



Analiza și Prognoza Seriilor de Timp

# Capitolul 1: Introducere în Serii de Timp

Fundamente și Concepte



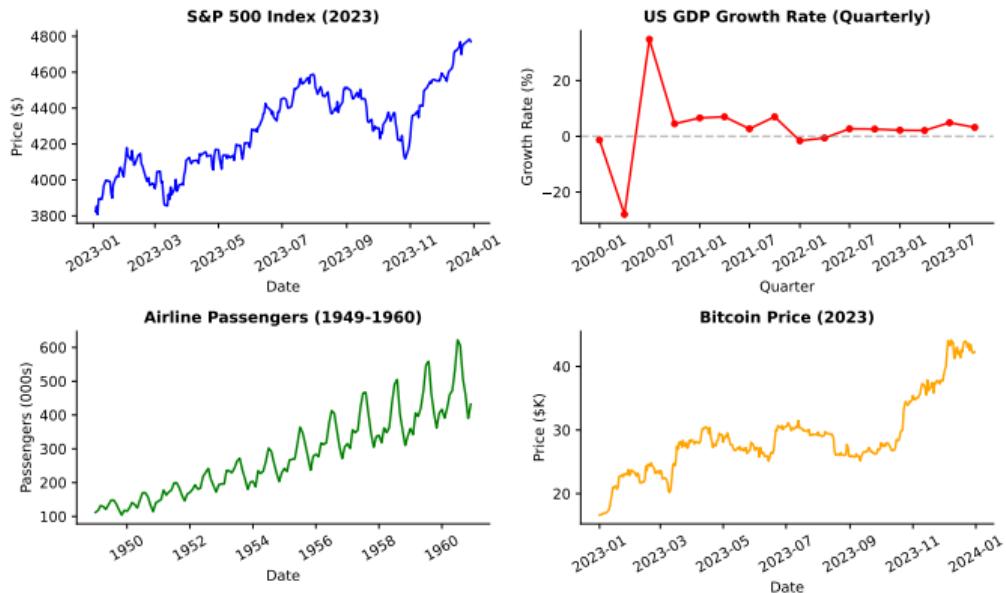
### La sfârșitul acestui capitol, veți fi capabili să:

1. Definiți seriile de timp și să le distingeți de datele transversale și de panel
2. Descompuneți seriile de timp în componente de trend, sezonalitate și reziduuri
3. Aplicați netezirea exponențială (SES, Holt, Holt-Winters, ETS)
4. Evaluați prognozele folosind MAE, RMSE, MAPE; separări train/validare/test
5. Modelați sezonalitatea folosind variabile dummy sau termeni Fourier
6. Gestionăți trendul și sezonalitatea prin eliminarea trendului și ajustare sezonieră
7. Înțelegeți procesele stochastice și staționaritatea
8. Calculați ACF/PACF și efectuați teste de staționaritate (ADF, KPSS)

# Structura Capitolului

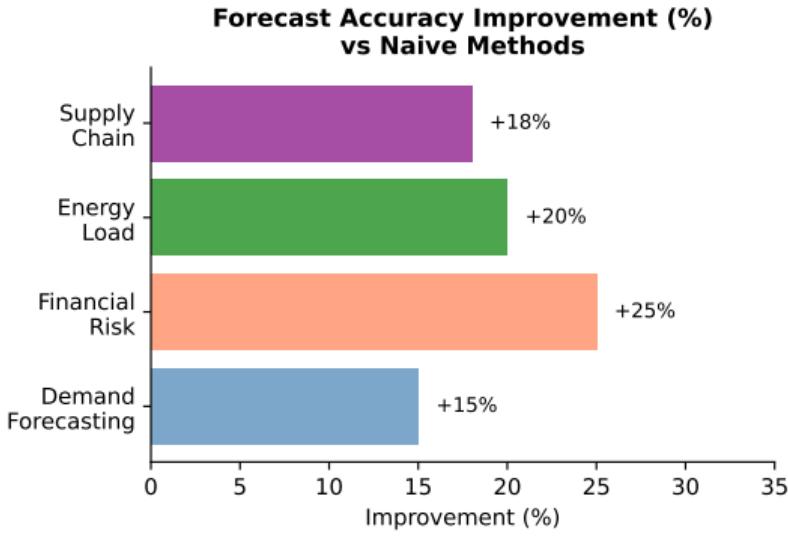
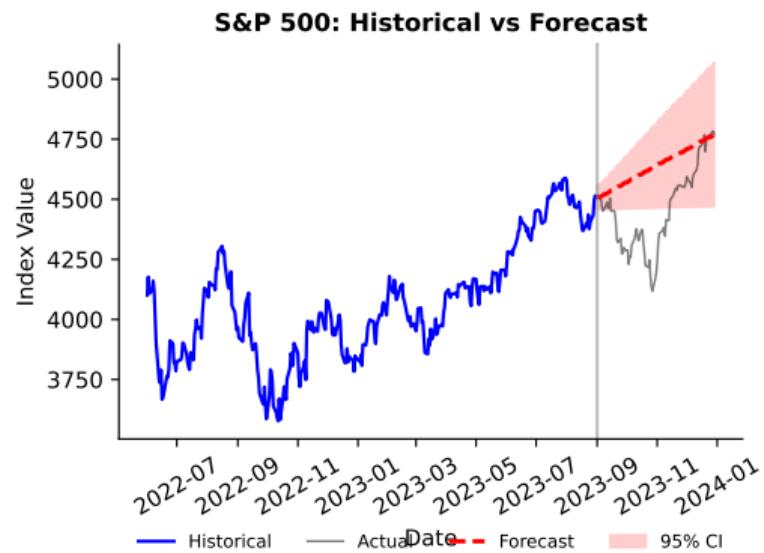
- 1 Ce Este o Serie de Timp?
- 2 Descompunerea Seriilor de Timp
- 3 Metode de Netezire Exponențială
- 4 Evaluarea Prognozei
- 5 Modelarea Sezonalității
- 6 Gestionarea Trendului și Sezonalității
- 7 Procese Stochastice
- 8 Staționaritatea
- 9 Zgomot Alb și Mers Aleatoriu
- 10 Funcții de Autocorelație
- 11 Operatorul Lag și Diferențierea
- 12 Testarea Staționarității
- 13 Aplicație pe Date Financiare
- 14 Rezumat
- 15 Quiz

# Seriile de Timp Sunt Prețutindeni



- **Finanțe**: Prețuri de acțiuni, cursuri de schimb, volume de tranzacționare
- **Economie**: PIB, șomaj, rate ale inflației
- **Afaceri**: Vânzări, trafic web, cererea clienților
- **Știință**: Temperatură, niveluri de poluare, indicatori vitali ai pacienților

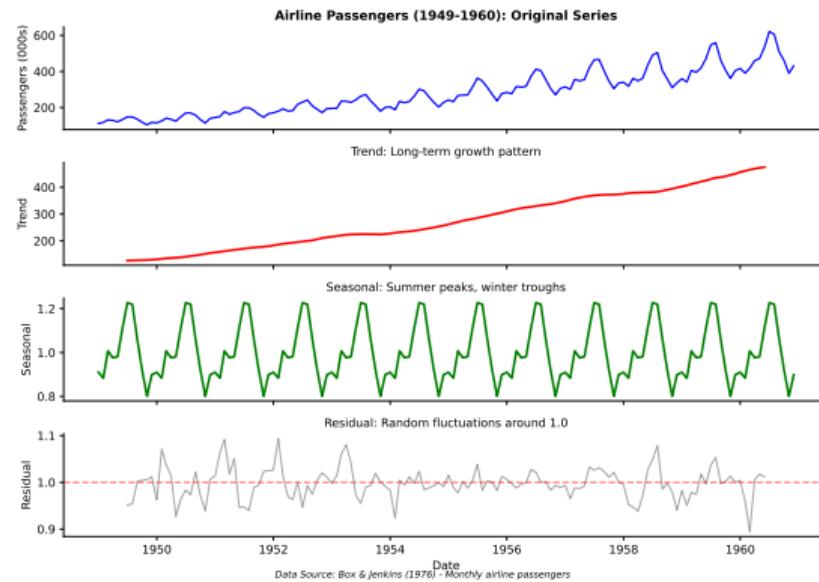
# De Ce Studiem Seriile de Timp?



## Obiectivul Principal: Prognoza

Folosiți tiparele istorice pentru a prezice valorile viitoare — esențial pentru planificarea afacerilor, gestionarea risurilor și decizii de politică.

# Înțelegerea Structurii Seriilor de Timp



## Descompunere

Orice serie de timp poate fi descompusă în componente interpretabile: trend, sezonialitate și zgromadire.

## Definiție 1 (Serie de Timp)

O serie de timp este o secvență de observații  $\{X_t\}$  indexate după timp:

$$\{X_t : t \in \mathcal{T}\}$$

unde  $\mathcal{T}$  este un set de indici reprezentând puncte temporale.

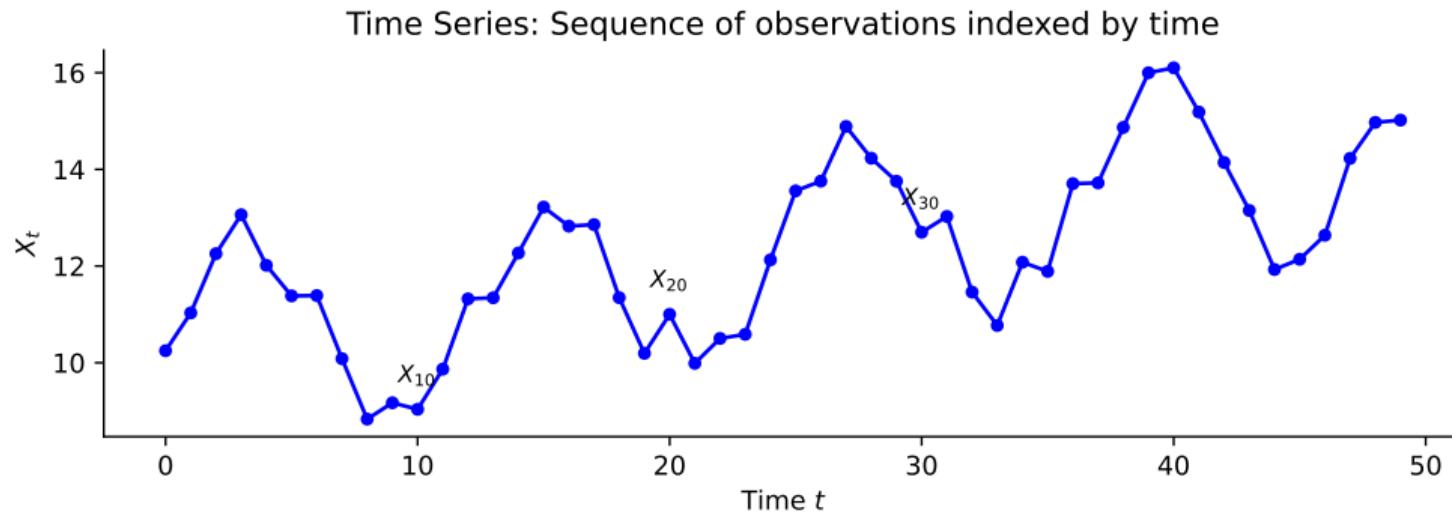
### Caracteristici cheie:

- **Ordonate:** Observațiile au o ordine temporală naturală
- **Dependente:** Observațiile consecutive sunt de obicei corelate
- **Discrete sau Continue:** Indexul temporal poate fi discret ( $t = 1, 2, 3, \dots$ ) sau continuu

### Notație:

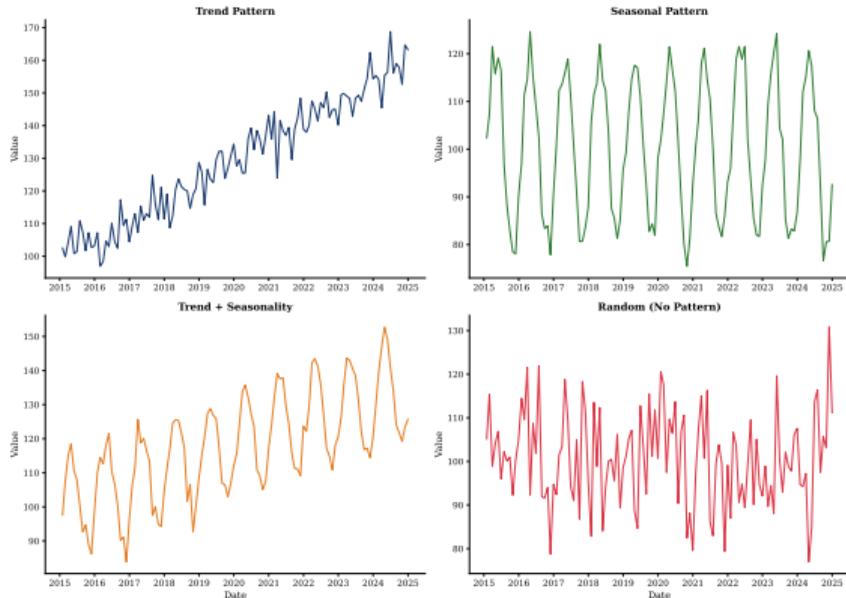
- $X_t$  = observația la momentul  $t$
- $\{X_t\}_{t=1}^T$  = serie de timp finită cu  $T$  observații

## Serie de Timp: Ilustrație Vizuală



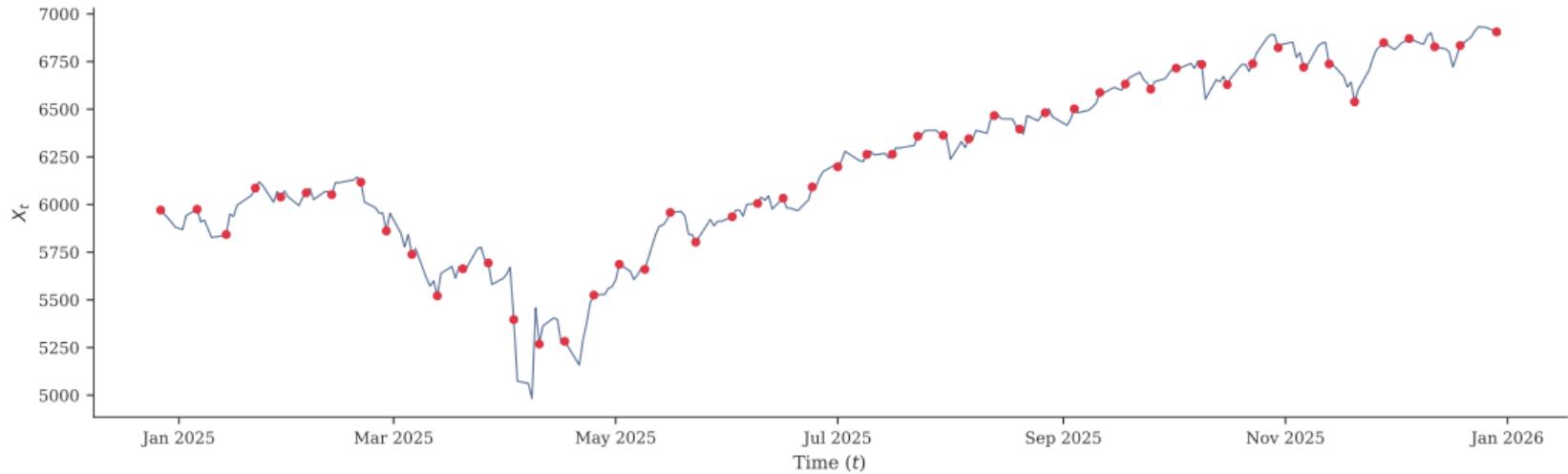
Fiecare punct  $X_t$  reprezintă o observație la momentul  $t$ . Secvența este ordonată și observațiile consecutive sunt de obicei corelate.

# Tipare Comune în Seriile de Timp



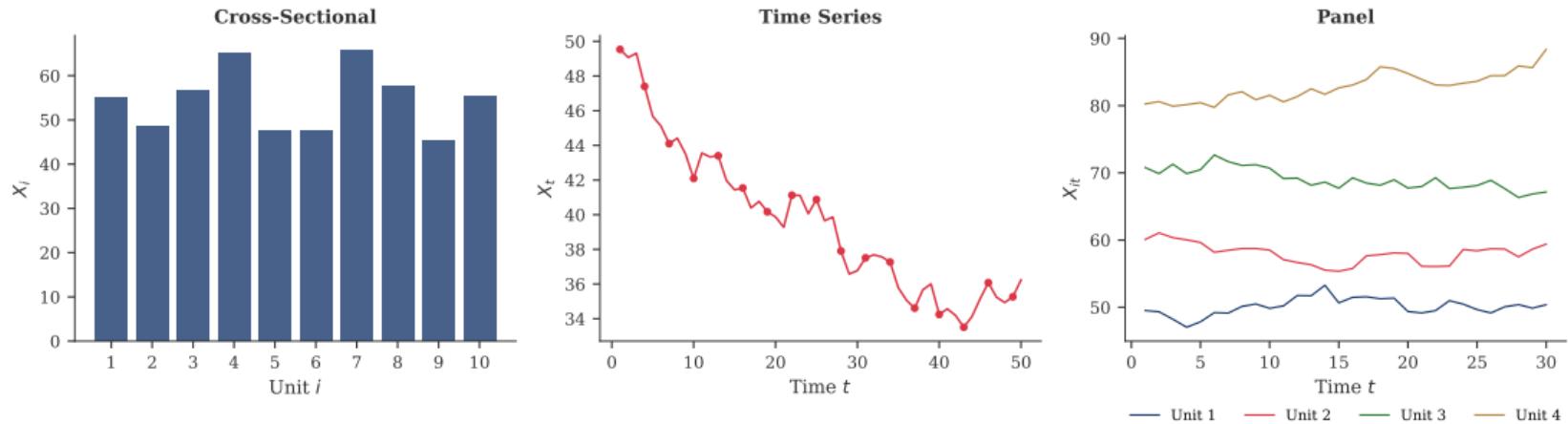
- **Trend:** Creștere sau scădere pe termen lung a datelor
- **Sezonalitate:** Tipare periodice regulate (de ex., lunar, trimestrial)
- **Aleatoriu:** Niciun tipar sistematic – fluctuații imprevizibile

## Serie de Timp: Definiție Vizuală



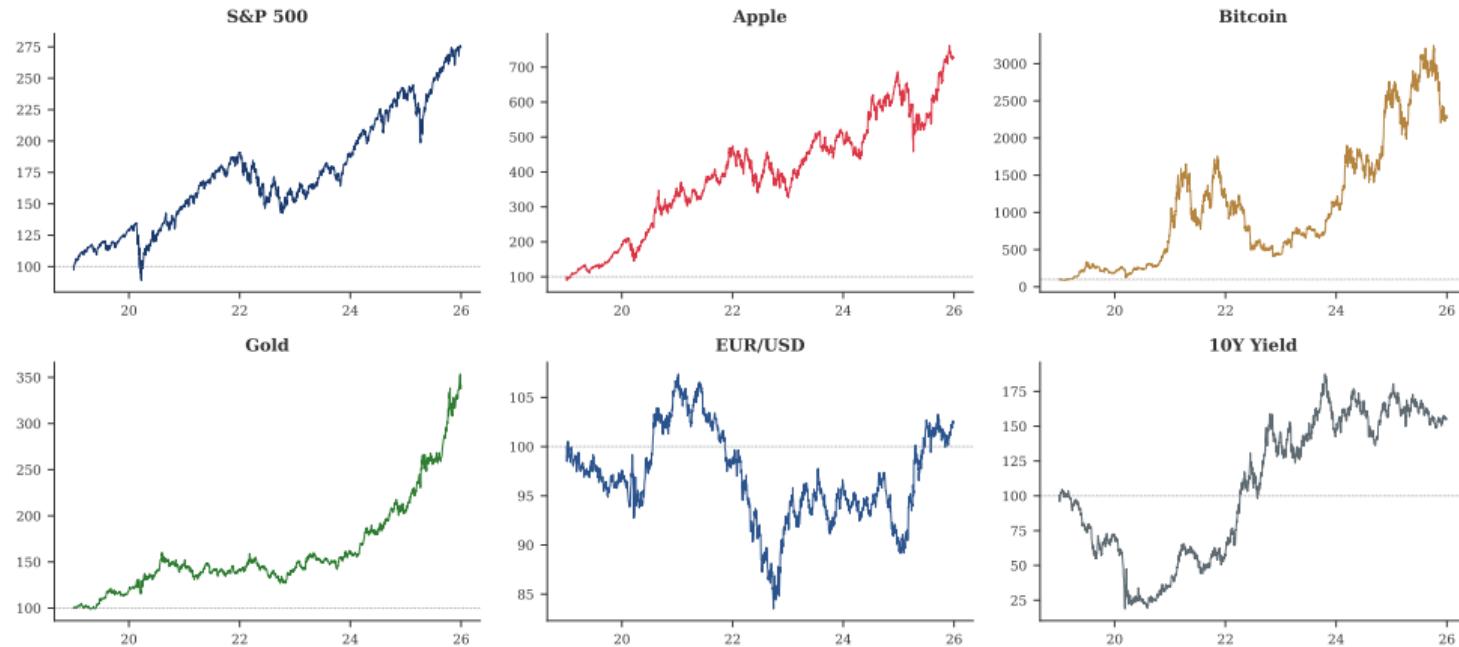
Fiecare punct  $X_t$  reprezintă o măsurătoare la momentul discret  $t$ . Date: S&P 500 (2024).

