ordonate în timp ✗
  - B. Observațiile consecutive sunt de obicei corelate ✗
  - C. **Observațiile sunt independente și identic distribuite ✓**
  - D. Datele au o ordonare temporală naturală ✗
- 
- Observațiile seriilor de timp sunt **dependente** (autocorelate), nu independente
  - Ipoteza i.i.d. este fundamentală pentru analiza transversală, dar este **încălcată** în seriile de timp
  - Această dependență temporală necesită **metode specializede**



## Test 2: Descompunere

### Întrebare

Când ar trebui să folosiți descompunerea multiplicativă în loc de cea aditivă?

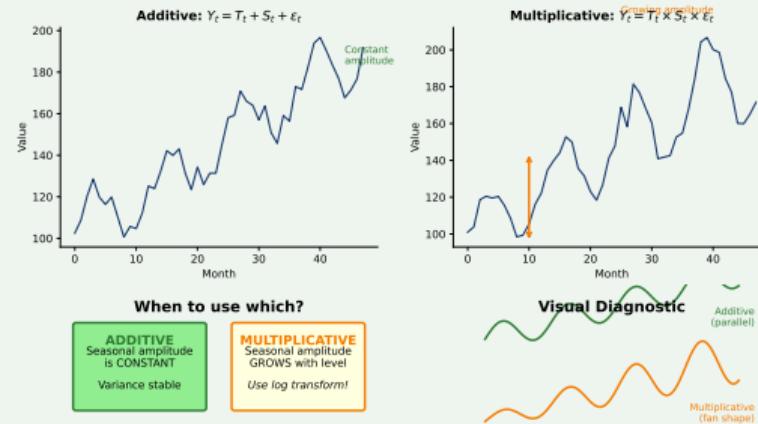
- A. Când modelul sezonier are amplitudine constantă
- B. Când varianța seriei este stabilă în timp
- C. Când fluctuațiile sezoniere cresc proporțional cu nivelul
- D. Când seria de timp nu are componentă de trend

*Răspunsul pe slide-ul următor...*



## Test 2: Răspuns

Răspuns: C – Când fluctuațiile sezoniere cresc proporțional cu nivelul



- Multiplicativă:**  $X_t = T_t \times S_t \times \varepsilon_t$
- Amplitudinea sezonieră **scalează cu nivelul** (model în evantai)



## Test 3: Netezire Exponențială

### Întrebare

În Netezirea exponențială Simplă cu  $\alpha = 0.9$ , ce se întâmplă?

- A. Prognozele sunt foarte netede și stabile
- B. Observațiile recente au foarte puțină pondere
- C. Prognozele reacționează rapid la schimbările recente
- D. Prognoza este în esență o medie pe termen lung

*Răspunsul pe slide-ul următor...*



## Test 3: Răspuns

Răspuns: C – Prognozele reacționează rapid la schimbările recente

$$\text{Cu } \alpha = 0.9: \hat{X}_{t+1} = 0.9X_t + 0.1\hat{X}_t$$

- $\alpha$  mare** (ex. 0.9): 90% pondere pe ultima observație
  - ▶ Prognoze foarte receptive la date noi
- $\alpha$  mic** (ex. 0.1): prognoze mai netede, mai stabile
  - ▶ Mediază peste mai mult istoric



## Test 4: Staționaritate

### Întrebare

Un proces de mers aleatoriu  $X_t = X_{t-1} + \varepsilon_t$  este:

- A. Strict staționar
- B. Slab staționar
- C. Nestăționar deoarece varianța crește cu timpul
- D. Staționar după adăugarea unei constante

*Răspunsul pe slide-ul următor...*



## Test 4: Răspuns

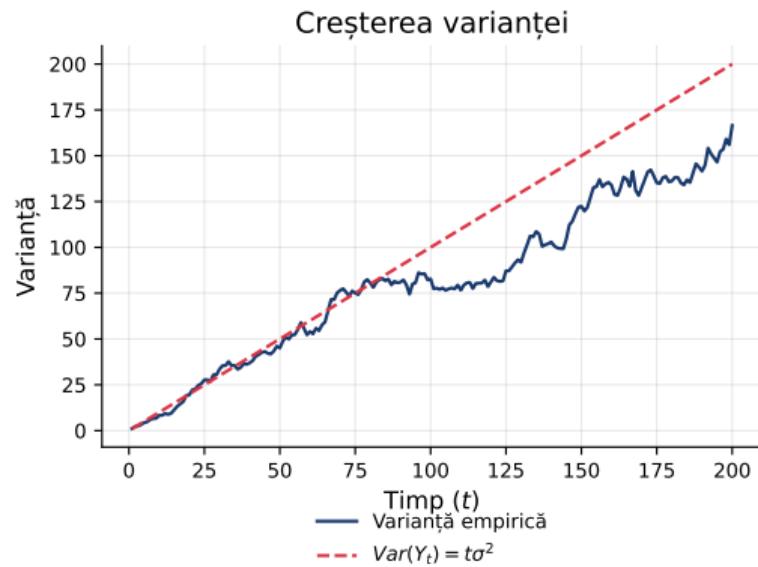
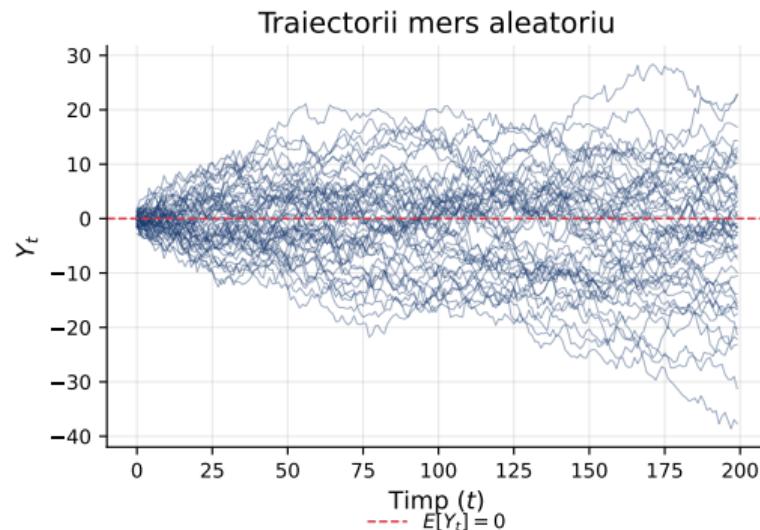
Răspuns: C – Nestaționar deoarece varianța crește cu timpul

Pentru mersul aleatoriu:  $X_t = \sum_{i=1}^t \varepsilon_i$

- $\mathbb{E}[X_t] = 0$  (medie constantă – OK)
- $\text{Var}(X_t) = t\sigma^2$  (varianța depinde de  $t$  – NU e OK!)
  - ▶ Varianța **nu** este constantă  $\Rightarrow$  încalcă staționaritatea
- Soluție:** diferențierea dă  $\Delta X_t = \varepsilon_t$  — staționară



## Vizual: Mers aleatoriu vs Staționar



- Traectoriile mersului aleatoriu rătăcesc imprevizibil
- Varianța crește liniar cu timpul  $\Rightarrow$  nestaționar



## Test 5: Teste pentru Rădăcină Unitară

### Întrebare

Rulați testele ADF și KPSS. ADF nu respinge  $H_0$ , iar KPSS respinge  $H_0$ . Ce concluzie rezultă?