# Tipuri de Date: Comparăție



Tip de Date	Unități ( $N$ )	Timp ( $T$ )	Exemplu
Transversale	Multe	1	Sondaj pe 1000 gospodării
Serie de timp	1	Multe	Preturi zilnice S&P 500
Panel	Multe	Multe	PIB pentru 50 țări, 20 ani

## Exemplu de Date de Tip Serie de Timp



Date financiare reale de la Yahoo Finance (2019–2025). Normalizeaza la baza 100.

# De Ce Descompunem o Serie de Timp?

**Descompunerea** separă o serie de timp în componente interpretabile:

## Obiective:

- Înțelegerea tiparelor subiacente
- Eliminarea sezonalității pentru modelare
- Identificarea direcției trendului
- Izolarea fluctuațiilor neregulate
- Îmbunătățirea acurateței programei

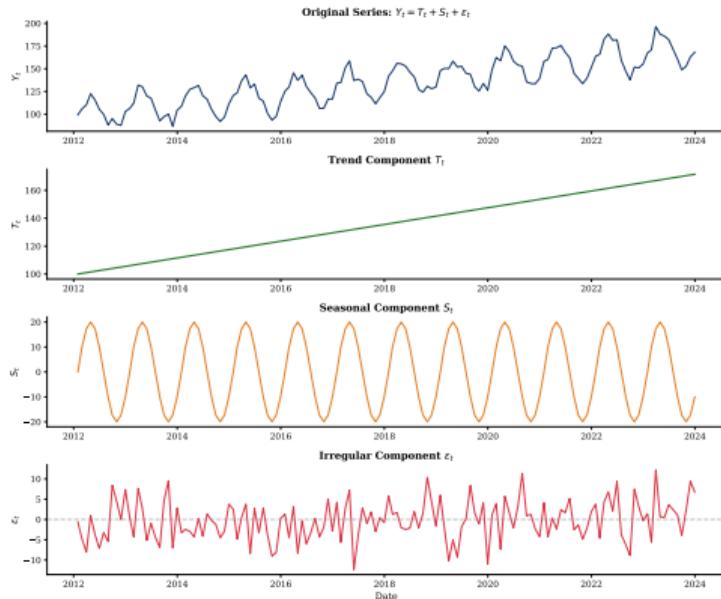
## Componente:

- $T_t$  = **Trend**: Mișcare pe termen lung
- $S_t$  = **Sezonalitate**: Tipar periodic regulat
- $C_t$  = **Ciclic**: Fluctuații ale ciclului de afaceri
- $\varepsilon_t$  = **Rezidual**: Zgomot aleatoriu

## Modele Clasice de Descompunere

- **Aditiv**:  $X_t = T_t + S_t + \varepsilon_t$
- **Multiplicativ**:  $X_t = T_t \times S_t \times C_t \times \varepsilon_t$

# Descompunerea Seriilor de Timp: Exemplu Vizual



- **Original:** Seria de timp observată cu toate componentele
- **Trend:** Mișcarea subiacentă pe termen lung extrasă prin netezire
- **Sezonalitate:** Tiparul periodic regulat care se repetă la fiecare ciclu
- **Rezidual:** Zgomotul aleatoriu după eliminarea trendului și sezonalității

## Componenta Ciclică

**Componenta ciclică**  $C_t$ : Fluctuații pe termen mediu (2–10 ani)

### Caracteristici:

- Fluctuații ale ciclului de afaceri
- Nicio perioadă fixă (spre deosebire de sezonialitate)
- Durata variază: 2–10 ani
- Amplitudinea variază în timp

### Exemple:

- Expansiuni/recesiuni economice
- Cicluri de credit
- Cicluri imobiliare
- Cicluri ale prețurilor materiilor prime

### Notă Practică

Adesea combinată cu trendul ca componentă **trend-ciclu** deoarece:

- Dificil de separat de trend cu date scurte
- Multe metode de descompunere estimatează  $T_t + C_t$  împreună

$$X_t = T_t + S_t + \varepsilon_t \quad (1)$$

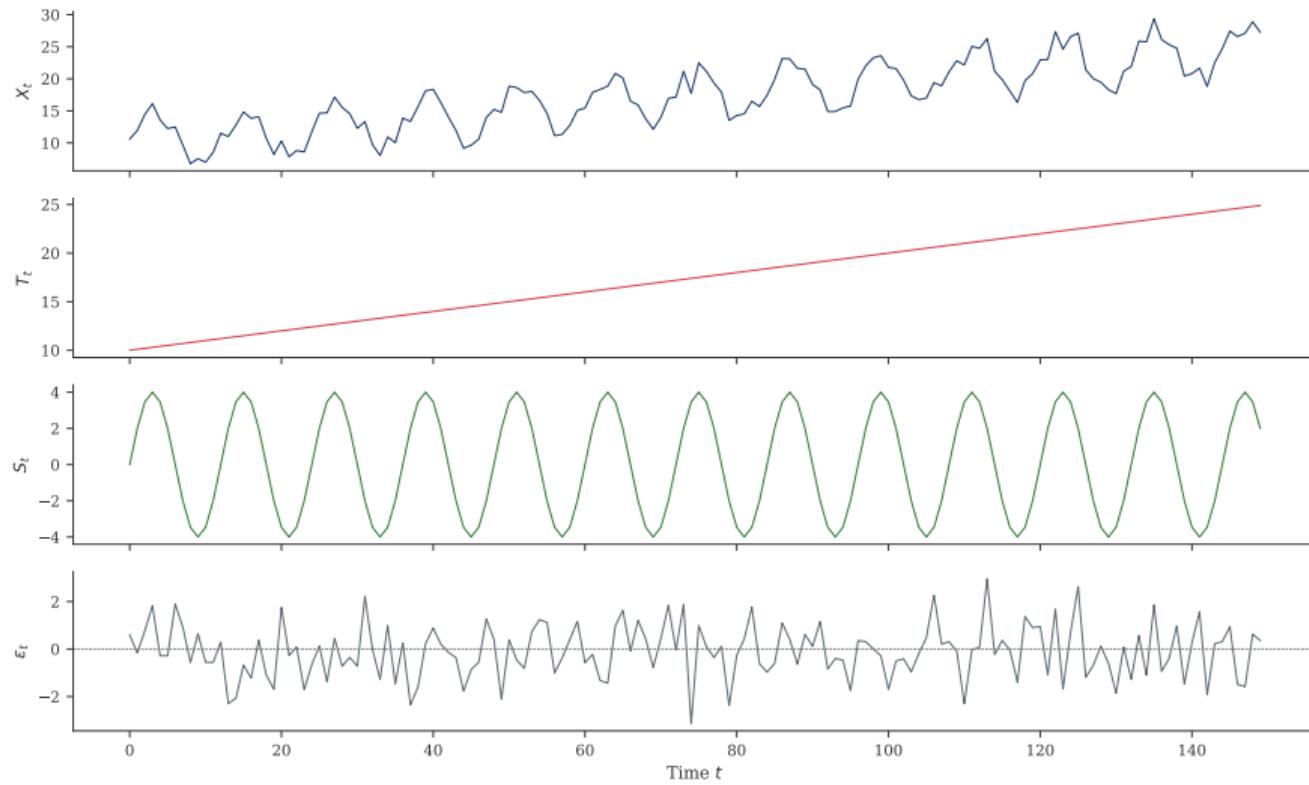
Când se utilizează:

- Fluctuațiile sezoniere sunt **constante** în timp
- Varianța seriei este **stabilă**

Proprietăți:

- $\mathbb{E}[\varepsilon_t] = 0$  (reziduuri cu medie zero)
- $\sum_{j=1}^s S_j = 0$  (sezonalitatea însumează la zero)
- Unitățile lui  $S_t$  sunt aceleași cu ale lui  $X_t$

## Descompunere Aditivă: Vizualizare



## Modelul de Descompunere Multiplicativă

$$X_t = T_t \times S_t \times \varepsilon_t \quad (2)$$

Când se utilizează:

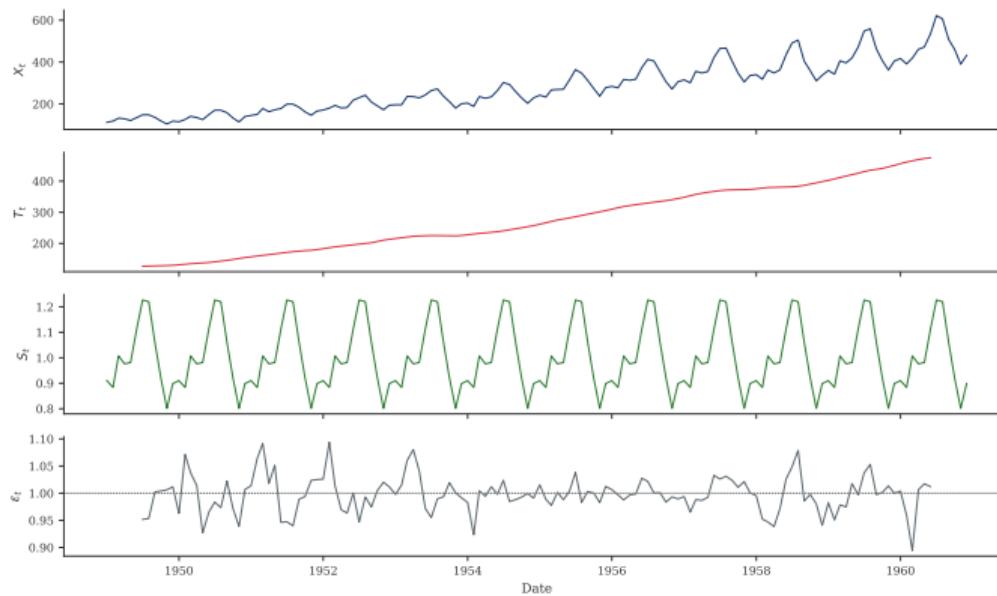
- Fluctuațiile sezoniere cresc odată cu nivelul seriei
- Varianța crește în timp

Proprietăți:

- $\mathbb{E}[\varepsilon_t] = 1$  (reziduuri centrate la 1)
- $\frac{1}{s} \sum S_j = 1$  (media sezonalității este 1)
- $S_t$  este un raport (adimensional)

Sfat: Transformarea logaritmică convertește la modelul aditiv.

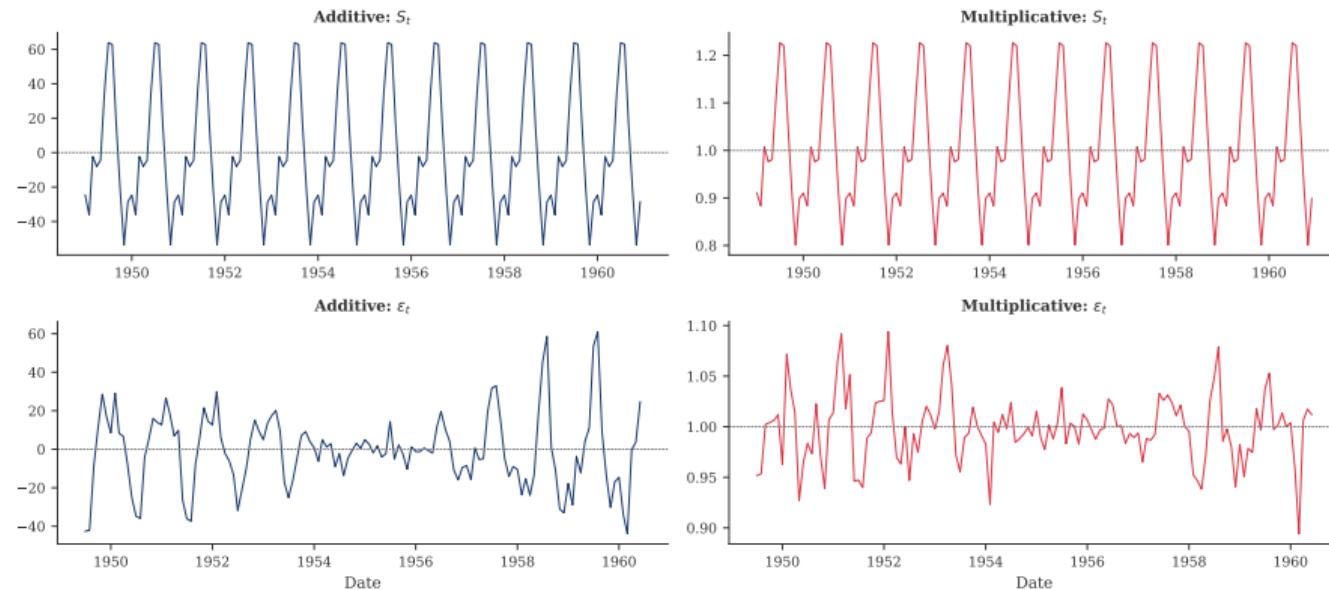
## Descompunere Multiplicativă: Date Reale



### Exemplu

Setul clasic Box-Jenkins pentru pasageri aerieni (1949–1960). Amplitudinea sezonieră crește cu nivelul.

## Aditivă vs Multiplicativă: Comparație



**Diferența cheie:** În modelul multiplicativ, componenta sezonala este un *raport* (centrat la 1), în timp ce în modelul aditiv este în *unități absolute* (centrat la 0).

## Definiție 2 (Media Mobilă Centrată)

Media mobilă centrată de ordin  $2q + 1$  este:

$$\hat{T}_t = \frac{1}{2q+1} \sum_{j=-q}^q X_{t+j}$$

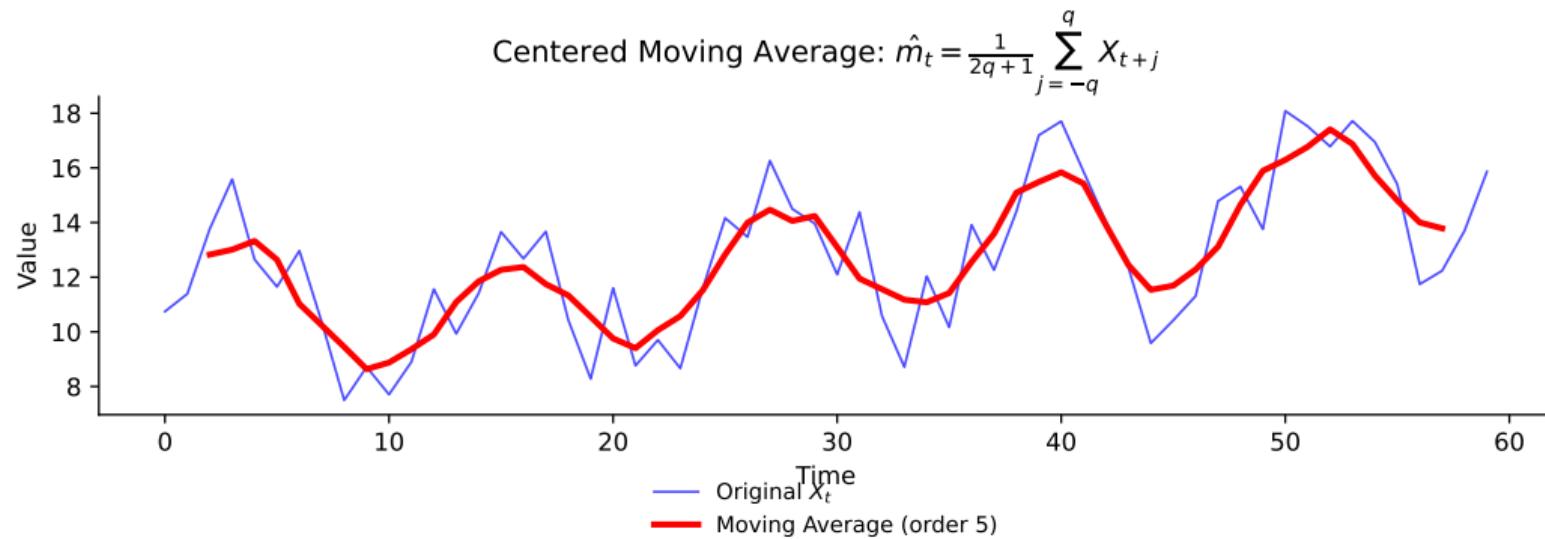
Pentru date sezoniere:

- Dacă perioada  $s$  este **impară**: medie simplă pe  $s$  observații
- Dacă perioada  $s$  este **pară** (de ex., 12): se folosește MA  $2 \times s$  cu ponderi înjumătățite la capete

Proprietăți:

- Netezește fluctuațiile sezoniere și aleatorii
- Fereastră mai mare  $\Rightarrow$  trend mai neted
- Compromis: pierdere de date la capete

## Media Mobilă Centrată: Ilustrație Vizuală



Media mobilă netezește fluctuațiile pe termen scurt, dezvăluind trendul subiacent.

# Algoritmul Descompunerii Clasice

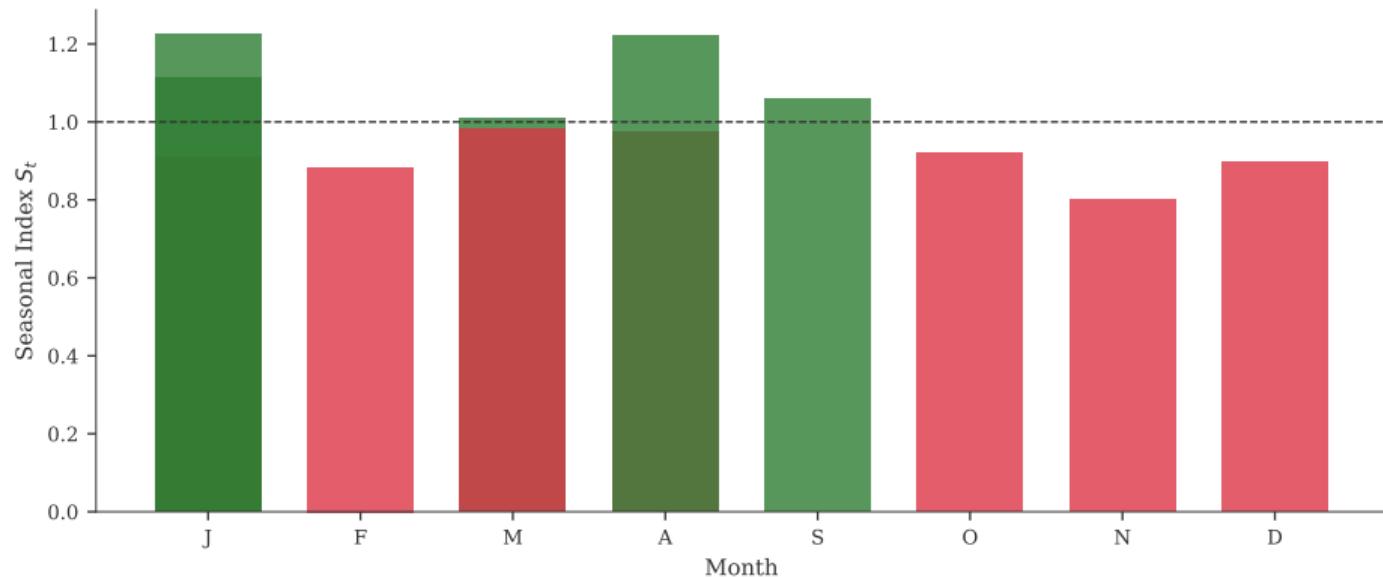
**Pași pentru Descompunerea Multiplicativă:**

- ① **Estimarea Trendului:**  $\hat{T}_t = MA_s(X_t)$
- ② **Eliminarea trendului:**  $D_t = X_t / \hat{T}_t$
- ③ **Estimarea Sezonalității:** Media  $D_t$  pentru fiecare sezon  $j$

$$\hat{S}_j = \text{media}(D_t \text{ pentru toate } t \text{ din sezonul } j)$$

- ④ **Normalizare:** Scalare astfel încât  $\frac{1}{s} \sum_{j=1}^s \hat{S}_j = 1$
- ⑤ **Calculul Reziduurilor:**  $\hat{\varepsilon}_t = X_t / (\hat{T}_t \times \hat{S}_t)$

## Indici Sezonieri: Interpretare



**Interpretare:**  $S_t > 1$  înseamnă activitate peste medie;  $S_t < 1$  înseamnă sub medie. Datele aeriene arată vârf de călătorii în iulie–august.