- A. Seria este staționară
- B. Seria are o rădăcină unitară (nestaționară)
- C. Rezultatele sunt neconcludente
- D. Sunt necesare teste suplimentare

*Răspunsul pe slide-ul următor...*



## Test 5: Răspuns

Răspuns: B – Seria are o rădăcină unitară (nestaționară)

- ADF:  $H_0$  = rădăcină unitară. Nu respingem  $\Rightarrow$  evidență PENTRU rădăcină unitară
- KPSS:  $H_0$  = staționară. Respingem  $\Rightarrow$  evidență ÎMPOTRIVA staționarității
  - ▶ Ambele teste confirmă: seria este **nestaționară**
- Următorul pas: diferențiați seria înainte de a modela cu ARMA



## Test 6: Evaluarea prognozei

### Întrebare

Care metrică este adecvată pentru compararea acurateții prognozei între serii de timp cu scale diferite?

- A. Eroarea Absolută Medie (MAE)
- B. Rădăcina Erorii Medii Pătratice (RMSE)
- C. Eroarea Absolută Medie Procentuală (MAPE)
- D. Eroarea Medie Pătratică (MSE)

*Răspunsul pe slide-ul următor...*



## Test 6: Răspuns

Răspuns: C – Eroarea Absolută Medie Procentuală (MAPE)

$$\text{MAPE} = \frac{100}{n} \sum \left| \frac{e_t}{X_t} \right| \text{ exprimă erorile ca procente.}$$

- MAE, RMSE, MSE sunt **dependente de scală** (unități ale lui  $X_t$ )
- MAPE este **independentă de scală** (întotdeauna în %)
- Atenție: MAPE devine instabilă când  $X_t \approx 0$



## Test 7: Tipuri de Trend

### Întrebare

Un trend determinist poate fi eliminat prin:

- A. Diferențiere
- B. Regresie pe timp
- C. Ajustare sezonieră
- D. Netezire cu medie mobilă

*Răspunsul pe slide-ul următor...*



## Test 7: Răspuns

Răspuns: B – Regresie pe timp

- Trend determinist:**  $Y_t = \alpha + \beta t + \varepsilon_t$  ( $\beta$  fix)
- Metoda de eliminare:** regresie  $Y_t$  pe  $t$ , analizați reziduurile  $\hat{\varepsilon}_t$
- De ce nu diferențiere?**
  - ▶ Diferențierea dă  $\Delta Y_t = \beta + \Delta \varepsilon_t$  — elimină trendul dar lasă o constantă
  - ▶ Diferențierea este corectă doar pentru trenduri *stochastice* (rădăcini unitare)



## Test 8: interpretarea ACF

### Întrebare

Dacă ACF-ul unei serii de timp descrește foarte lent (rămâne semnificativ pentru multe lag-uri), aceasta sugerează:

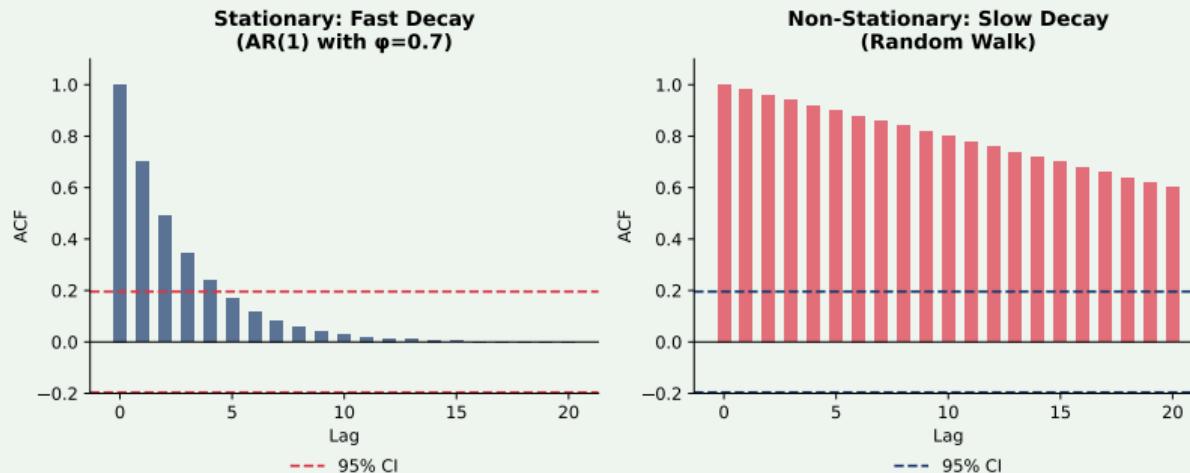
- A. Seria este zgomot alb
- B. Seria este probabil nestaționară
- C. Seria nu are autocorelație
- D. Seria este perfect predictibilă

*Răspunsul pe slide-ul următor...*



## Test 8: Răspuns

Răspuns: B – Seria este probabil nestaționară



- Staționară: ACF descrește rapid ( $\rho_k = \phi^k \rightarrow 0$ )
- Nestaționară: ACF rămâne aproape de 1  $\Rightarrow$  diferențiere necesară



## Test 9: Metoda Holt

### Întrebare

Netezirea exponențială Holt diferă de SES prin adăugarea:

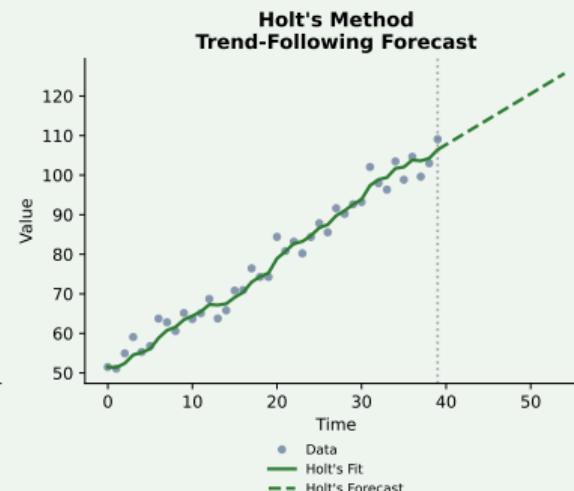
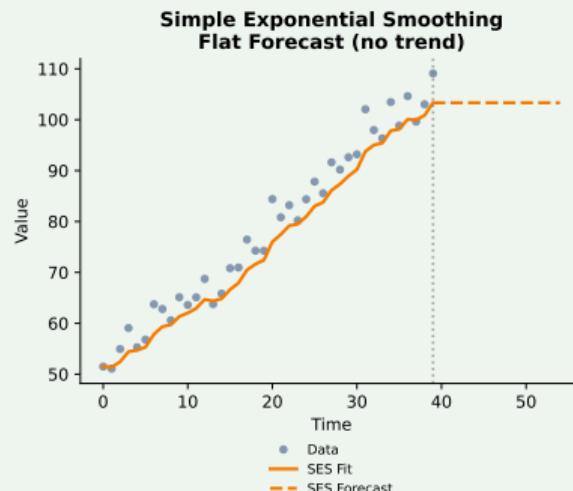
- A. O componentă sezonieră
- B. O componentă de trend
- C. O componentă ciclică
- D. O componentă neregulată

*Răspunsul pe slide-ul următor...*



## Test 9: Răspuns

Răspuns: B – O componentă de trend



- **Holt:**  $L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + b_{t-1}); \quad b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1}$
- **Prognoză:**  $\hat{Y}_{t+h} = L_t + h \cdot b_t$



## Test 10: Zgomot alb

### Întrebare

Care proprietate NU este necesară pentru ca un proces să fie zgomot alb?

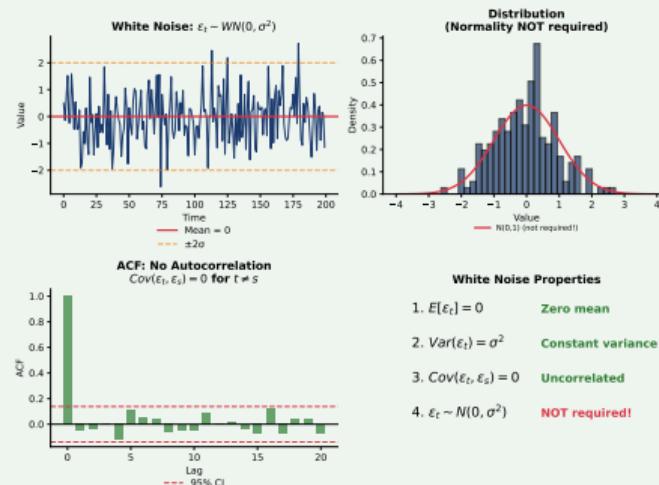
- A.  $\mathbb{E}[\varepsilon_t] = 0$
- B.  $\text{Var}(\varepsilon_t) = \sigma^2$  (constantă)
- C.  $\text{Cov}(\varepsilon_t, \varepsilon_s) = 0$  pentru  $t \neq s$
- D.  $\varepsilon_t \sim N(0, \sigma^2)$

*Răspunsul pe slide-ul următor...*



## Test 10: Răspuns

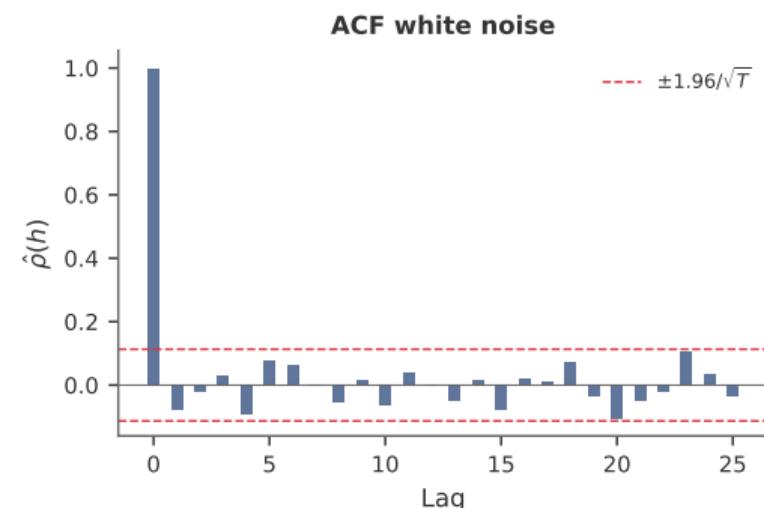
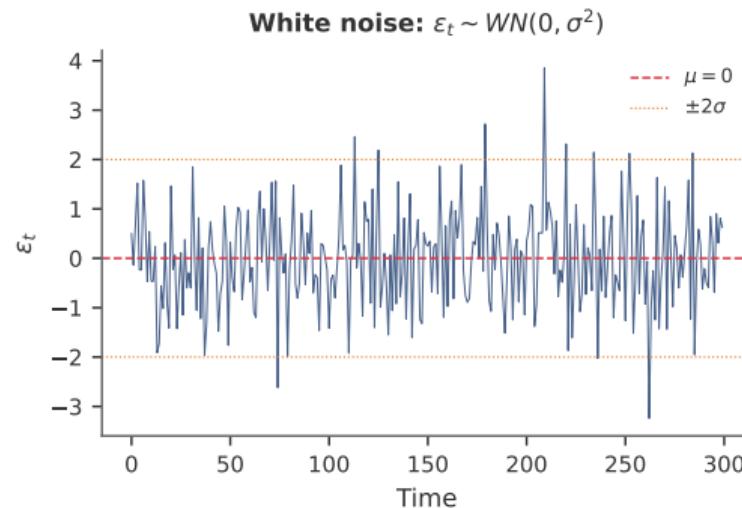
Răspuns: D – Normalitatea NU este necesară



- Zgomot alb:** medie zero, varianță constantă, necorelat
- Zgomot alb Gaussian:** adăugă normalitate  $\Rightarrow$  independent (nu doar necorelat)



## Vizual: Proprietățile Zgomotului Alb



- Stânga:** zgomotul alb fluctuează în jurul lui zero
- Dreapta:** ACF nu arată autocorelație (toate valorile  $\approx 0$  după lag 0)

## Test 11: Orizont de Prognoză

### Întrebare

Pe măsură ce orizontul de prognoză  $h$  crește, ce se întâmplă de obicei cu intervalele de prognoză?

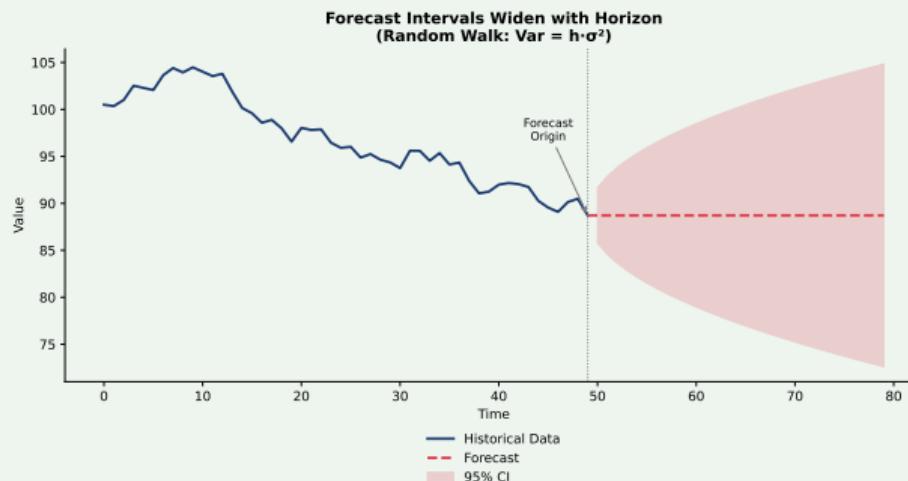
- A. Devin mai înguste
- B. Rămân la aceeași lățime
- C. Devin mai largi
- D. Dispar

*Răspunsul pe slide-ul următor...*



## Test 11: Răspuns

Răspuns: C – Devin mai largi



- Mers aleatoriu:  $\text{Var} = h\sigma^2$  (crește liniar)
- IC 95%:  $\hat{Y}_{t+h} \pm 1.96\sqrt{h}\sigma$  (se lărgește cu  $\sqrt{h}$ )



## Test 12: Detectarea Sezonalității

### Întrebare

ACF-ul arată vârfuri semnificative la lag-urile 12, 24 și 36 pentru date lunare. Aceasta sugerează:

- A. Fără sezonalitate
- B. Sezonialitate anuală
- C. Sezonialitate săptămânală
- D. Zgomot aleatoriu

*Răspunsul pe slide-ul următor...*



## Test 12: Răspuns

Răspuns: B – Sezonalitate anuală

- Recunoașterea modelului:**
  - ▶ Lag 12: corelație cu aceeași lună de anul trecut
  - ▶ Lag 24: aceeași lună de acum doi ani
  - ▶ Lag 36: aceeași lună de acum trei ani
- Perioada sezonieră:**  $s = 12$  (date lunare cu ciclu anual)
- Modele comune:** vânzări retail (decembrie), consum de energie (vară/iarnă), turism



## Test 13: Validare Încrucișată în Seriile de timp

### Întrebare

De ce nu putem folosi validarea încrucișată standard k-fold pentru seriile de timp?

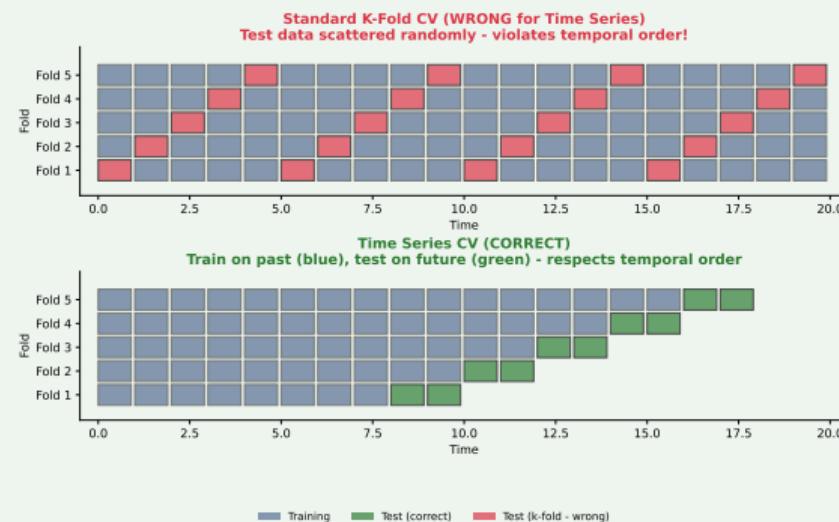
- A. Datele seriilor de timp sunt prea mici
- B. Ar încălca ordonarea temporală (viitorul prezicând trecutul)
- C. Validarea încrucișată este întotdeauna invalidă
- D. Seriile de timp nu au nevoie de validare

*Răspunsul pe slide-ul următor...*



## Test 13: Răspuns

Răspuns: B – Ar încălca ordonarea temporală



Principiu: datele viitoare nu pot fi folosite pentru a prezice trecutul! Se recomandă CV cu fereastră mobilă/în expansiune.



## Test 14: Limitarea MAPE

### Întrebare

MAPE (Eroarea Absolută Medie Procentuală) NU ar trebui folosită când:

- A. Comparați modele pe același set de date
- B. Valorile reale pot fi zero sau aproape de zero
- C. Prognozați prețuri de acțiuni
- D. Datele au un trend

*Răspunsul pe slide-ul următor...*



## Test 14: Răspuns

Răspuns: B – Când valorile reale pot fi zero sau aproape de zero

- Formula:**  $MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right|$
- Problema:**  $Y_t \approx 0 \Rightarrow MAPE \rightarrow \infty$
- Alternative:**
  - ▶ **SMAPE:**  $\frac{200\%}{n} \sum \frac{|Y_t - \hat{Y}_t|}{|Y_t| + |\hat{Y}_t|}$  (mărginită 0–200%)
  - ▶ **MASE:**  $\frac{1}{n} \sum \frac{|e_t|}{\frac{1}{n-1} \sum |Y_t - Y_{t-1}|}$  (fără scală)



## Adevărat sau Fals? (Setul 1)

### Întrebare

Marcați fiecare afirmație ca Adevărat (A) sau Fals (F):

1. O serie de timp cu medie constantă este întotdeauna staționară.

\_\_\_\_\_

2. Varianța unui mers aleatoriu crește liniar cu timpul.

\_\_\_\_\_

3. Prognozele SES sunt întotdeauna plate (constante pentru toate orizonturile).

\_\_\_\_\_

4. Testele ADF și KPSS au aceeași ipoteză nulă.

\_\_\_\_\_

5. RMSE mai mic înseamnă întotdeauna prognoze mai bune.

\_\_\_\_\_

6. Autocorelația la lag 0 este întotdeauna egală cu 1.

\_\_\_\_\_

*Răspunsul pe slide-ul următor...*



## Adevărat sau Fals: Răspunsuri (Setul 1)

### Răspunsuri

1. Medie constantă  $\Rightarrow$  staționară.
2. Varianța mersului aleatoriu crește liniar cu timpul.
3. Prognozele SES sunt plate (constante  $\forall h$ ).
4. ADF și KPSS au aceeași ipoteză nulă.
5. RMSE mai mic  $\Rightarrow$  prognoze mai bune.
6.  $\rho(0) = 1$  întotdeauna.