## Definiție 3 (STL - Descompunere Sezonieră-Trend folosind LOESS)

STL folosește regresia ponderată local (LOESS) pentru a estima componentele:

$$X_t = T_t + S_t + R_t$$

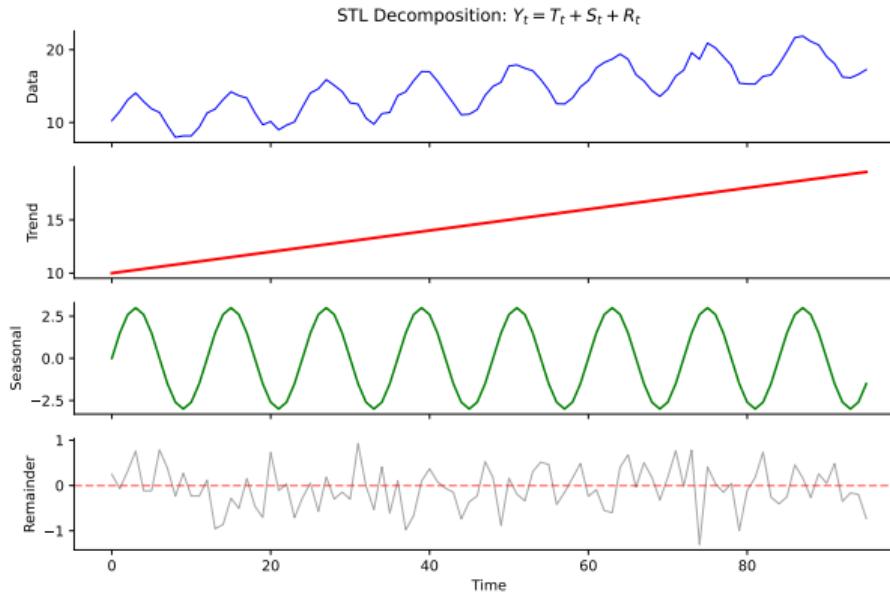
### Avantaje față de descompunerea clasică:

- Gestionează **orice perioadă sezonieră** (nu doar 4 sau 12)
- Componența sezonieră poate să se schimbe **în timp**
- **Robust** la valori aberante (cu opțiunea robust=True)
- Oferă estimări **netede** ale trendului

### Parametri cheie:

- period: Perioada sezonieră (de ex., 12 pentru lunar)
- seasonal: Fereastră pentru netezirea sezonieră (întreg impar)
- robust: Folosește ajustare robustă pentru a reduce ponderea valorilor aberante

## Descompunerea STL: Ilustrație Vizuală



### Observație Cheie

STL separă seria în trend, sezonialitate și rest folosind LOESS.

# Netezirea Exponențială: Prezentare Generală

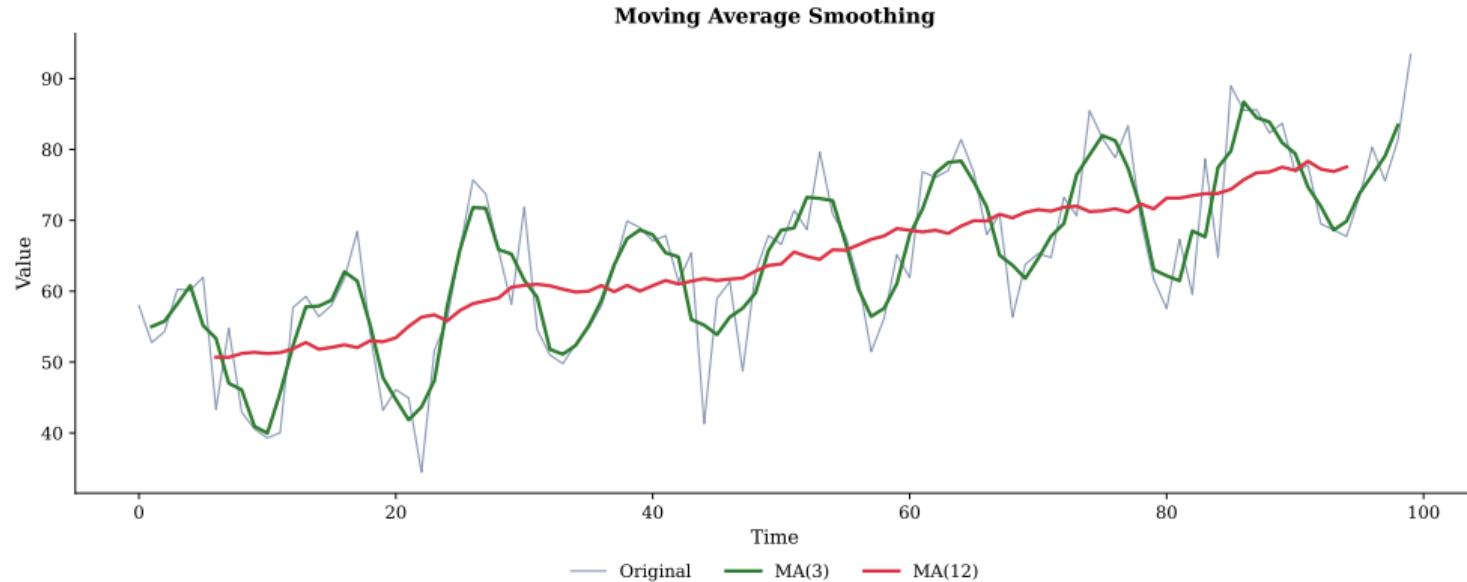
Metodele de **netezire exponențială** produc prognoze bazate pe medii ponderate ale observațiilor trecute, cu ponderi care scad exponențial.

## De Ce Netezire Exponențială?

- Metode de prognoză simple dar eficiente
- Observațiile mai recente primesc ponderi mai mari
- Gestionează trendul și sezonalitatea
- Baza pentru modelele ETS

## Trei metode principale:

- ① **Netezire Exponențială Simplă (SES)**: Doar nivel
- ② **Metoda Holt**: Nivel + Trend
- ③ **Holt-Winters**: Nivel + Trend + Sezonalitate



- **Fereastră mică** (de ex., 5): Receptivă la schimbări dar zgomotoasă
- **Fereastră mare** (de ex., 30): Mai netedă dar mai lentă în reacție
- Compromis între reducerea zgomotului și întârzierea în detectarea schimbărilor

## Netezirea Exponențială Simplă (SES)

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

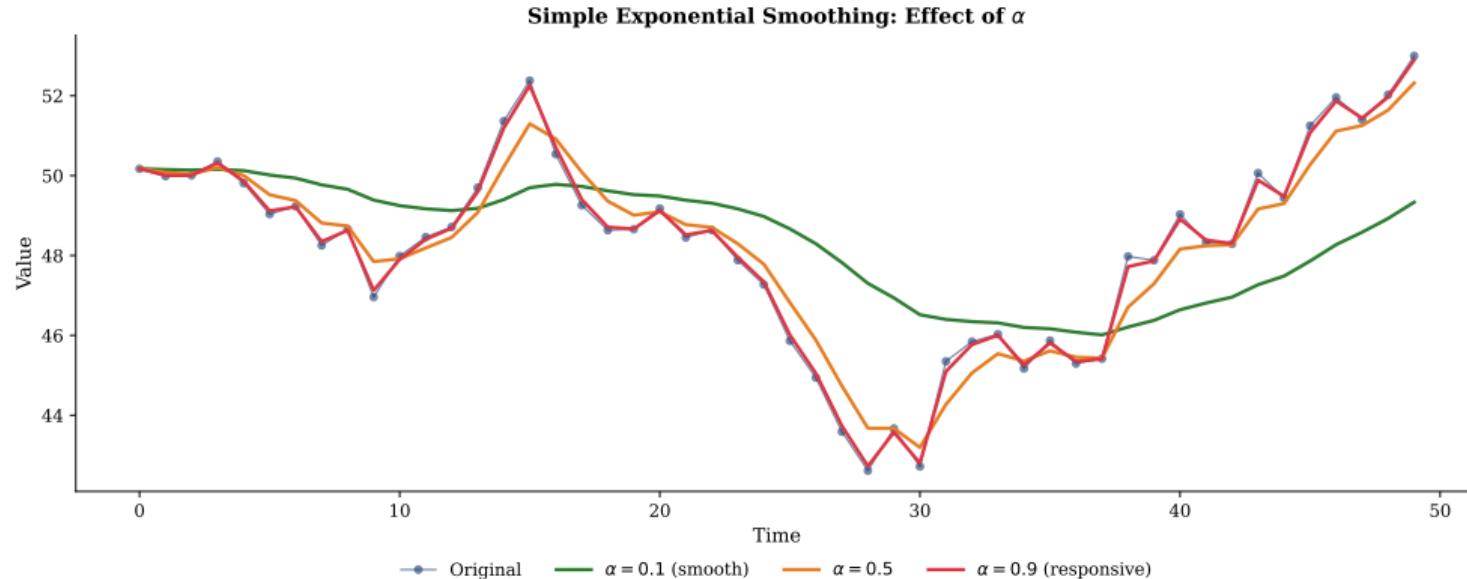
unde  $\alpha \in (0, 1)$  este **parametrul de netezire**.

**Cum funcționează:**

- Ponderile scad exponențial în trecut
- $\alpha$  mare: receptiv la schimbările recente
- $\alpha$  mic: prognoze mai netede, mai stabile

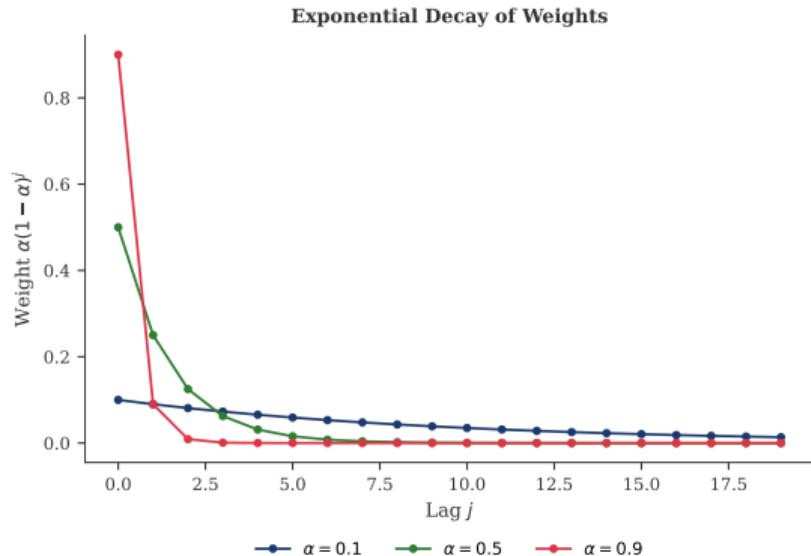
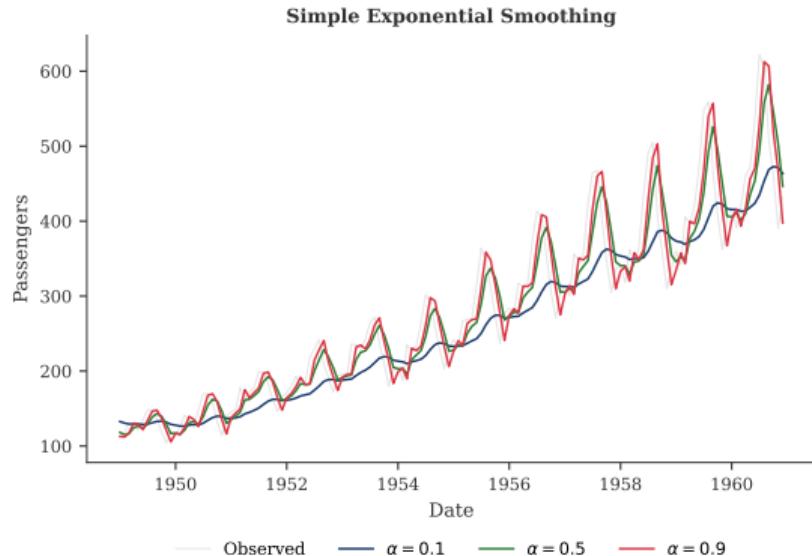
**Forma de nivel:**  $\ell_t = \alpha X_t + (1 - \alpha) \ell_{t-1}$

## Netezirea Exponențială: Efectul lui Alpha



- $\alpha$  mic (de ex., 0.1): Mai multă pondere pe trecut – mai neted, adaptare mai lentă
- $\alpha$  mare (de ex., 0.9): Mai multă pondere pe recent – receptiv, mai volatil
- Alegeti  $\alpha$  în funcție de cât de rapid se schimbă procesul subiacent

## Netezirea Exponențială Simplă: Efectul lui $\alpha$



Un  $\alpha$  mai mic produce programe mai netede; un  $\alpha$  mai mare urmează datele mai îndeaproape.

## Metoda Holt cu Trend Liniar

Extinde SES pentru a captura **trendul liniar** folosind două ecuații:

**Nivel:**  $\ell_t = \alpha X_t + (1 - \alpha)(\ell_{t-1} + b_{t-1})$

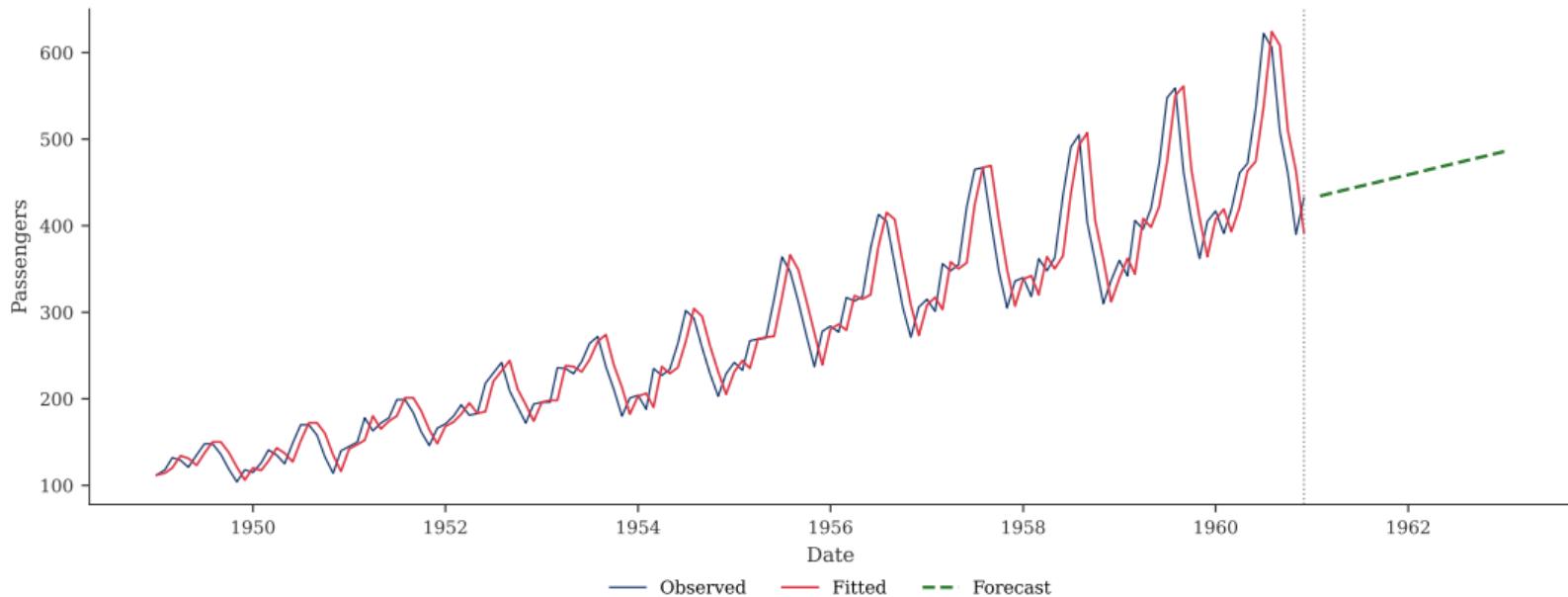
**Trend:**  $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1}$

**Prognoză:**  $\hat{X}_{t+h|t} = \ell_t + h \cdot b_t$

**Parametri:**

- $\alpha \in (0, 1)$ : Parametru de netezire pentru nivel
- $\beta^* \in (0, 1)$ : Parametru de netezire pentru trend
- $\ell_t$ : Nivelul estimat la momentul  $t$
- $b_t$ : Trendul (panta) estimat la momentul  $t$

## Metoda Holt: Vizualizare



Metoda Holt captează atât nivelul cât și trendul, proiectându-le în orizontul de prognoză.

## Metoda Sezonieră Holt-Winters

Extinde metoda Holt pentru a include **sezonalitatea** cu trei ecuații:

**Nivel:**  $\ell_t = \alpha(X_t - S_{t-s}) + (1 - \alpha)(\ell_{t-1} + b_{t-1})$

**Trend:**  $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1}$

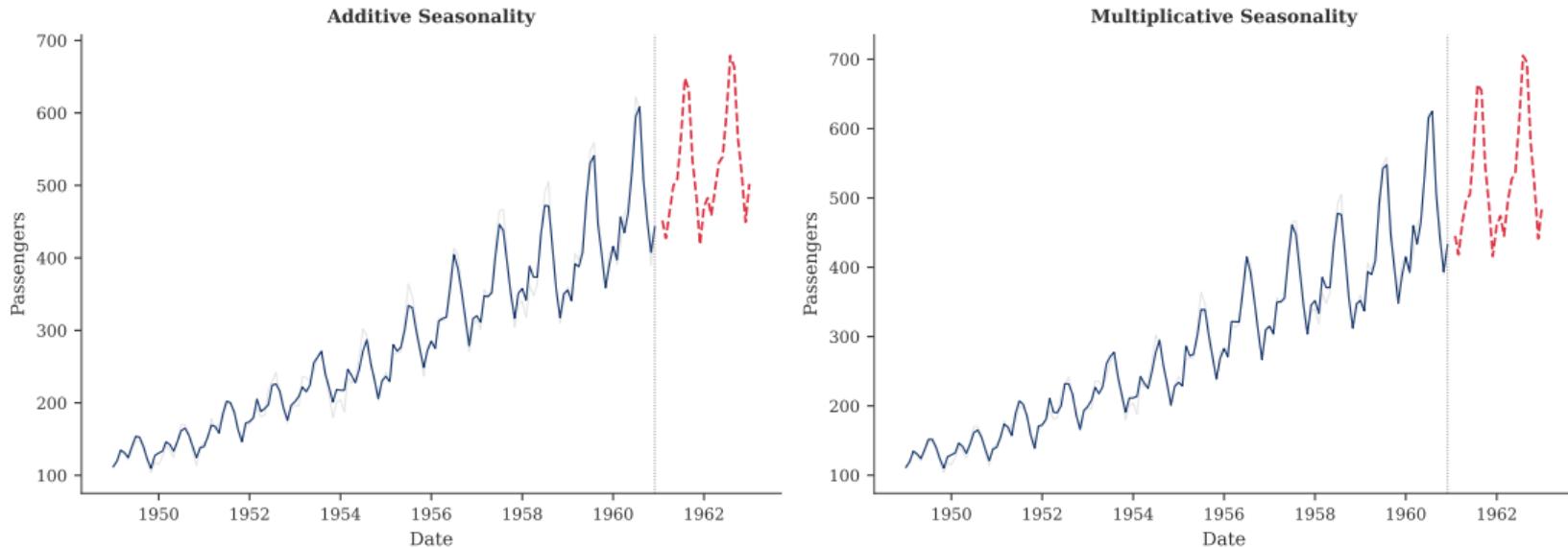
**Sezonalitate:**  $S_t = \gamma(X_t - \ell_t) + (1 - \gamma)S_{t-s}$

**Prognosă:**  $\hat{X}_{t+h|t} = \ell_t + h \cdot b_t + S_{t+h-s(k+1)}$

**Parametri:**

- $\alpha$ : Netezire nivel
- $\beta^*$ : Netezire trend
- $\gamma$ : Netezire sezonalitate
- $s$ : Perioada sezonieră (de ex., 12 pentru lunar)

## Holt-Winters: Captarea Sezonalității



Holt-Winters descompune seria și produce prognoze sezoniere.