**FALS** — Trebuie și varianță constantă și covarianță dependentă doar de lag.

**ADEVĂRAT** —  $\text{Var}(X_t) = t\sigma^2$ .

**ADEVĂRAT** — SES nu are trend:  $\hat{X}_{t+h} = L_t$  pentru orice  $h$ .

**FALS** — ADF:  $H_0$  = rădăcină unitară. KPSS:  $H_0$  = staționară. Opuse!

**FALS** — Dependenta de scală; poate supraajusta la valori extreme.

**ADEVĂRAT** —  $\rho(0) = \gamma(0)/\gamma(0) = 1$  prin definiție.



## Adevărat sau Fals? (Setul 2)

### Întrebare

Marcați fiecare afirmație ca Adevărat (A) sau Fals (F):

1. ACF-ul unui proces AR(1) staționar descrește exponențial.

\_\_\_\_\_

2. Zgomotul alb este întotdeauna distribuit normal.

\_\_\_\_\_

3. Diferențierea poate face o serie nestaționară să devină staționară.

\_\_\_\_\_

4. PACF-ul unui proces MA(1) se anulează după lag 1.

\_\_\_\_\_

5. Ar trebui să folosiți întotdeauna setul de test pentru ajustarea hiperparametrilor.

\_\_\_\_\_

6. Holt-Winters este potrivit pentru date fără sezonalitate.

\_\_\_\_\_

*Răspunsul pe slide-ul următor...*



## Adevărat sau Fals: Răspunsuri (Setul 2)

### Răspunsuri

1. ACF-ul unui AR(1) staționar descrește exponențial.
2. Zgomotul alb este întotdeauna distribuit normal.

autocorelație. Gaussian = caz special.

3. Diferențierea poate face o serie nestaționară staționară.
4. PACF-ul unui MA(1) se anulează după lag 1.
5. Setul de test trebuie folosit pentru ajustarea hiperparametrilor.
6. Holt-Winters este potrivit pentru date fără sezonalitate.

**ADEVĂRAT** —  $\rho(h) = \phi^h$ , descrește exponențial.

**FALS** — Necesară doar medie zero, varianță constantă, fără

**ADEVĂRAT** — Elimină trendurile stochastice (rădăcinile unitare).

**FALS** — ACF-ul se anulează pentru MA. PACF descrește exponențial.

**FALS** — Setul de validare pentru ajustare; setul de test doar pentru evaluarea finală.

**FALS** — Fără sezonalitate: metoda Holt sau SES.



## Exercițiu 1: Netezire Exponențială Simplă

**Enunț:** Date fiind următoarele date și  $\alpha = 0.3$ :

$t$	1	2	3	4	5
$X_t$	10	12	11	14	13

Cu valoarea inițială  $\hat{X}_1 = X_1 = 10$ , determinați:

- a) Prognozele  $\hat{X}_2, \hat{X}_3, \hat{X}_4, \hat{X}_5$
- b) Prognoza pentru  $t = 6$ :  $\hat{X}_6$
- c) Erorile de prognoză  $e_t = X_t - \hat{X}_t$  pentru  $t = 2, 3, 4, 5$
- d) MAE și RMSE

**Formula:**  $\hat{X}_{t+1} = \alpha X_t + (1 - \alpha)\hat{X}_t$



## Exercițiu 1: Soluție

Folosind  $\hat{X}_{t+1} = 0.3X_t + 0.7\hat{X}_t$ :

$t$	1	2	3	4	5	6
$X_t$	10	12	11	14	13	?
$\hat{X}_t$	10	10	10.6	10.72	11.70	<b>12.09</b>
$e_t$	-	2	0.4	3.28	1.30	-

Calcule:

- $\hat{X}_2 = 0.3(10) + 0.7(10) = 10$
- $\hat{X}_3 = 0.3(12) + 0.7(10) = 10.6$
- $\hat{X}_4 = 0.3(11) + 0.7(10.6) = 10.72$
- $\hat{X}_5 = 0.3(14) + 0.7(10.72) = 11.70$
- $\hat{X}_6 = 0.3(13) + 0.7(11.70) = \mathbf{12.09}$

$$\text{MAE} = \frac{|2| + |0.4| + |3.28| + |1.30|}{4} = 1.745 \quad \text{RMSE} = \sqrt{\frac{4 + 0.16 + 10.76 + 1.69}{4}} = 2.04$$

## Exercițiu 2: Autocovarianță

**Enunț:** Pentru un proces staționar cu:

- $\mathbb{E}[X_t] = 5$
- $\gamma(0) = 4$  (variantă)
- $\gamma(1) = 2$
- $\gamma(2) = 1$

Calculați:

- a) Funcția de autocorelație  $\rho(0), \rho(1), \rho(2)$
- b)  $\text{Cov}(X_t, X_{t-1})$
- c)  $\text{Corr}(X_5, X_7)$
- d) Dacă  $X_t = 6$ , care este  $\mathbb{E}[X_{t+1}|X_t = 6]$  presupunând AR(1)?

## Exercițiu 2: Soluție

a) Autocorelații:

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}$$

- $\rho(0) = \gamma(0)/\gamma(0) = 1$
- $\rho(1) = \gamma(1)/\gamma(0) = 2/4 = 0.5$
- $\rho(2) = \gamma(2)/\gamma(0) = 1/4 = 0.25$

b)  $\text{Cov}(X_t, X_{t-1}) = \gamma(1) = 2$  (prin staționaritate, covarianța la lag 1)

c)  $\text{Corr}(X_5, X_7) = \rho(|7 - 5|) = \rho(2) = 0.25$

d) Pentru AR(1) cu  $\phi = \rho(1) = 0.5$ :

$$\mathbb{E}[X_{t+1}|X_t] = \mu + \phi(X_t - \mu) = 5 + 0.5(6 - 5) = 5.5$$



### Exercițiu 3: Proprietățile Mersului Aleatoriu

**Enunț:** Considerați un mers aleatoriu  $X_t = X_{t-1} + \varepsilon_t$  unde  $\varepsilon_t \sim WN(0, 4)$  și  $X_0 = 100$ .

Calculați:

- a)  $\mathbb{E}[X_{10}]$
- b)  $\text{Var}(X_{10})$
- c)  $\text{Cov}(X_5, X_{10})$
- d) Intervalul de încredere de 95% pentru  $X_{100}$
- e) Dacă  $X_5 = 108$ , care este prognoza optimală pentru  $X_6$ ?



### Exercițiu 3: Soluție

**Mers aleatoriu:**  $X_t = X_0 + \sum_{i=1}^t \varepsilon_i$  cu  $\sigma^2 = 4$

a)  $\mathbb{E}[X_{10}] = X_0 = 100$  (media rămâne la valoarea de pornire)

b)  $\text{Var}(X_{10}) = 10 \times \sigma^2 = 10 \times 4 = 40$

c)  $\text{Cov}(X_5, X_{10}) = \min(5, 10) \times \sigma^2 = 5 \times 4 = 20$

d) Pentru  $X_{100}$ :

$\mathbb{E}[X_{100}] = 100$ ,  $\text{Var}(X_{100}) = 400$ ,  $SD = 20$

IC 95%:  $100 \pm 1.96 \times 20 = [60.8, 139.2]$

e) Prognoza optimală:  $\hat{X}_6 = X_5 = 108$

(Proprietate a mersului aleatoriu: prognoza optimală este ultima valoare observată)

## Exercițiu Python 1: Import și Vizualizare



**Cerintă:** Importați datele S&P 500 și realizați un grafic de bază al seriei de timp.

### Cod inițial

```
import yfinance as yf
import matplotlib.pyplot as plt
sp500 = yf.download('^GSPC', start='2020-01-01', end='2025-01-01')
# TODO: Reprezentați grafic prețurile de închidere
# TODO: Adăugați titlu și etichete
# TODO: Calculați și afișați statistici de bază
```

### Întrebări:

1. Care este media și deviația standard a randamentelor?
2. Seria pare staționară? Argumentați.



## Exercițiu Python 2: Descompunere

**Cerință:** Aplicați descompunerea STL pe datele privind pasagerii aerieni.

### Cod inițial

```
from statsmodels.tsa.seasonal import STL  
import pandas as pd  
  
# Încărcați pasagerii aerieni  
url = 'https://raw.githubusercontent.com/..../airline.csv'  
airline = pd.read_csv(url, parse_dates=['Month'],  
                      index_col='Month')  
  
# TODO: Aplicați descompunerea STL cu period=12  
# TODO: Reprezentați grafic toate componentele  
# TODO: Ce procent din varianță este explicat de trend?
```

**Indicație:** `STL(data, period=12).fit()`



## Exercițiu Python 3: Netezire Exponențială

**Cerintă:** Comparați metodele SES, Holt și Holt-Winters pe date reale.

### Cod inițial

```
from statsmodels.tsa.holtwinters import (SimpleExpSmoothing,
                                           ExponentialSmoothing)

# Împărțiți datele: 80% antrenare, 20% test
train = airline[:'1958']
test = airline['1959':]

# TODO: Ajustați SES, Holt și Holt-Winters
# TODO: Generați prognoze pentru perioada de test
# TODO: Calculați RMSE pentru fiecare metodă
# TODO: Care metodă are cele mai bune performanțe? De ce?
```



## Exercițiu Python 4: Testarea Staționarității



**Cerință:** Verificați staționaritatea prin testele ADF și KPSS.

### Cod inițial

```
from statsmodels.tsa.stattools import adfuller, kpss
prices = sp500['Close']
returns = prices.pct_change().dropna()
# TODO: Rulați testul ADF pe prețuri și randamente
# TODO: Rulați testul KPSS pe prețuri și randamente
# TODO: Interpretăți rezultatele
# ADF: adfuller(series) | KPSS: kpss(series, regression='c')
```

### Întrebări:

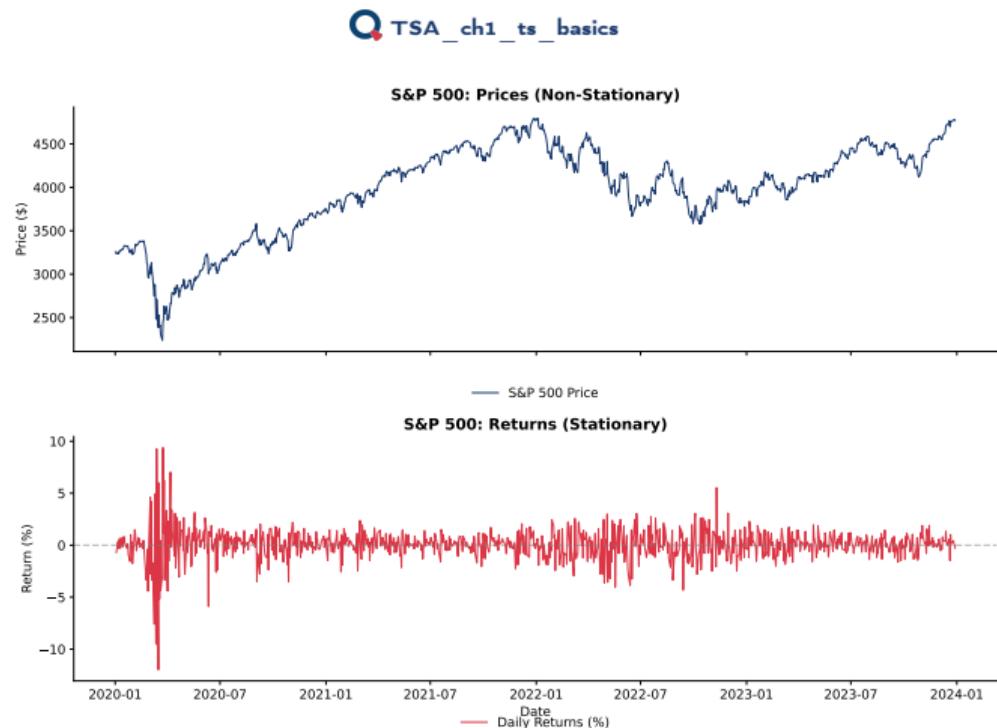
1. Prețurile sunt staționare? Randamentele sunt staționare?
2. Rezultatele ADF și KPSS converg?