## Definiție 4 (Modele ETS)

Cadrul **ETS** generalizează netezirea exponențială cu structură explicită de eroare:

$$\text{ETS}(E, T, S)$$

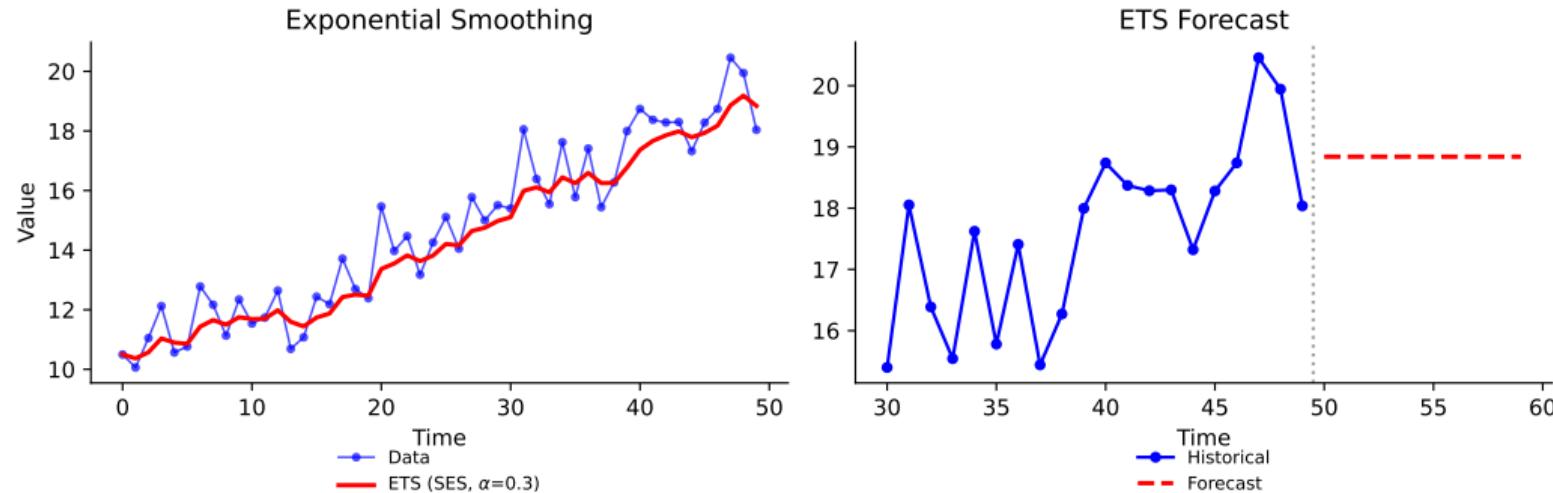
unde fiecare componentă poate fi:

Componentă	N	A	M
Eroare (E)	–	Aditivă	Multiplicativă
Trend (T)	Niciunul	Aditiv	Multiplicativ
Sezonalitate (S)	Niciuna	Aditivă	Multiplicativă

**Exemple:**

- $\text{ETS}(A,N,N) = \text{Netezire Exponențială Simplă}$
- $\text{ETS}(A,A,N) = \text{Metoda Liniară Holt}$
- $\text{ETS}(A,A,A) = \text{Holt-Winters Aditivă}$
- $\text{ETS}(M,A,M) = \text{Erori multiplicitative, trend aditiv, sezonalitate multiplicativă}$

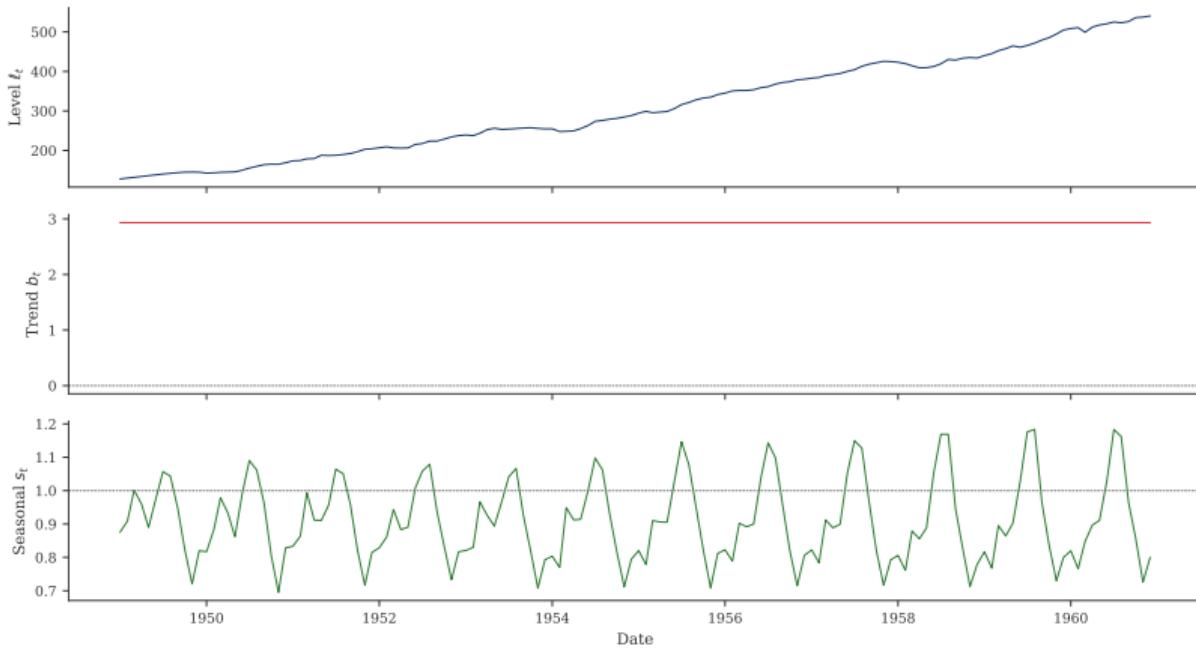
## ETS: Ilustrație Netezire Exponențială



### Interpretare

Modelele ETS folosesc observații ponderate exponențial pentru prognoză. Ponderile scad pe măsură ce observațiile devin mai vechi.

## Selectia Modelului ETS



### Interpretare

Cadrul ETS oferă o modalitate sistematică de a alege cel mai bun model folosind AIC/BIC.

## Parametrul de Amortizare

Introduce  $\phi \in (0, 1)$  pentru a preveni supra-proiecția

### Ecuații

Nivel:  $\ell_t = \alpha X_t + (1 - \alpha)(\ell_{t-1} + \phi b_{t-1})$

Trend:  $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)\phi b_{t-1}$

Prognoză:  $\hat{X}_{t+h|t} = \ell_t + \phi \frac{1 - \phi^h}{1 - \phi} b_t$

### Observație Cheie

- Când  $h \rightarrow \infty$ : prognoza  $\rightarrow$  constantă
- Previne extrapolarea nerealistă
- Adesea cel mai bun pentru orizonturi lungi

# Metrici de Acuratețe a Prognozei

**Eroarea de Prognoză:**  $e_t = X_t - \hat{X}_t$  (real minus prezis)

**Dependente de Scară:**

- MAE =  $\frac{1}{n} \sum |e_t|$
- MSE =  $\frac{1}{n} \sum e_t^2$
- RMSE =  $\sqrt{\text{MSE}}$

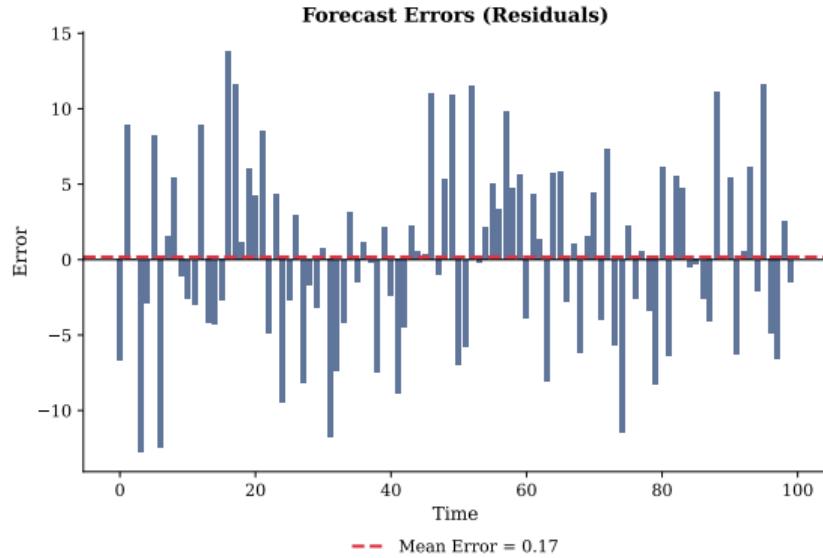
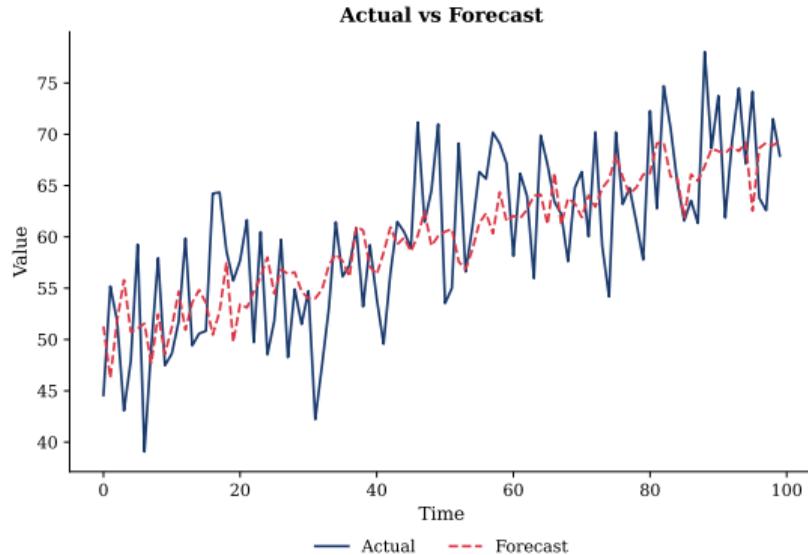
**Independente de Scară:**

- MAPE =  $\frac{100}{n} \sum \left| \frac{e_t}{X_t} \right|$
- sMAPE (simetric)

Pe care să o folosim?

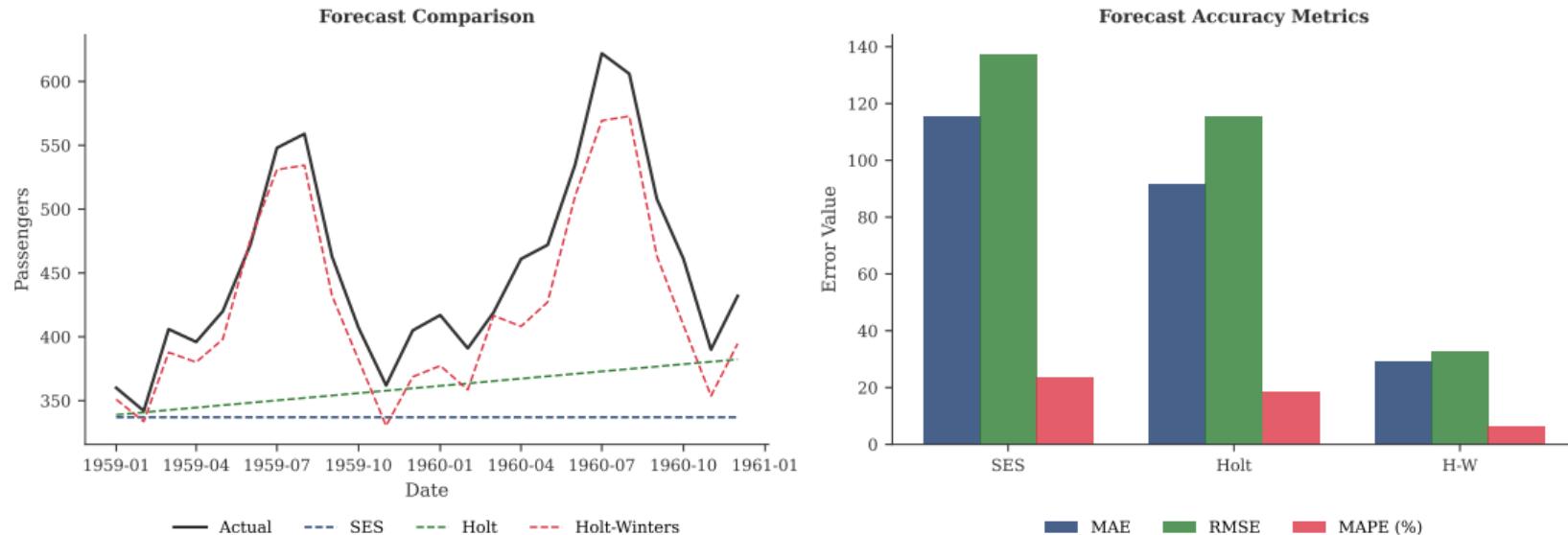
- Aceeași serie: RMSE, MAE
- Comparație între serii: MAPE, sMAPE

## Evaluarea Prognozei: Exemplu Vizual



- **Sus:** Valori reale vs. valori prognozate – evaluare vizuală a potrivirii
- **Jos:** Reziduurile ar trebui să fie centrate în jurul lui zero fără tipar
- Prognozele bune au reziduuri mici, aleatorii cu varianță constantă

# Compararea Metodelor de Prognoză



## Interpretare

**Stânga:** Compararea prognozelor SES, Holt și Holt-Winters. **Dreapta:** Metrii de eroare pentru fiecare metodă.

## Proprietăți Reziduuri

Prognozele bune ar trebui să aibă reziduuri:

- ① **Medie zero:**  $\mathbb{E}[e_t] = 0$
- ② **Necorelate:**  $\text{Cov}(e_t, e_{t-k}) = 0$
- ③ **Varianță constantă:**  $\text{Var}(e_t) = \sigma^2$
- ④ **Distribuite normal**

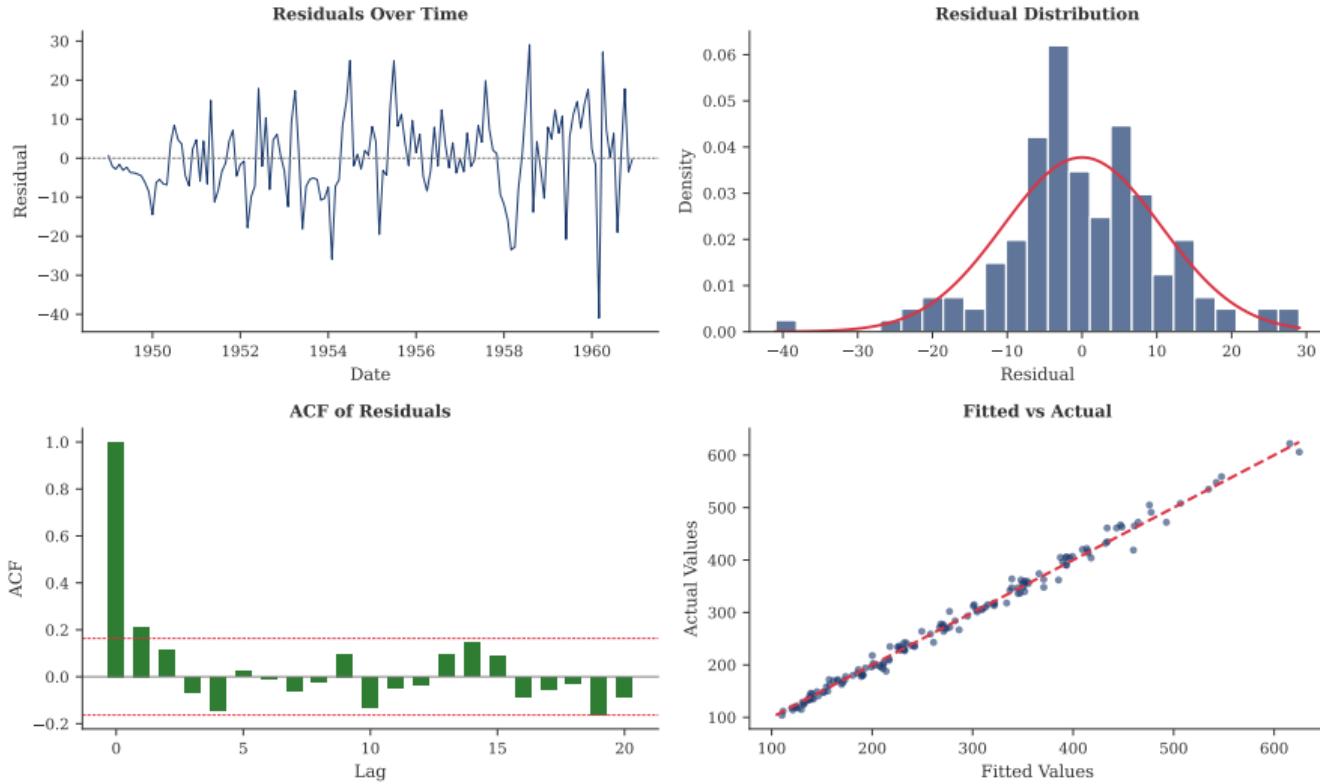
## Teste de Diagnostic

**Testul Ljung-Box** (autocorelație):

$$Q = T(T+2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{T-k} \sim \chi_h^2$$

**Testul Jarque-Bera** (normalitate)

# Diagnosticarea Reziduurilor: Vizualizare



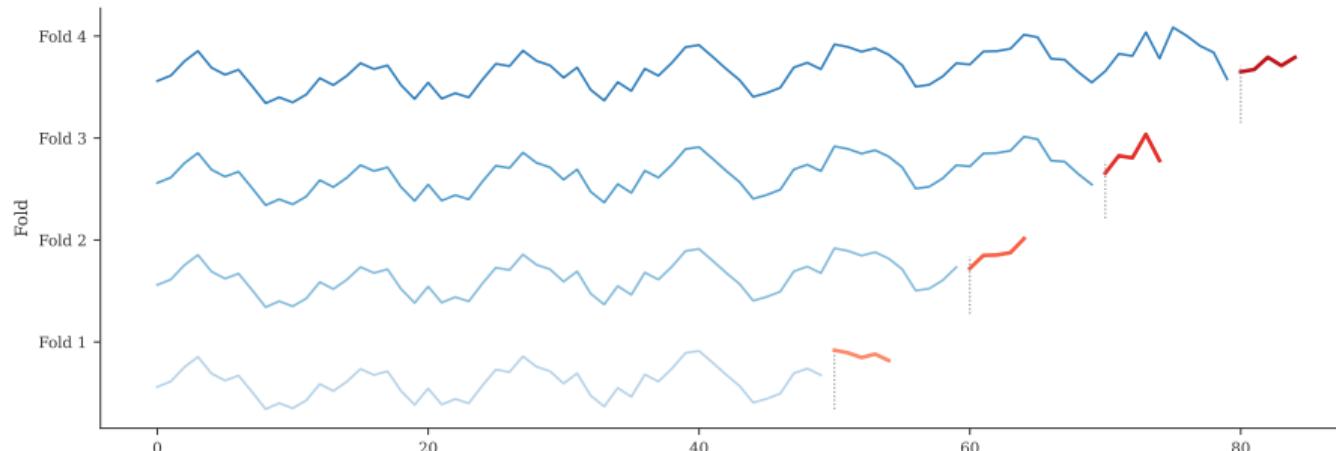
# Validarea Încrucișată pentru Serii de Timp

## Important

CV standard nu funcționează pentru serii de timp (dependentă temporală).

## CV cu Origine Mobilă (Ferestre în Expansiune)

- ① Antrenare pe  $\{X_1, \dots, X_t\}$ , prognoză  $\hat{X}_{t+h}$
- ② Incrementare  $t$ , repetare



# Separarea Train / Validare / Test

Separare în trei părți pentru dezvoltarea modelului:

## Set de Antrenare

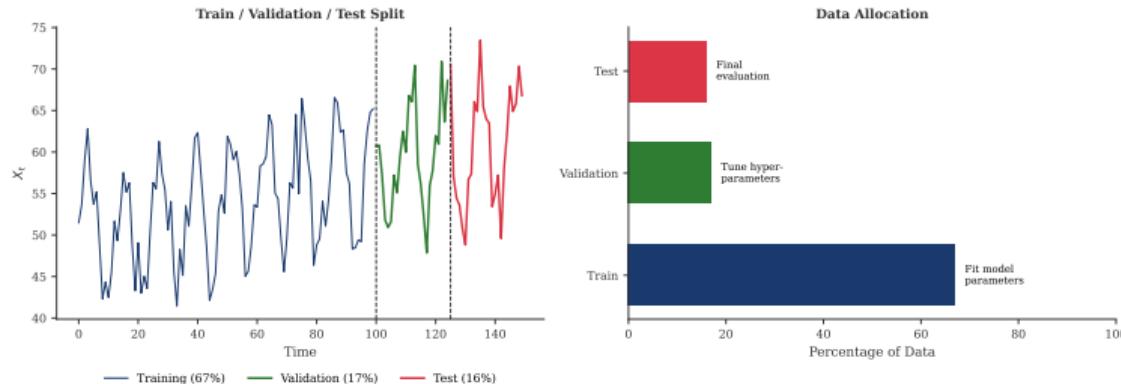
- Potrivirea parametrilor modelului
- Cea mai mare porțiune (60–80%)
- Folosit pentru estimare

## Set de Validare

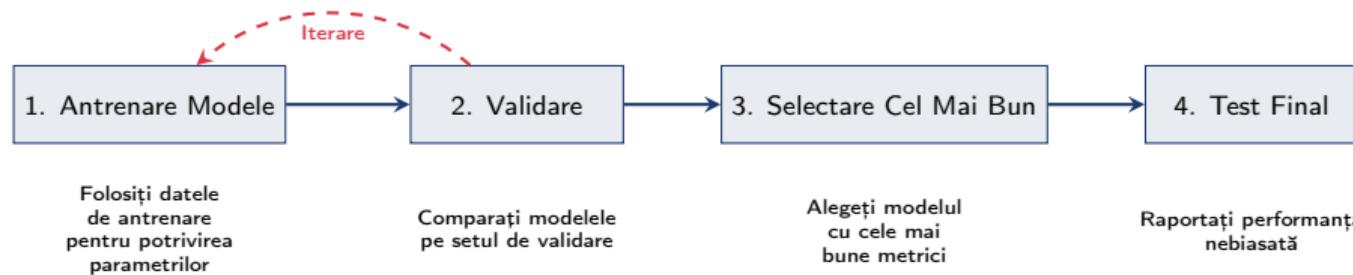
- Ajustarea hiperparametrilor
- Compararea modelelor
- Selectarea celei mai bune abordări

## Set de Test

- Doar evaluare finală
- Niciodată folosit pentru ajustare
- Performanță nebiasată



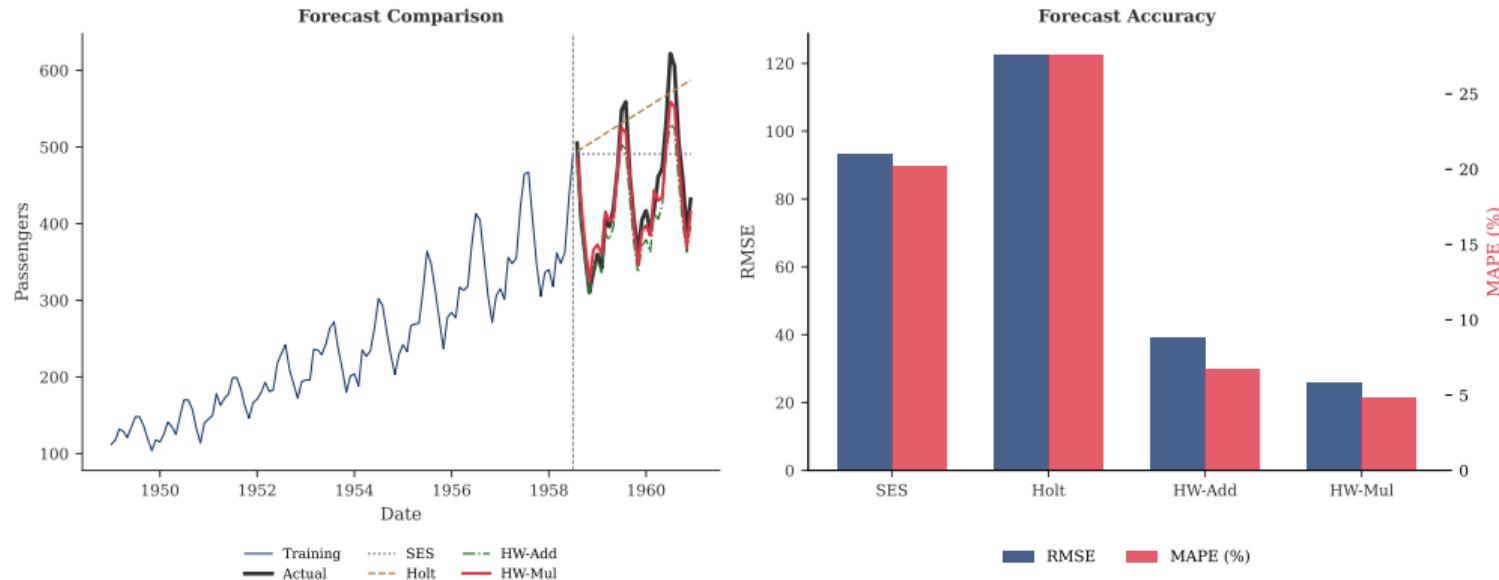
# Fluxul de Lucru pentru Dezvoltarea Modelului



## Regulă Critică

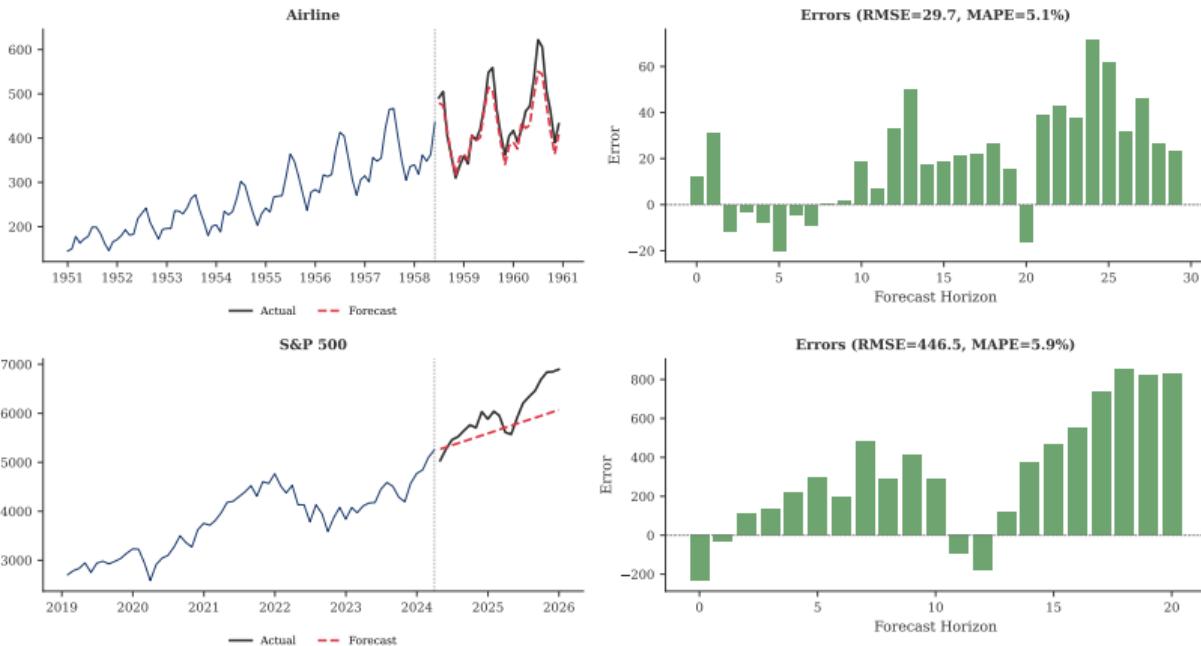
Niciodată nu folosiți setul de test pentru selecția modelului! Aceasta cauzează *surgere de date* și estimări excesiv de optimiste ale performanței.

# Date Reale: Compararea Prognozelor



Date pasageri aerieni: Holt-Winters Multiplicativ performează cel mai bine pentru date sezoniere.

# Performanța Prognozei pe Diferite Seturi de Date



Serii diferite necesită modele diferite. Datele sezoniere au nevoie de metode sezoniere.

## 1. Variabile Dummy:

$$X_t = \mu + \sum_{j=1}^{s-1} \gamma_j D_{jt} + \varepsilon_t$$

- $D_{jt} = 1$  dacă  $t$  în sezonul  $j$
- $s - 1$  parametri
- Orice tipar sezonier

## 2. Termeni Fourier:

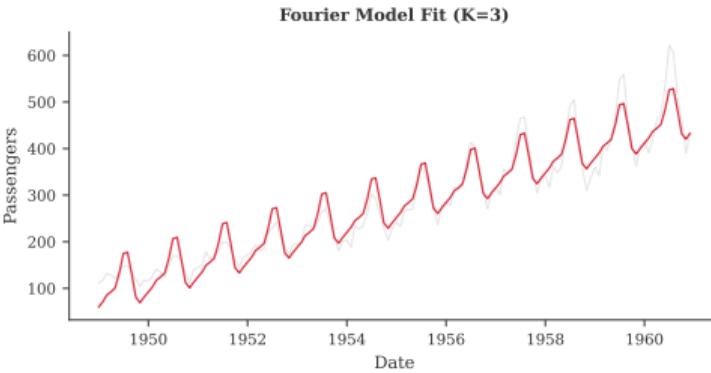
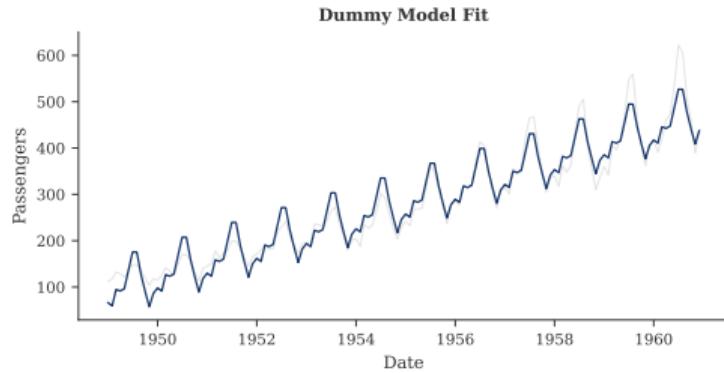
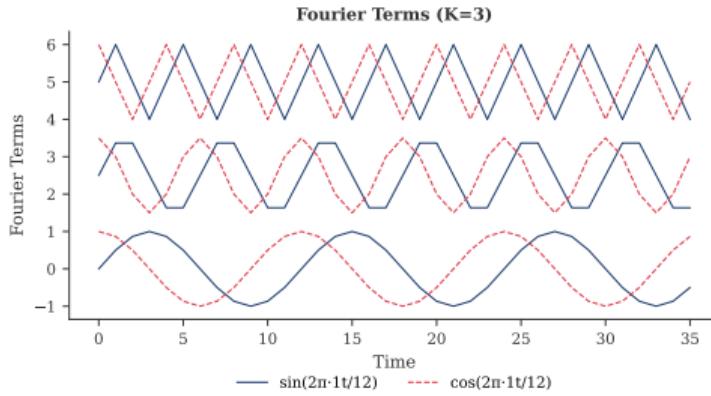
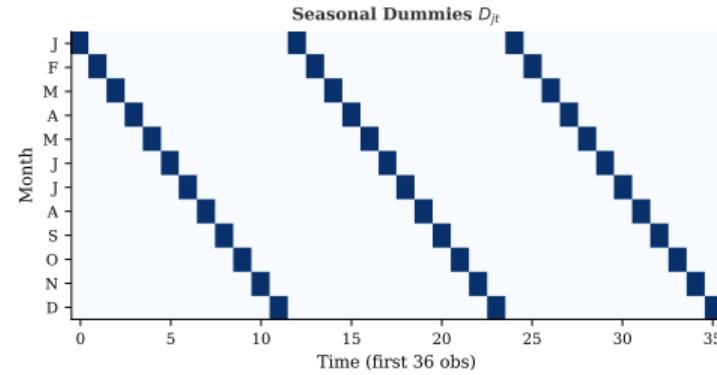
$$X_t = \mu + \sum_{k=1}^K [\alpha_k \sin(\cdot) + \beta_k \cos(\cdot)]$$

- Funcții sinusoidale
- $2K$  parametri
- Tipare netede

## Compromis

Dummy: orice tipar, mai mulți parametri. Fourier: neted, mai puțini parametri.

# Variabile Dummy vs Termeni Fourier



## Alegerea între Dummy și Fourier

Criteriu	Dummy	Fourier
Parametri (lunar)	11	$2K$ (adesea 4–6)
Tipar sezonier	Orice formă	Neted/sinusoidal
Interpretare	Directă (efekte lunare)	Componente de frecvență
Sezoane de înaltă frecvență	Mulți parametri	Eficient
Sezonalitate multiplă	Complex	Ușor (adăugare termeni)

### Ghiduri:

- Folosiți **dummy** când tiparul sezonier este neregulat sau aveți nevoie de coeficienți interpretabili
- Folosiți **Fourier** pentru tipare netede, sezonalitate de înaltă frecvență (zilnică, orară) sau perioade sezoniere multiple
- **Termenii Fourier** sunt folosiți în modelele TBATS și Facebook Prophet

# De Ce Eliminăm Trendul și Sezonialitatea?

Înainte de modelare, adesea trebuie să facem seria staționară:

## Motive pentru eliminarea trendului:

- Cerința de staționaritate
- Focus pe fluctuații
- Evitarea regresiei false
- Permiterea inferenței valide

## Motive pentru desezonalizare:

- Dezvăluirea trendului subiacent
- Comparare între sezoane
- Simplificarea modelării
- Focus pe componenta neregulată

## Important

După modelarea seriei fără trend/desezonalizate, trebuie să **inversăm transformarea** pentru prognoză.

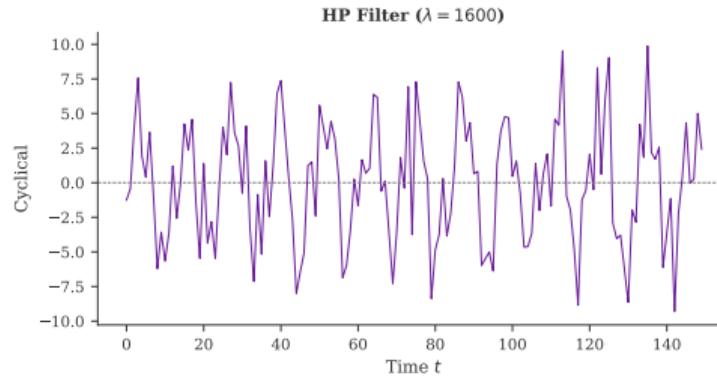
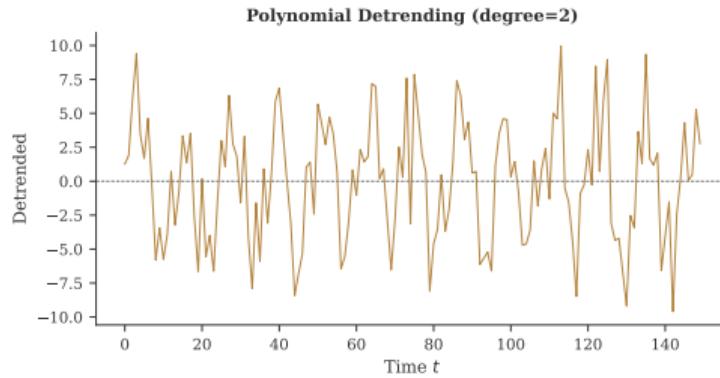
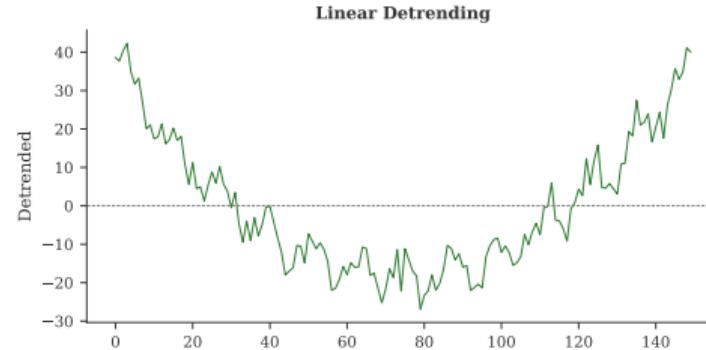
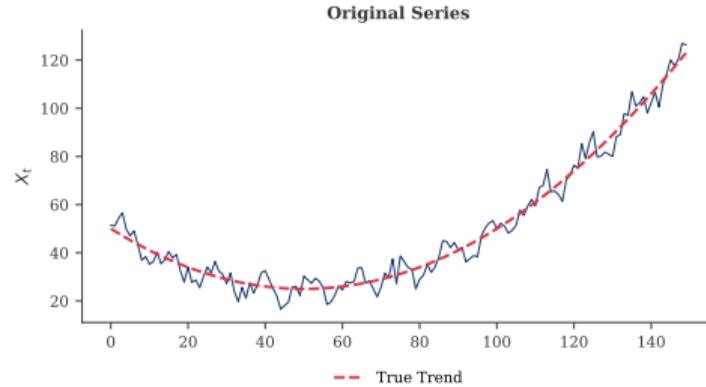
## Şase abordări comune pentru eliminarea trendului:

- ① Diferențiere:  $\Delta X_t = X_t - X_{t-1}$
- ② Regresie liniară: Potrivire  $\hat{T}_t = \hat{\beta}_0 + \hat{\beta}_1 t$
- ③ Polinomială: Potrivire polinom de ordin superior
- ④ Filtru HP: Echilibru între potrivire și netezime
- ⑤ Media mobilă:  $\hat{T}_t = MA_q(X_t)$
- ⑥ LOESS: Regresie polinomială locală

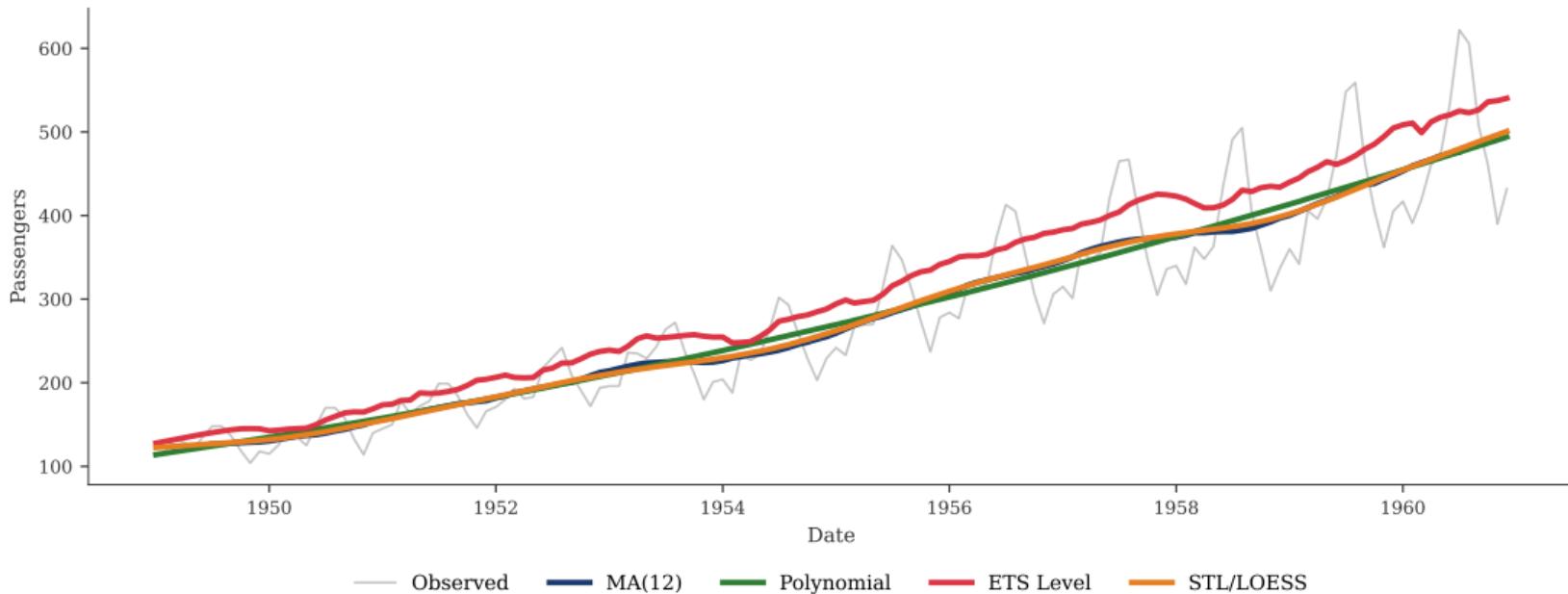
Alegerea depinde de:

- Natura trendului (determinist vs stochastic)
- Scopul (prognoză vs analiză)

# Metode de Eliminare a Trendului: Comparăție



## Estimarea Trendului: Abordări Multiple



Metodele diferite captează trendul la niveluri variate de netezime.