## Studiu de Caz: Indicele S&P 500

### Observații

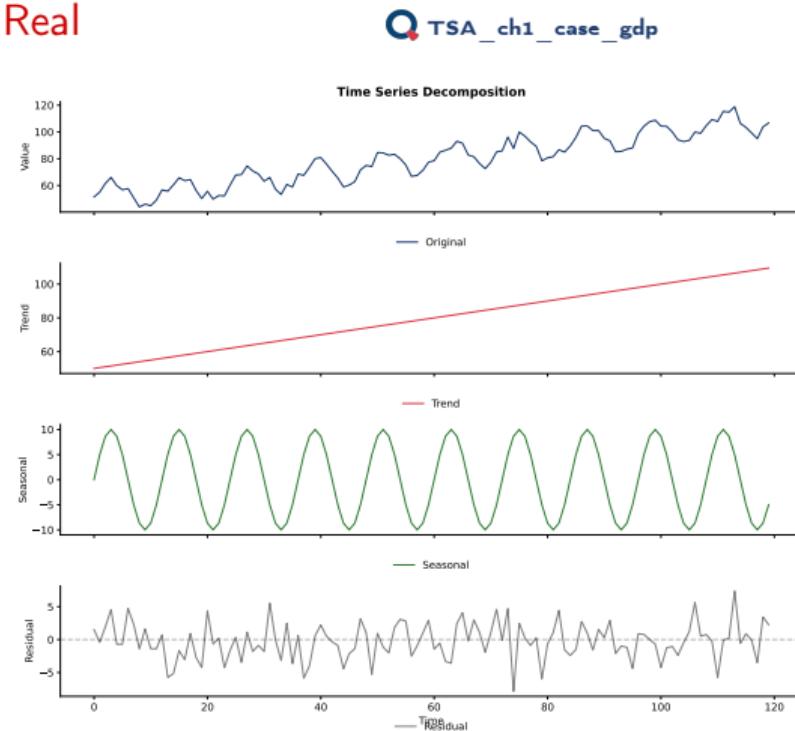
- Sus:** Prețuri S&P 500 — trend ascendent clar (nestaționar)
- Jos:** Randamente  $r_t = \log(P_t/P_{t-1})$  — staționare
- Fluctuații în jurul mediei zero
- Grupare a volatilității vizibilă



## Descompunerea Seriilor de Timp: Exemplu Real

### Observații

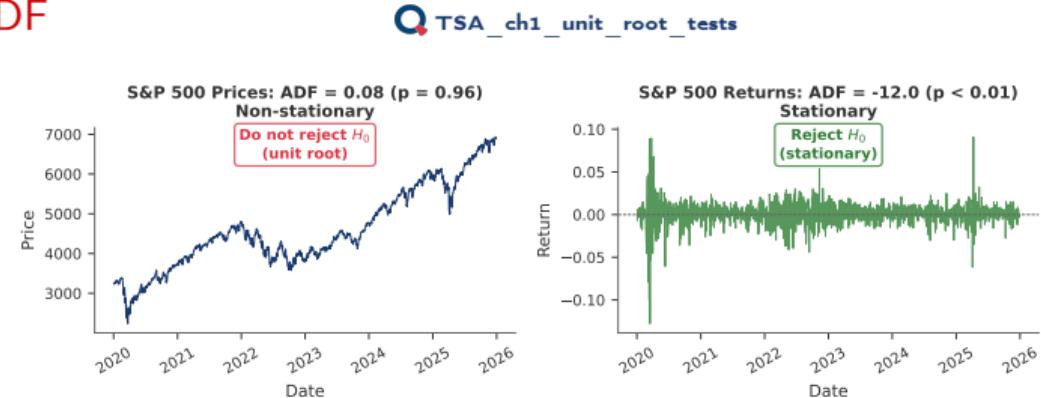
- **Trend:** Direcția pe termen lung
- **Sezonalitate:** Modele periodice regulate
- **Reziduu:** Ce rămâne după eliminarea trendului și sezonalității
- Descompunerea  $\Rightarrow$  înțelegerea structurii înainte de modelare



## Testarea Staționarității: Rezultate ADF

### Observații

- ADF compară statistica de test cu valorile critice
- Stat. test < val. critică  $\Rightarrow$  respingem  $H_0$  (staționară)
- **Preturi:** ADF  $> -2.86 \Rightarrow$  nestăționară
- **Randamente:** ADF  $< -2.86 \Rightarrow$  staționară



## Comparație Staționaritate: Prețuri vs Randamente

### Rezultate Test ADF

Serie	Statistică ADF	valoare-p	Concluzie
Prețuri S&P 500	-0.82	0.812	Nestaționară
Randamente S&P 500	-45.3	< 0.001	Staționară

### Observație Cheie

Prețurile financiare sunt de obicei  $I(1)$  – integrate de ordinul 1.

Luând diferențe de ordinul întâi (randamente) se obține staționaritatea.

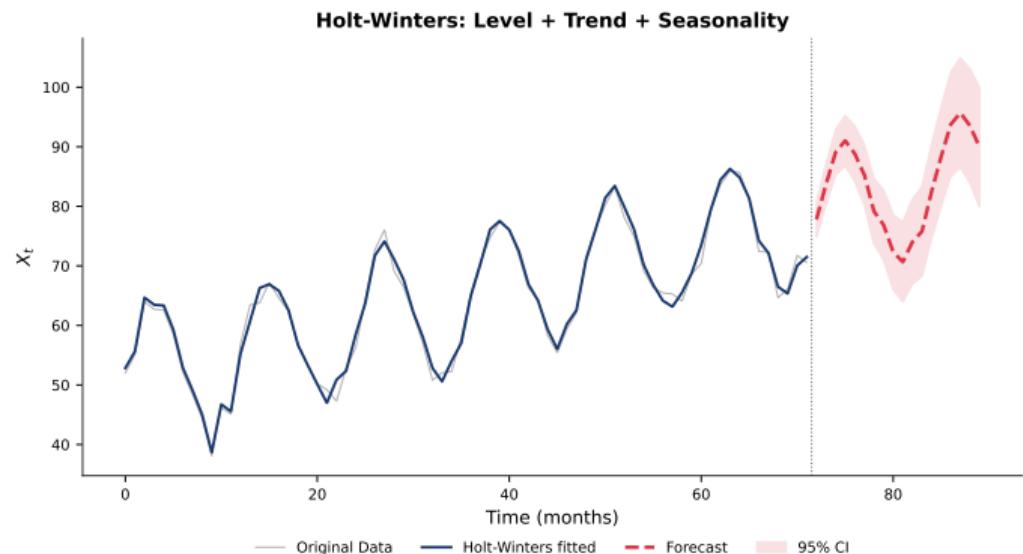
De aceea modelăm **randamentele**, nu prețurile!



## Prognoză cu Netezire Exponențială

### Observații

- Holt-Winters: date cu trend + sezonalitate
- $\alpha, \beta, \gamma$  controlează adaptabilitatea
- Captează continuarea trendului + model sezonier
- Simplă dar eficientă pentru aplicații de business



## Întrebare de Discuție 1

### Scenariu

- Date lunare de vânzări pentru o companie de retail
- Sezonialitate clară (vânzări ridicate în decembrie) + trend ascendent
- Vârfurile sezoniere au devenit mai mari în timp

### Discutați:

1. Descompunere aditivă sau multiplicativă? Argumentați.
2. Ce metodă de netezire exponențială ați recomanda?
3. Cum evaluați performanța prognozei?
4. Ce riscuri implică alegerea greșită a descompunerii?