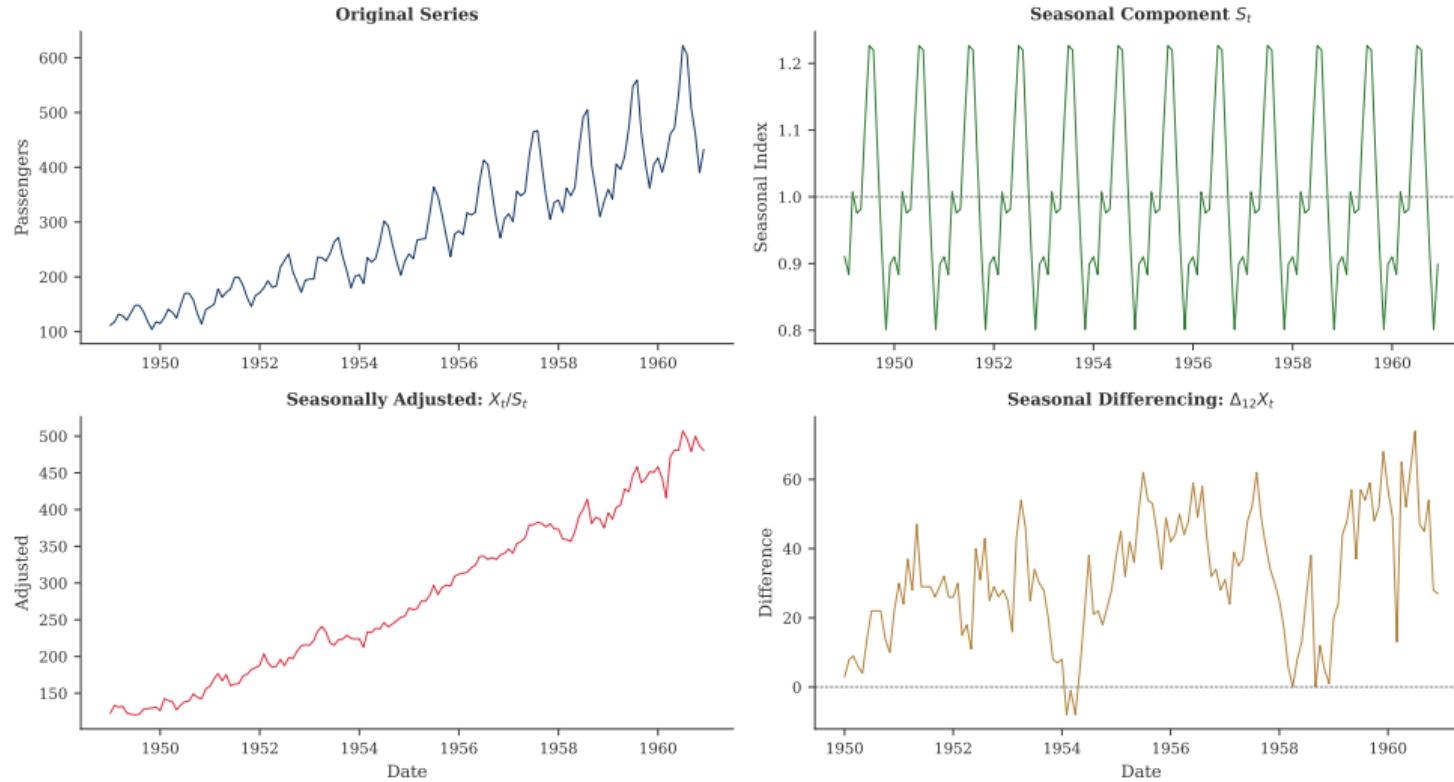
# Metode de Eliminare a Sezonalității

Patru abordări pentru eliminarea sezonalității:

- ① Diferențiere sezonieră:  $\Delta_s X_t = X_t - X_{t-s}$
- ② Împărțire (multiplicativ):  $X_t^{adj} = X_t / \hat{S}_t$
- ③ Scădere (aditiv):  $X_t^{adj} = X_t - \hat{S}_t$
- ④ X-13ARIMA-SEATS: Metodă statistică guvernamentală

Perioada sezonieră  $s$ : Lunar  $\Rightarrow s = 12$ ; Trimestrial  $\Rightarrow s = 4$

# Ajustare Sezonieră: Vizualizare



## Trend Determinist vs Stochastic

### Trend Determinist:

$$X_t = \beta_0 + \beta_1 t + \varepsilon_t$$

- Trendul este o funcție de timp
- Eliminarea trendului prin regresie
- $\varepsilon_t$  este staționar

### Trend Stochastic:

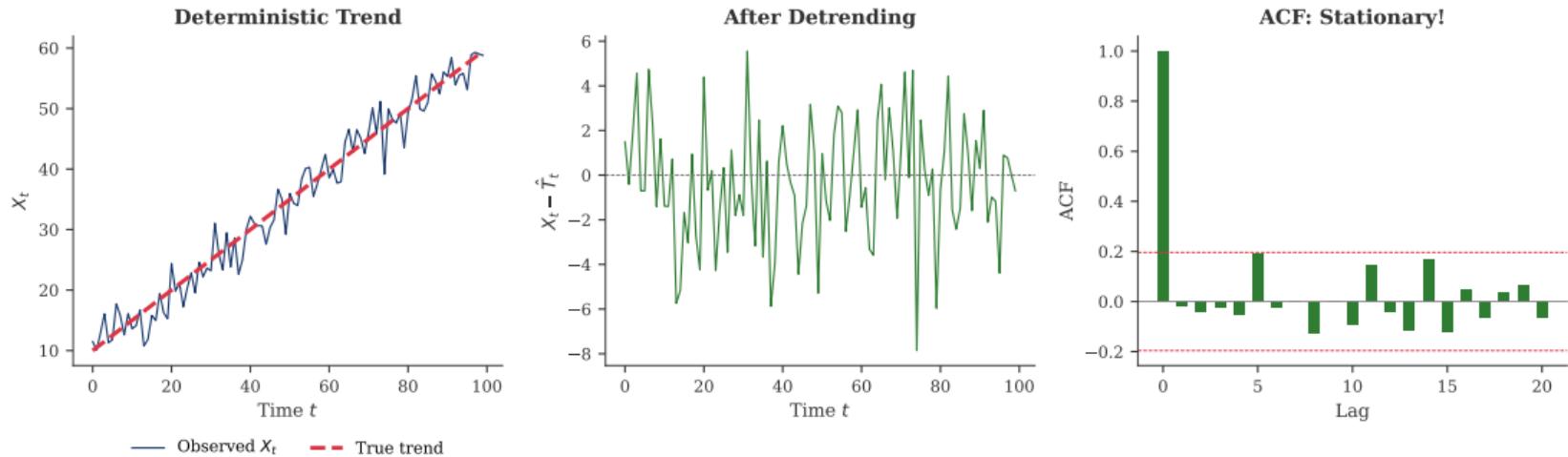
$$X_t = X_{t-1} + \varepsilon_t$$

- Componentă de mers aleatoriu
- Eliminarea trendului prin diferențiere
- $\Delta X_t$  este staționar

### Metoda Greșită = Probleme

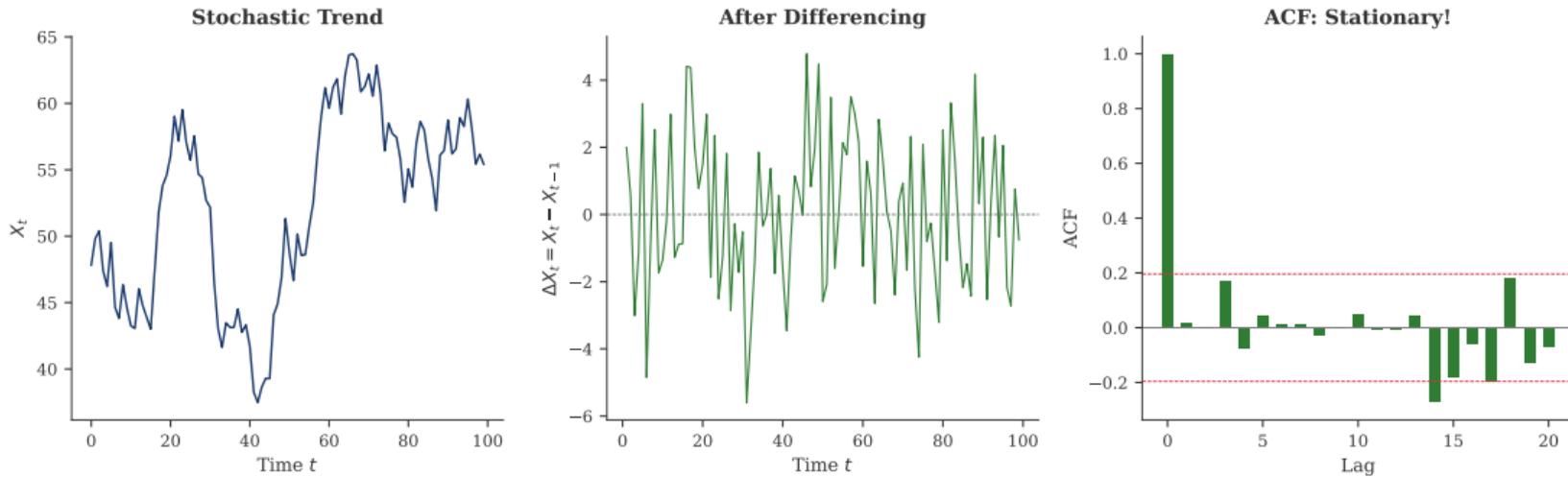
- Diferențierea trendului determinist  $\Rightarrow$  supra-diferențiere
- Regresie pe trend stochastic  $\Rightarrow$  regresie falsă

## Exemplu: Trend Determinist



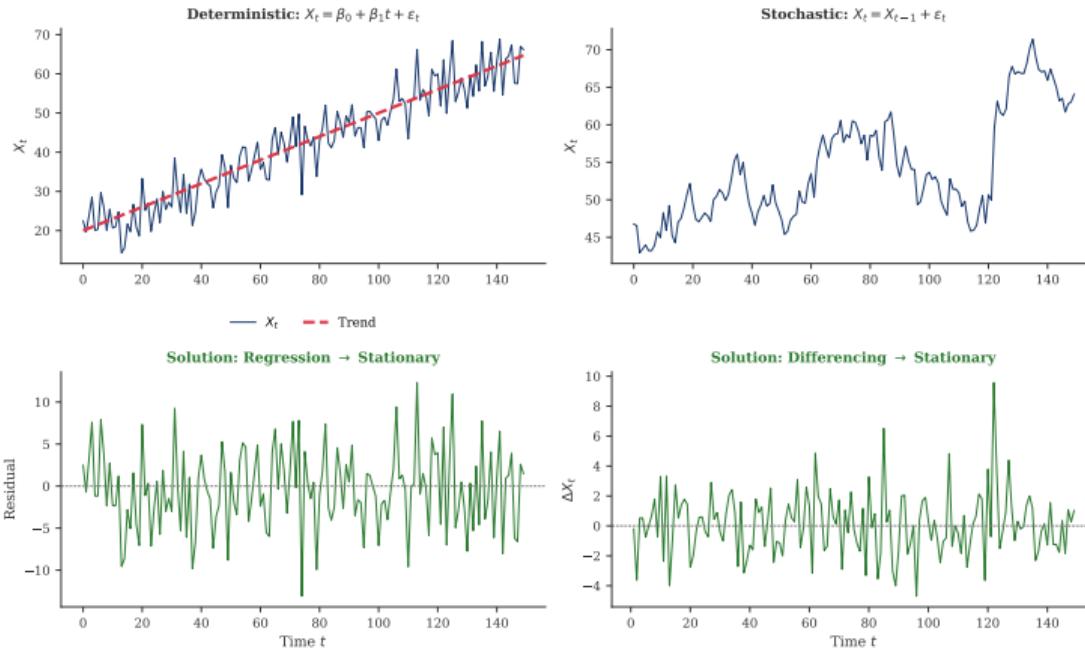
Cheie: Folosiți **regresia** pentru a elimina trendul → reziduurile sunt staționare (ACF scade rapid).

## Exemplu: Trend Stochastic (Mers Aleatoriu)



**Cheie:** Folosiți **diferențierea** pentru a elimina trendul → diferențele sunt staționare (zgomot alb).

# Comparație Alăturată



**Rețineți:** Determinist  $\rightarrow$  regresie. Stochastic  $\rightarrow$  diferențiere.

## Definiție 5 (Proces Stochastic)

Un proces stochastic este o colecție de variabile aleatoare indexate după timp:

$$\{X_t(\omega) : t \in \mathcal{T}, \omega \in \Omega\}$$

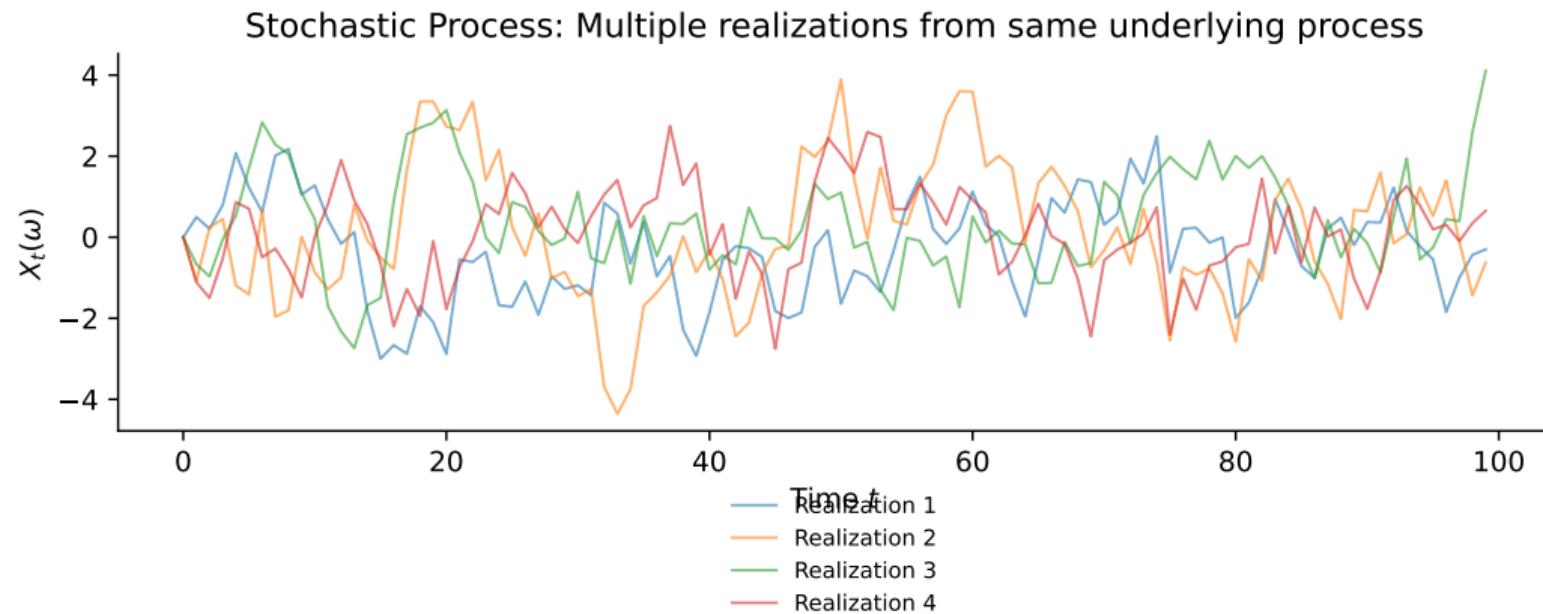
unde  $\Omega$  este spațiul de selecție al rezultatelor posibile.

### Două perspective:

- $\omega$  fix: O realizare sau *traекторie de selecție*  $\{X_t(\omega)\}_{t \in \mathcal{T}}$
- $t$  fix: O variabilă aleatoare  $X_t$  cu distribuția  $F_t(x)$

**Observație cheie:** O serie de timp pe care o observăm este **o realizare** a procesului stochastic subiacent. Folosim această singură realizare pentru a deduce proprietățile procesului.

## Proces Stochastic: Ilustrație Vizuală



Fiecare linie este o realizare diferită din același proces stochastic subiacent. Observăm doar o realizare dar vrem să înțelegem procesul.

# Momentele unui Proces Stochastic

Primele două momente caracterizează proprietățile slabe:

**Funcția de Medie:**  $\mu_t = \mathbb{E}[X_t]$

**Funcția de Autocovarianță (ACVF):**

$$\gamma(t, s) = \text{Cov}(X_t, X_s) = \mathbb{E}[(X_t - \mu_t)(X_s - \mu_s)]$$

**Funcția de Autocorelație (ACF):**

$$\rho(t, s) = \frac{\gamma(t, s)}{\sqrt{\text{Var}(X_t) \cdot \text{Var}(X_s)}}$$

**Proprietăți:**  $\rho(t, s) \in [-1, 1]$  și  $\rho(t, t) = 1$

## De Ce Contează Staționaritatea

**Staționaritatea** este o ipoteză fundamentală pentru analiza seriilor de timp:

### Fără Staționaritate:

- Media, varianța se schimbă în timp
- Trecutul poate să nu prezică viitorul
- Metodele standard eșuează
- Corelații false

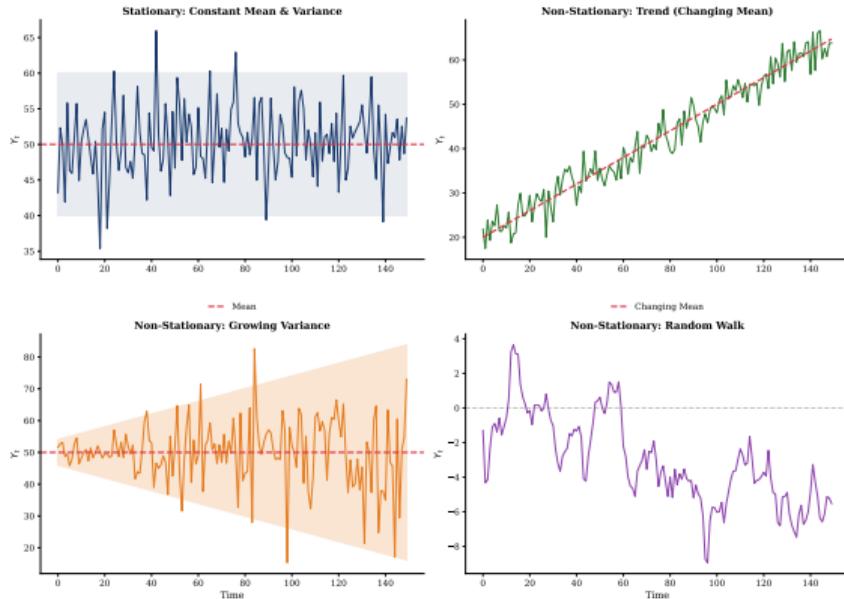
### Cu Staționaritate:

- Proprietăți statistice constante
- Putem estima din o singură realizare
- Inferență validă posibilă
- Modelele sunt semnificative

### Principiu Cheie

Majoritatea modelelor de serii de timp (ARMA, ARIMA, etc.) necesită staționaritate. Seriile nestaționare trebuie transformate (de ex., diferențiere) înainte de modelare.

# Staționar vs Nestăționar: Comparație Vizuală



- **Staționar:** Medie și varianță constantă – fluctuează în jurul unui nivel fix
- **Nestăționar:** Media și/sau varianța se schimbă în timp
- Inspecția vizuală este primul pas; testele formale (ADF, KPSS) confirmă

### Definiție 6 (Staționaritate Strictă (Puternică))

Un proces  $\{X_t\}$  este **strict staționar** dacă pentru toți  $k$ , toți  $t_1, \dots, t_k$  și toți  $h$ :

$$(X_{t_1}, X_{t_2}, \dots, X_{t_k}) \stackrel{d}{=} (X_{t_1+h}, X_{t_2+h}, \dots, X_{t_k+h})$$

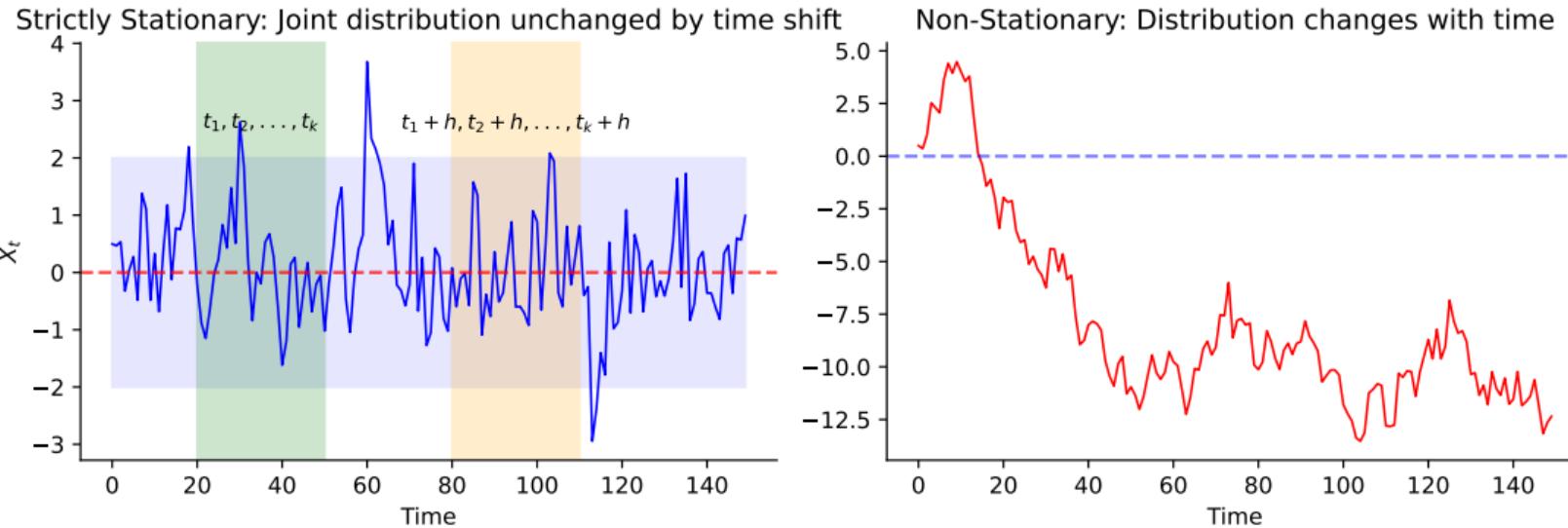
**Interpretare:** Distribuția comună a oricărei colecții de observații este **invariantă la deplasări temporale**.

**Implicații:**

- Toate distribuțiile marginale  $F_{X_t}(x)$  sunt identice
- $\mathbb{E}[X_t] = \mu$  (medie constantă)
- $\text{Var}(X_t) = \sigma^2$  (varianță constantă)
- Distribuțiile comune depind doar de *diferențele temporale*

**Notă:** Staționaritatea strictă este o condiție puternică, adesea impractică de verificat.

## Staționaritate Strictă: Ilustrație Vizuală



Staționar: oricare două ferestre au aceeași distribuție comună. Nestaționar: distribuția se schimbă în timp.

## Staționaritate Slabă (de Covarianță)

### Definiție 7 (Staționaritate Slabă)

Un proces  $\{X_t\}$  este **slab staționar** (sau staționar de covarianță) dacă:

- ①  $\mathbb{E}[X_t] = \mu$  (medie constantă)
- ②  $\text{Var}(X_t) = \sigma^2 < \infty$  (varianță constantă, finită)
- ③  $\text{Cov}(X_t, X_{t+h}) = \gamma(h)$  (covarianță depinde doar de lag-ul  $h$ )

**Proprietate cheie:** Autocovarianța este o funcție doar de lag:

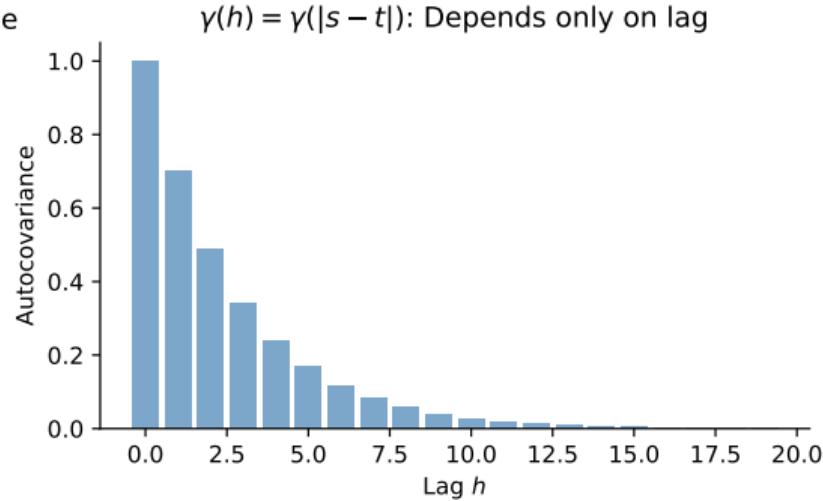
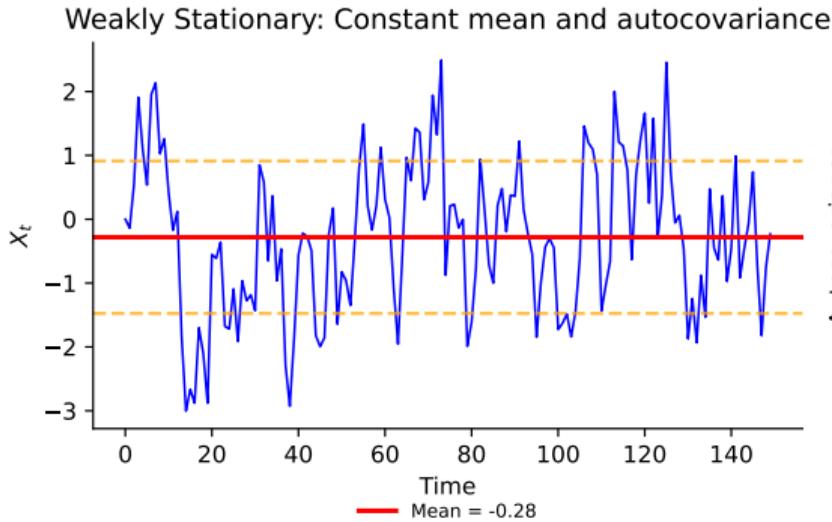
$$\gamma(h) = \text{Cov}(X_t, X_{t+h}) = \mathbb{E}[(X_t - \mu)(X_{t+h} - \mu)]$$

**Funcția de autocorelație:**

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{\text{Cov}(X_t, X_{t+h})}{\text{Var}(X_t)}$$

Notă:  $\rho(0) = 1$  și  $\rho(h) = \rho(-h)$  (simetrie)

## Staționaritate Slabă: Ilustrație Vizuală



Stânga: medie și varianță constantă. Dreapta: autocovarianță depinde doar de lag-ul  $h$ , nu de timpul  $t$ .

## Proprietățile Funcției de Autocovarianță

Pentru un proces slab staționar, ACVF  $\gamma(h)$  satisfac:

- ① **Simetrie:**  $\gamma(h) = \gamma(-h)$
- ② **Maxim la zero:**  $|\gamma(h)| \leq \gamma(0)$
- ③ **Definit nenegativ**

**Implicație:** Nu orice funcție poate fi o funcție de autocovarianță.

## Definiție 8 (Zgomot Alb)

Un proces  $\{\varepsilon_t\}$  este **zgomot alb**, notat  $\varepsilon_t \sim WN(0, \sigma^2)$ , dacă:

- ❶  $\mathbb{E}[\varepsilon_t] = 0$  pentru toți  $t$
- ❷  $\text{Var}(\varepsilon_t) = \sigma^2$  pentru toți  $t$
- ❸  $\text{Cov}(\varepsilon_t, \varepsilon_s) = 0$  pentru  $t \neq s$

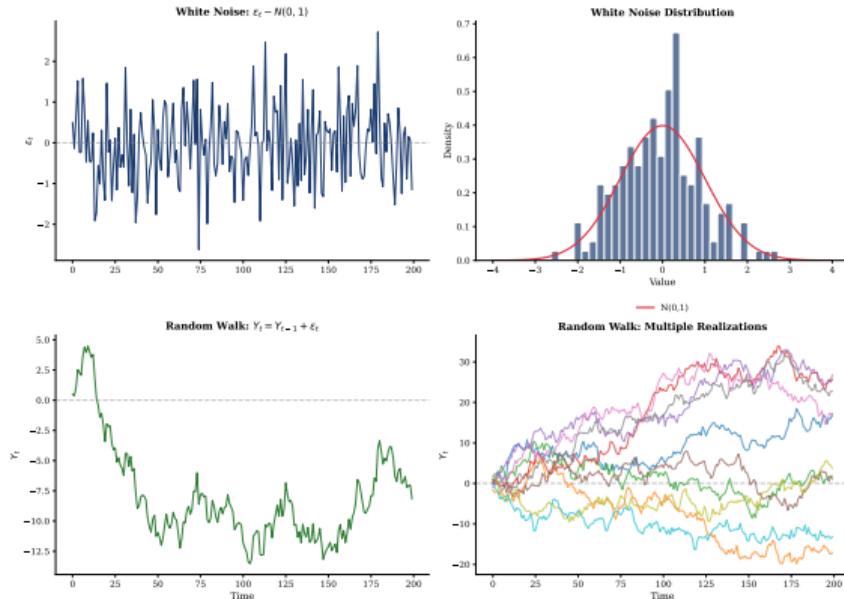
ACF al Zgomotului Alb:

$$\rho(h) = \begin{cases} 1 & \text{dacă } h = 0 \\ 0 & \text{dacă } h \neq 0 \end{cases}$$

Tipuri:

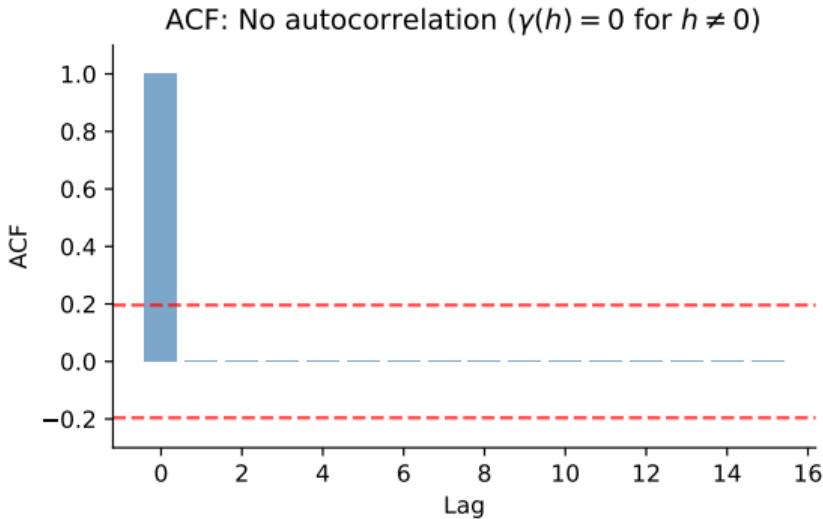
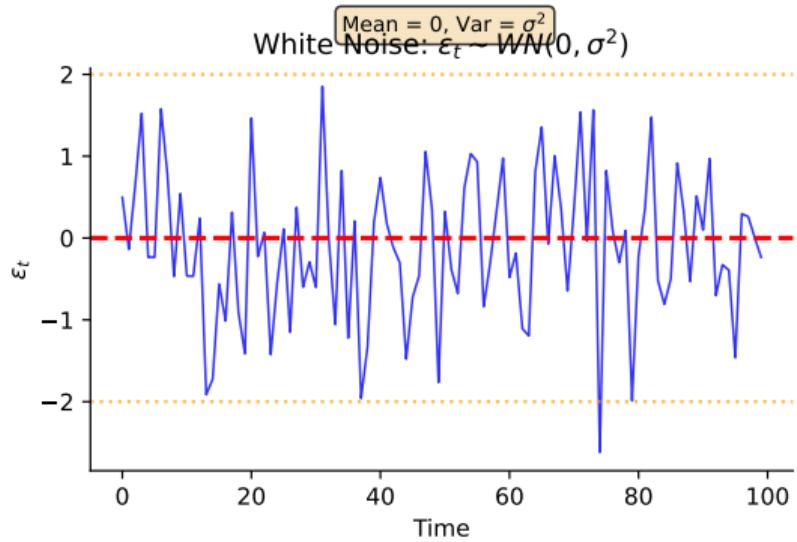
- **Zgomot alb slab:** Necorelat (condițiile de mai sus)
- **Zgomot alb puternic:** Independent și identic distribuit (i.i.d.)
- **Zgomot alb Gaussian:**  $\varepsilon_t \stackrel{iid}{\sim} N(0, \sigma^2)$

# Zgomot Alb vs Mers Aleatoriu: Comparație



- **Zgomot alb:** Fluctuează în jurul lui zero – staționar, varianță constantă
- **Mers aleatoriu:** Suma cumulativă a zgomotului alb – rătăcește, nestaționar
- Mersul aleatoriu este cel mai simplu proces nestaționar (rădăcină unitate)

## Zgomot Alb: Ilustrație Vizuală



Stânga: zgomotul alb fluctuează în jurul lui zero cu varianță constantă. Dreapta: ACF arată nicio autocorelație (toate zero după lag 0).

## Procesul de Mers Aleatoriu

**Definiție:**  $X_t = X_{t-1} + \varepsilon_t$  unde  $\varepsilon_t \sim WN(0, \sigma^2)$ ,  $X_0 = 0$

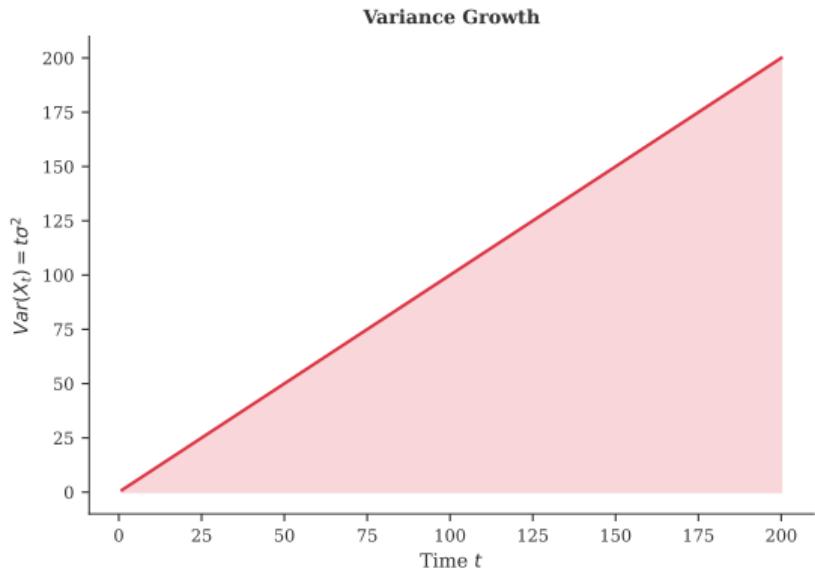
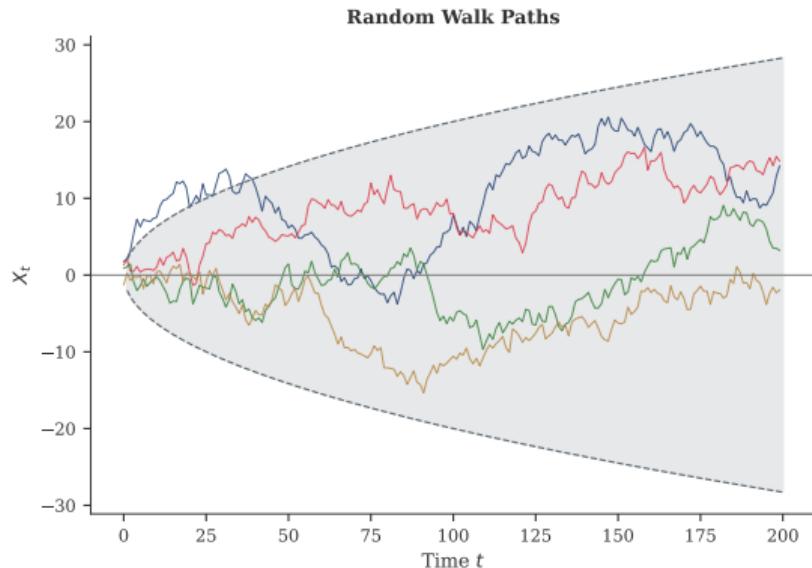
**Forma explicită:**  $X_t = \sum_{i=1}^t \varepsilon_i$

**Proprietăți:**

- $\mathbb{E}[X_t] = 0$  (medie constantă)
- $\text{Var}(X_t) = t\sigma^2$  (varianța crește în timp!)
- $\text{Cov}(X_t, X_s) = \min(t, s) \cdot \sigma^2$

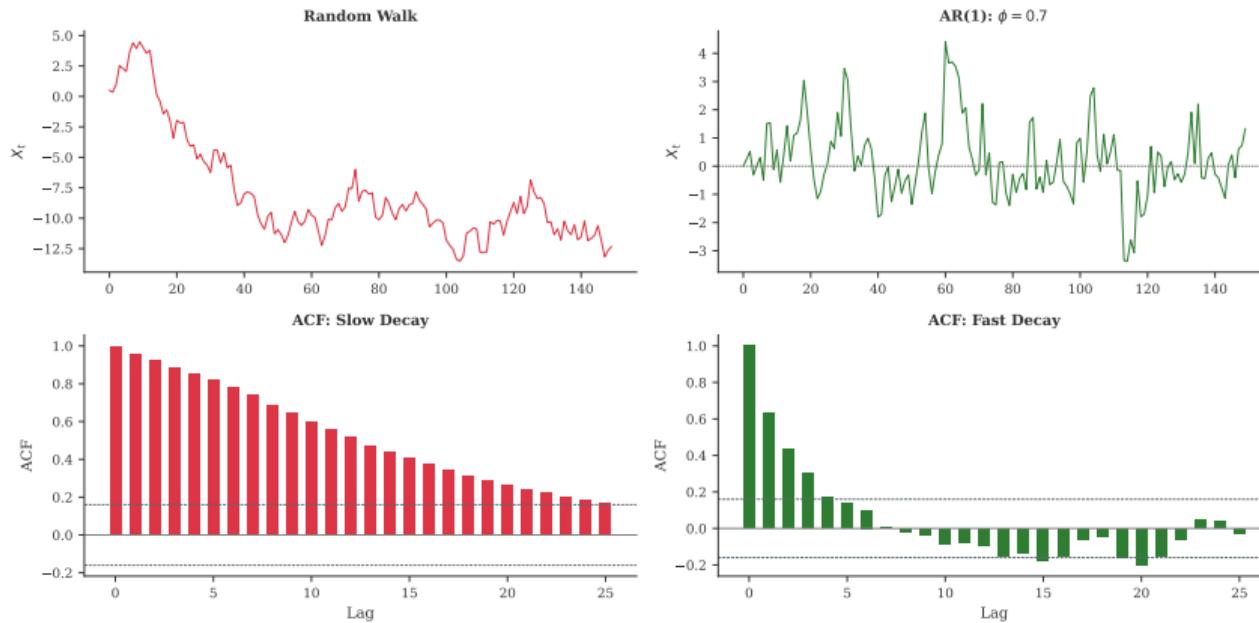
Nestaționar!

Mersul aleatoriu **nu este staționar** deoarece varianța depinde de  $t$ .



**Stânga:** Traекторii multiple divergă în timp. **Dreapta:** Varianța crește liniar:  $\text{Var}(X_t) = t\sigma^2$ .

## Staționar vs Nestaționar: Comparație



**Diagnostic cheie:** ACF al procesului staționar scade rapid; ACF al mersului aleatoriu scade foarte lent.

## Funcția de Autocorelație din Eșantion

ACF din eșantion la lag-ul  $h$ :

$$\hat{\rho}(h) = \frac{\sum_{t=1}^{T-h} (x_t - \bar{x})(x_{t+h} - \bar{x})}{\sum_{t=1}^T (x_t - \bar{x})^2}$$

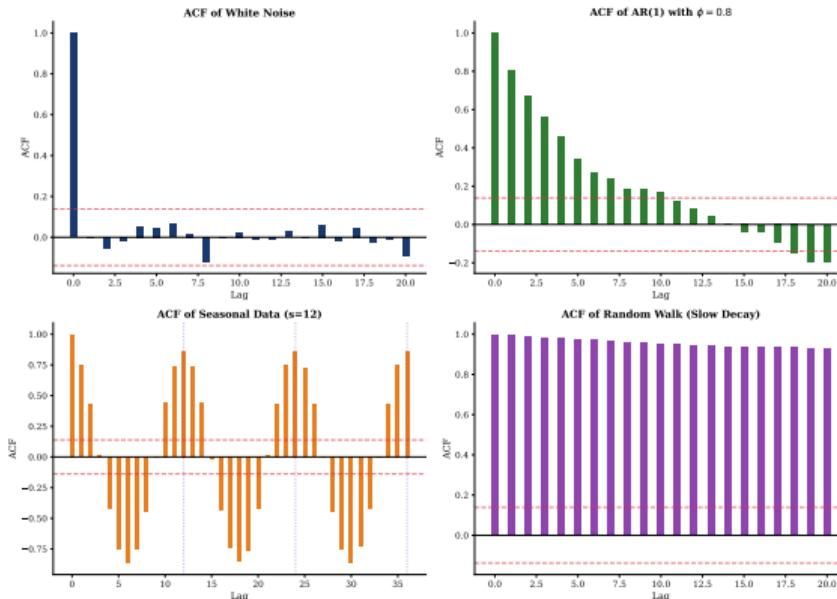
Proprietăți:

- $\hat{\rho}(0) = 1$  întotdeauna
- $|\hat{\rho}(h)| \leq 1$

Test de semnificație: Sub zgromot alb,  $\hat{\rho}(h) \approx N(0, 1/T)$

Limite 95%:  $\pm 1.96/\sqrt{T}$

## Tipare ACF pentru Diferite Procese



- **Zgomot alb:** ACF scade la zero imediat (nicio dependență)
- **AR(1):** ACF scade exponentiațial – indică structură autoregresivă
- **Sezonier:** ACF arată vârfuri la lag-uri sezoniere (de ex., 12, 24 pentru lunar)
- **Mers aleatoriu:** ACF scade foarte lent – semn de nestaționaritate

## Funcția de Autocorelație Parțială (PACF)

**PACF**  $\phi_{hh}$ : Corelația dintre  $X_t$  și  $X_{t+h}$  după eliminarea efectului liniar al  $X_{t+1}, \dots, X_{t+h-1}$ .