## Întrebare de Discuție 2

### Scenariu

Un coleg afirmă:

- „Am rulat testul ADF pe datele mele de prețuri de acțiuni”
- „Am obținut o valoare-p de 0.65”
- „Deci datele mele sunt staționare și pot ajusta direct un model ARMA”

### Discuții:

1. Unde greșește raționamentul?
2. Care sunt ipotezele testului ADF?
3. Ce pași ar trebui urmați înainte de a estima un model ARMA?
4. Ce rol joacă testul KPSS în clarificarea situației?



## Întrebare de Discuție 3

### Scenariu

- Construiți un model de prognoză: MAPE de 2%
- Managerul este impresionat și vrea implementare imediată

### Discuții:

1. Ce verificări sunt necesare înainte de implementare?
2. Împărțirea antrenare/validare/test este corectă?
3. Există riscul contaminării datelor (*data leakage*)?
4. Ce verificări suplimentare sunt necesare?
5. Cum monitorizați performanța modelului în producție?



## Întrebare de Discuție 4

### Scenariu

Prognoza cererii zilnice de electricitate pentru săptămâna următoare:

- Modele zilnice puternice (vârfuri la ora 18)
- Modele săptămânale (mai scăzut în weekend)
- Modele anuale (mai ridicat vara/iarna)

### Discuții:

1. Cum abordați sezonalitatea multiplă?
2. Este Holt-Winters adecvat? Argumentați.
3. Care este avantajul termenilor Fourier în acest caz?
4. Cum organizați eșantioanele de antrenare/validare/test?



## AI în Analiza Serilor de Timp

### De ce folosim instrumente AI în acest curs?

Asistenții AI (Claude, ChatGPT, GitHub Copilot) pot genera cod și analize. Rolul vostru este să **evaluați, interpretați și criticați** — competențe pe care AI-ul nu le poate înlocui.

#### Obiective de învățare:

- Scrieți prompturi precise pentru sarcini econometrice
- Identificați erori în analiza statistică generată de AI
- Interpretăți rezultatele critic folosind conceptele din curs
- Comparați soluțiile AI cu metodologia manuală

#### Important

AI-ul este un **instrument**, nu un substitut pentru înțelegere. Trebuie să puteți explica *de ce* fiecare pas este corect.



## Exercițiu AI 1: Ingineria prompturilor

### Sarcina

Scriți un prompt pentru un asistent AI care să realizeze descompunerea STL pe datele lunare de pasageri aerienii. Apoi evaluați rezultatul.

### Comparați aceste prompturi:

1. „Descompune această serie de timp” Vag
2. „Aplică descompunerea STL pe datele lunare de pasageri cu period=12, trasează componenta de trend, sezonieră și reziduală” Precis

### Evaluati rezultatul AI:

- Perioada sezonieră este corectă?
- Descompunerea este aditivă sau multiplicativă? Este adecvată?
- Reziduurile sunt aproximativ zgomot alb?
- Ați alege o altă metodă? De ce?



## Exercițiu AI 2: Găsiți erorile

### Scenariu

Un asistent AI a produs această analiză a prețurilor acțiunilor:

1. A trasat seria de prețuri brute
2. A calculat ACF — a găsit autocorelație puternică
3. A concluzionat: „autocorelație ridicată = prognoză bună”
4. A ajustat Holt-Winters cu  $\alpha = 0.9$ ,  $\beta = 0.1$
5. A raportat MAPE = 1.2% pe setul de antrenare

### Identificați toate erorile:

1. AI-ul a testat staționaritatea? Ce test este necesar?
2. Este corect „ACF ridicat  $\Rightarrow$  prognoză bună”?
3. Ce implică  $\alpha = 0.9$  despre netezire?
4. De ce MAPE pe setul de antrenare este înselător?
5. Ce ar trebui făcut diferit?



## Exercițiu AI 3: Om vs. AI — Prognoză

### Sarcina

Folosiți setul de date AirPassengers din statsmodels.

#### Pasul 1 — Abordare manuală:

- Examinați datele vizual (trend, sezonalitate)
- Alegeti tipul de descompunere cu justificare
- Aplicați Holt-Winters, selectați parametrii
- Evaluati pe setul de test (ultimele 24 de luni)

#### Pasul 2 — Abordare AI:

- Cereti unui AI să „prognozeze pasagerii aerienii pe 24 de luni”
- Înregistrați metodologia și codul AI-ului

#### Pasul 3 — Comparați și reflectați:

- Care abordare a obținut RMSE/MAPE mai mic?
- AI-ul a verificat staționaritatea? A folosit o împărțire corectă?
- Ce ati învățat voi ce AI-ul a omis (sau invers)?



## Concluzii

- Seriile de timp sunt dependente**
  - ▶ Nu i.i.d. ca datele transversale — autocorelația este cheia
- Alegeți descompunerea cu înțelepciune**
  - ▶ Multiplicativă când amplitudinea sezonieră crește cu nivelul
- Înțelegeți parametrii de netezire**
  - ▶  $\alpha$  mare = reactiv,  $\alpha$  mic = neted
- Testați staționaritatea**
  - ▶ Folosiți atât ADF cât și KPSS împreună
- Evaluare corectă**
  - ▶ Nu ajustați niciodată pe setul de test!
- Mersul aleatoriu este nestaționar**
  - ▶ Varianța crește cu timpul:  $\text{Var}(X_t) = t\sigma^2$

### Următorul Seminar

Identificarea, estimarea și prognoza modelelor ARMA/ARIMA



## Surse de Date și Software

### Instrumente Software:

- statsmodels – Modele statistice pentru Python
- pandas – Manipulare date și serii de timp
- matplotlib, seaborn – Vizualizare
- scipy – Funcții statistice

### Date și Exemple:

- Serii de timp simulate pentru ilustrații
- Exemple bazate pe Hyndman & Athanasopoulos (2021)



## 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> — Repository de cod pentru statistică
- **Quantinar:** <https://quantinar.com> — Platformă de învățare metode cantitative
- **GitHub TSA:** [https://github.com/QuantLet/TSA/tree/main/TSA\\_Ch1](https://github.com/QuantLet/TSA/tree/main/TSA_Ch1) — Cod Python pentru acest seminar



# Vă Mulțumim!

Întrebări?

*Graficele au fost generate folosind Python (statsmodels, matplotlib)*

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