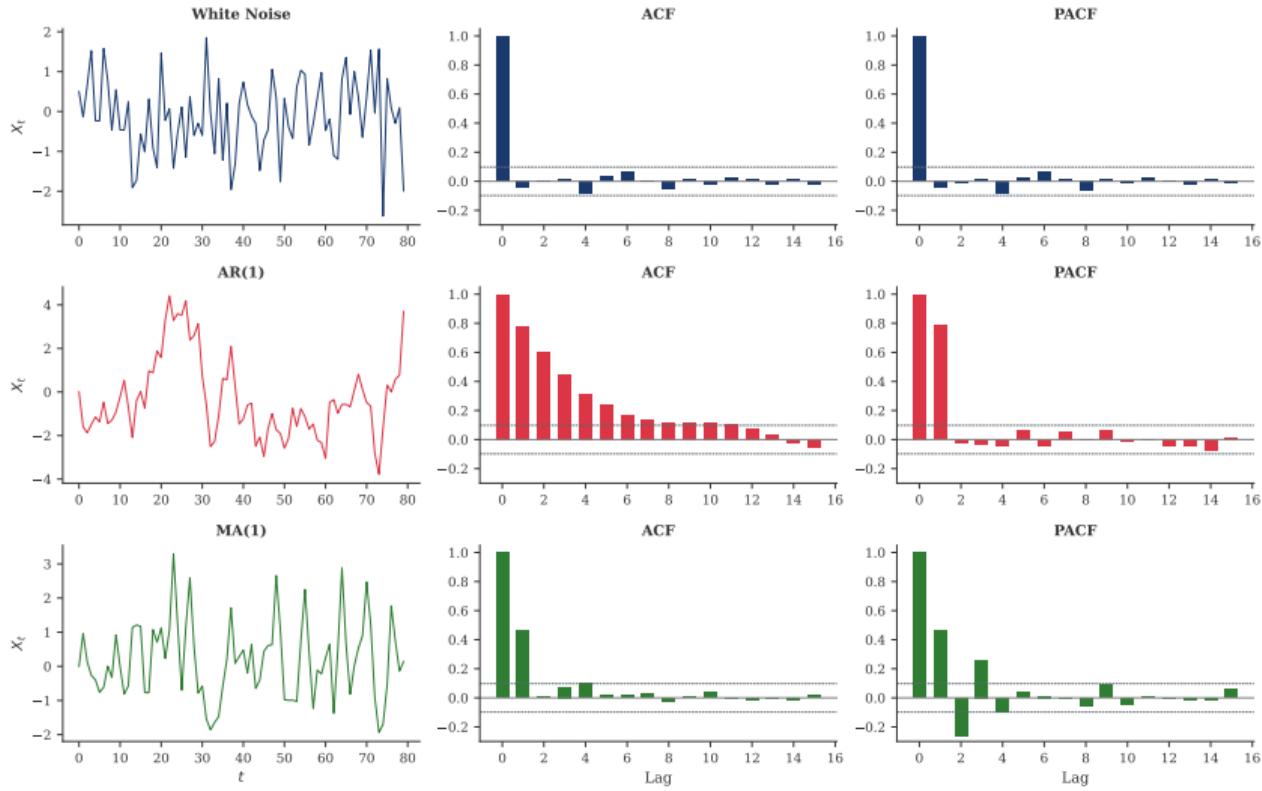
**Interpretare:**

- $\phi_{11} = \rho(1)$  (același ca ACF la lag 1)
- $\phi_{22} =$  corelația lui  $X_t, X_{t+2}$  controlând pentru  $X_{t+1}$
- Măsoară dependența *directă* la lag-ul  $h$

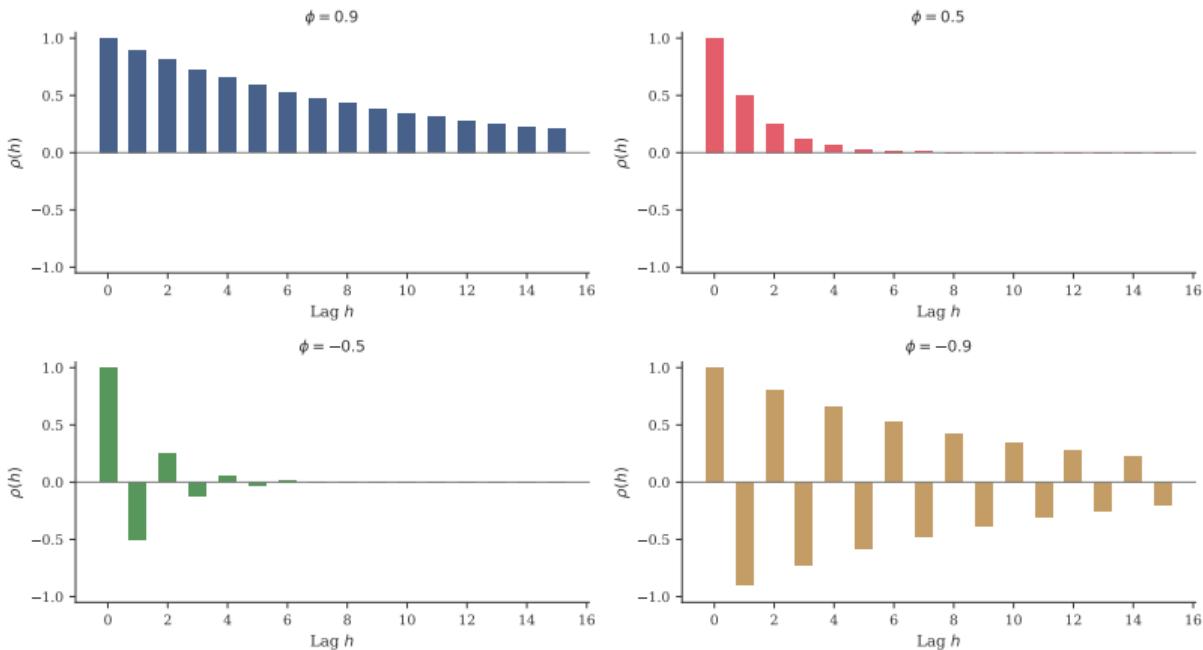
**Aplicație cheie:** Identificarea ordinului AR

- Pentru  $AR(p)$ : PACF se **întrerupe** după lag-ul  $p$
- Pentru  $MA(q)$ : ACF se **întrerupe** după lag-ul  $q$

# Tipare ACF și PACF



## ACF Teoretic pentru AR(1)



Pentru AR(1):  $X_t = \phi X_{t-1} + \varepsilon_t$ , ACF teoretic este  $\rho(h) = \phi^h$ .

## Definiție 9 (Operatorul Lag)

**Operatorul lag** (sau operatorul de întârziere)  $L$  este definit prin:

$$LX_t = X_{t-1}$$

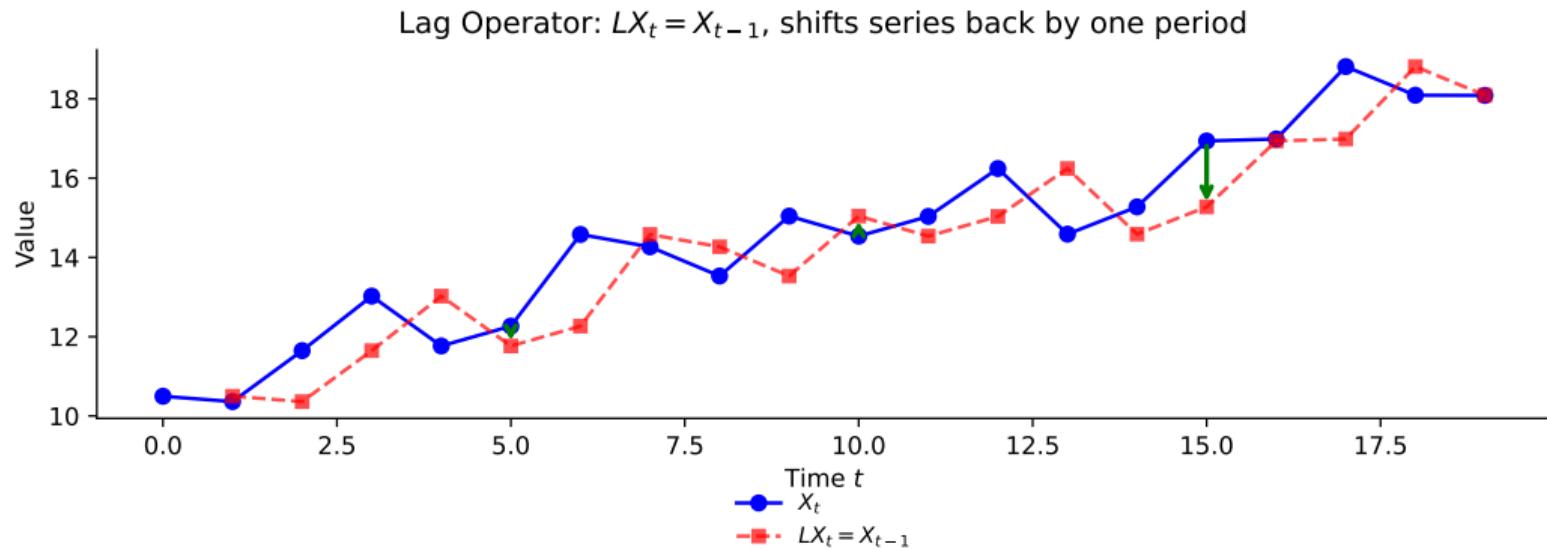
### Proprietăți:

- $L^k X_t = X_{t-k}$  (întârziere cu  $k$  perioade)
- $L^0 = I$  (identitate)
- $(1 - \phi L)X_t = X_t - \phi X_{t-1}$

### Exemple:

- AR(1):  $(1 - \phi L)X_t = \varepsilon_t$
- MA(1):  $X_t = (1 + \theta L)\varepsilon_t$
- AR( $p$ ):  $(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p)X_t = \varepsilon_t$

## Operatorul Lag: Ilustrație Vizuală



Operatorul lag  $L$  deplasează fiecare observație înapoi cu o perioadă de timp:  $LX_t = X_{t-1}$ .

**Prima diferență:**  $\Delta X_t = X_t - X_{t-1} = (1 - L)X_t$

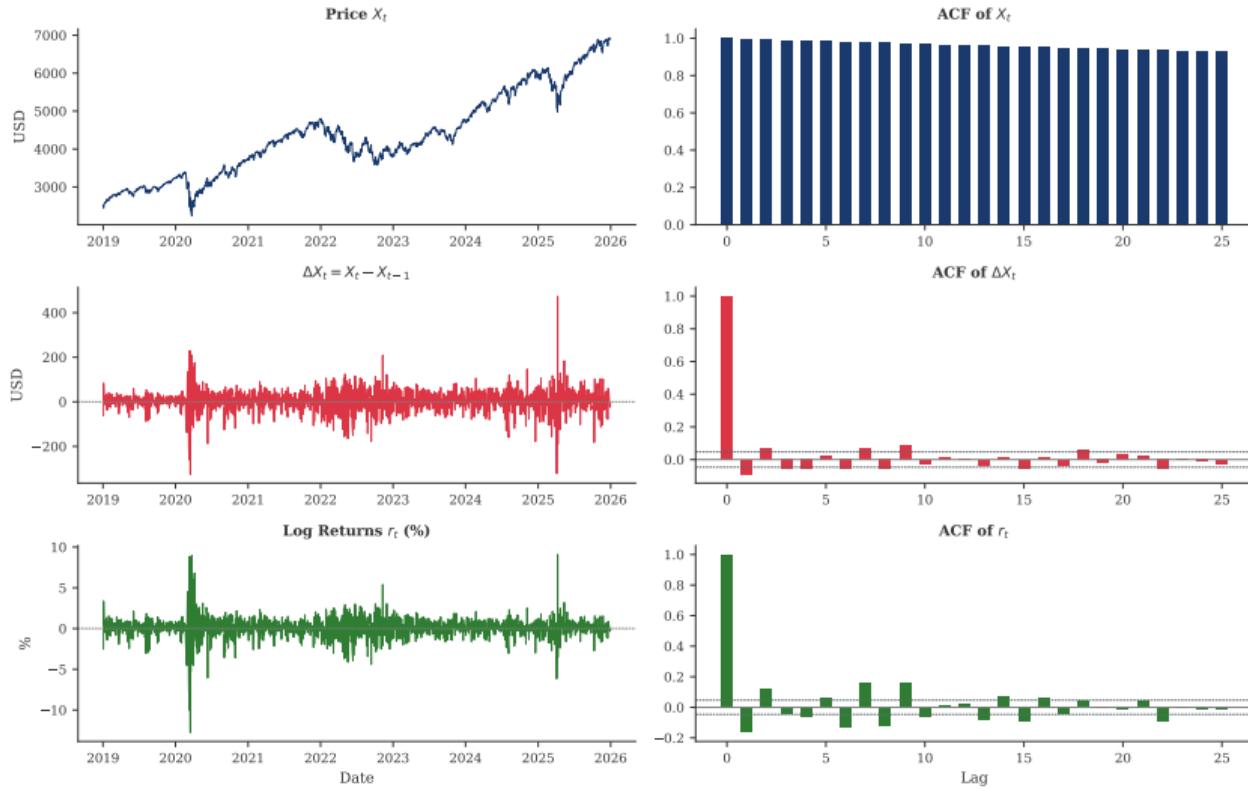
**De ce diferențiem?**

- Elimină trendul și rădăcina unitate
- Mers aleatoriu:  $\Delta X_t = \varepsilon_t$  (zgomot alb)

**Proces integrat:**  $X_t \sim I(d)$  dacă  $\Delta^d X_t$  este staționar

- $I(0)$ : Staționar (nu necesită diferențiere)
- $I(1)$ : Necesită o diferențiere
- $I(2)$ : Necesită două diferențieri

# Efectul Diferențierii: S&P 500



# Testul Augmented Dickey-Fuller (ADF)

**Model:**  $\Delta X_t = \alpha + \gamma X_{t-1} + \sum_{i=1}^p \delta_i \Delta X_{t-i} + \varepsilon_t$

**Ipoteze:**

- $H_0: \gamma = 0$  (rădăcină unitate)
- $H_1: \gamma < 0$  (staționar)

**Statistica de test:**

$$\tau = \frac{\hat{\gamma}}{SE(\hat{\gamma})}$$

**Decizie:**

- $\tau <$  valoare critică  $\Rightarrow$  Respingem  $H_0 \Rightarrow$  Stationar
- $\tau \geq$  valoare critică  $\Rightarrow$  Nestaționar

Valori critice: distribuția Dickey-Fuller (nu normală)

## Testul KPSS

**Model:**  $X_t = \xi t + r_t + \varepsilon_t$  unde  $r_t = r_{t-1} + u_t$

**Ipoteze (opuse față de ADF):**

- $H_0: \sigma_u^2 = 0$  (staționar)
- $H_1: \sigma_u^2 > 0$  (rădăcină unitate)

**Statistica de test:**

$$LM = \frac{\sum_{t=1}^T S_t^2}{T^2 \hat{\sigma}^2}$$

unde  $S_t = \sum_{i=1}^t \hat{e}_i$

**Decizie:**

- $LM >$  valoare critică  $\Rightarrow$  Respingem  $H_0 \Rightarrow$  **Nestaționar**
- $LM \leq$  valoare critică  $\Rightarrow$  **Staționar**

**Notă:** KPSS complementează ADF—folosiți ambele pentru concluzii robuste.

# Folosirea ADF și KPSS Împreună

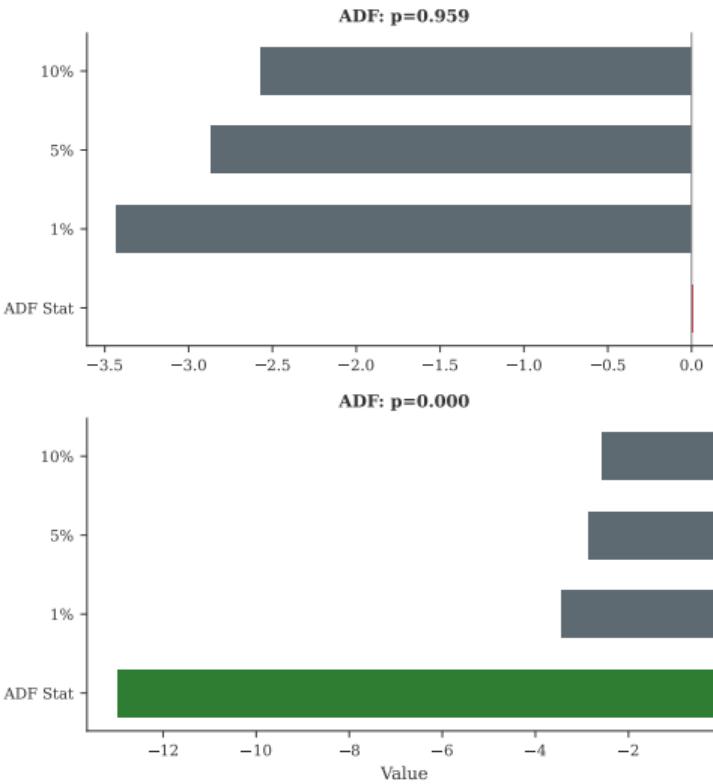
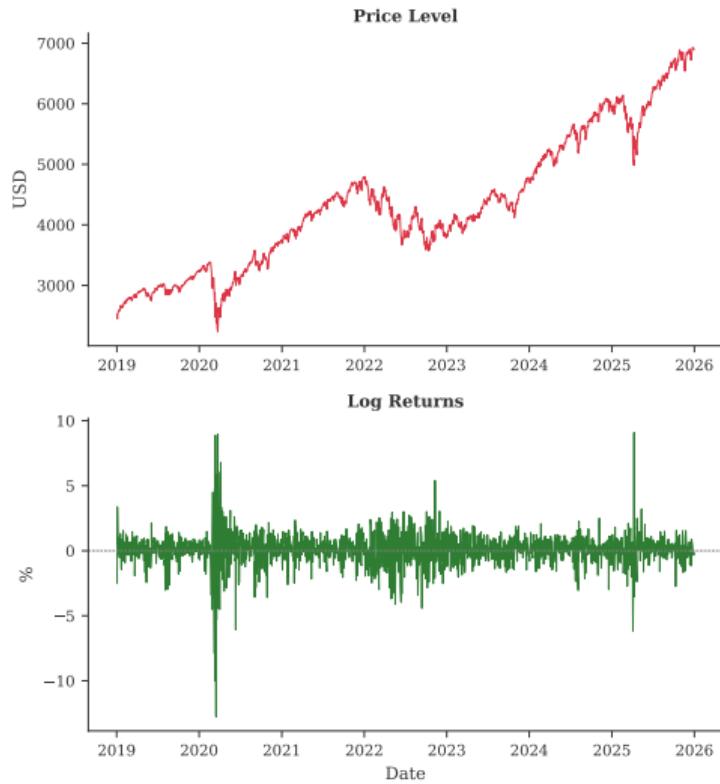
Testare confirmatorie pentru concluzii robuste:

ADF	KPSS	Concluzie
Respingem $H_0$	Nu respingem $H_0$	Staționar
Nu respingem $H_0$	Respingem $H_0$	Rădăcină Unitate
Respingem $H_0$	Respingem $H_0$	Neconcludent
Nu respingem $H_0$	Nu respingem $H_0$	Neconcludent

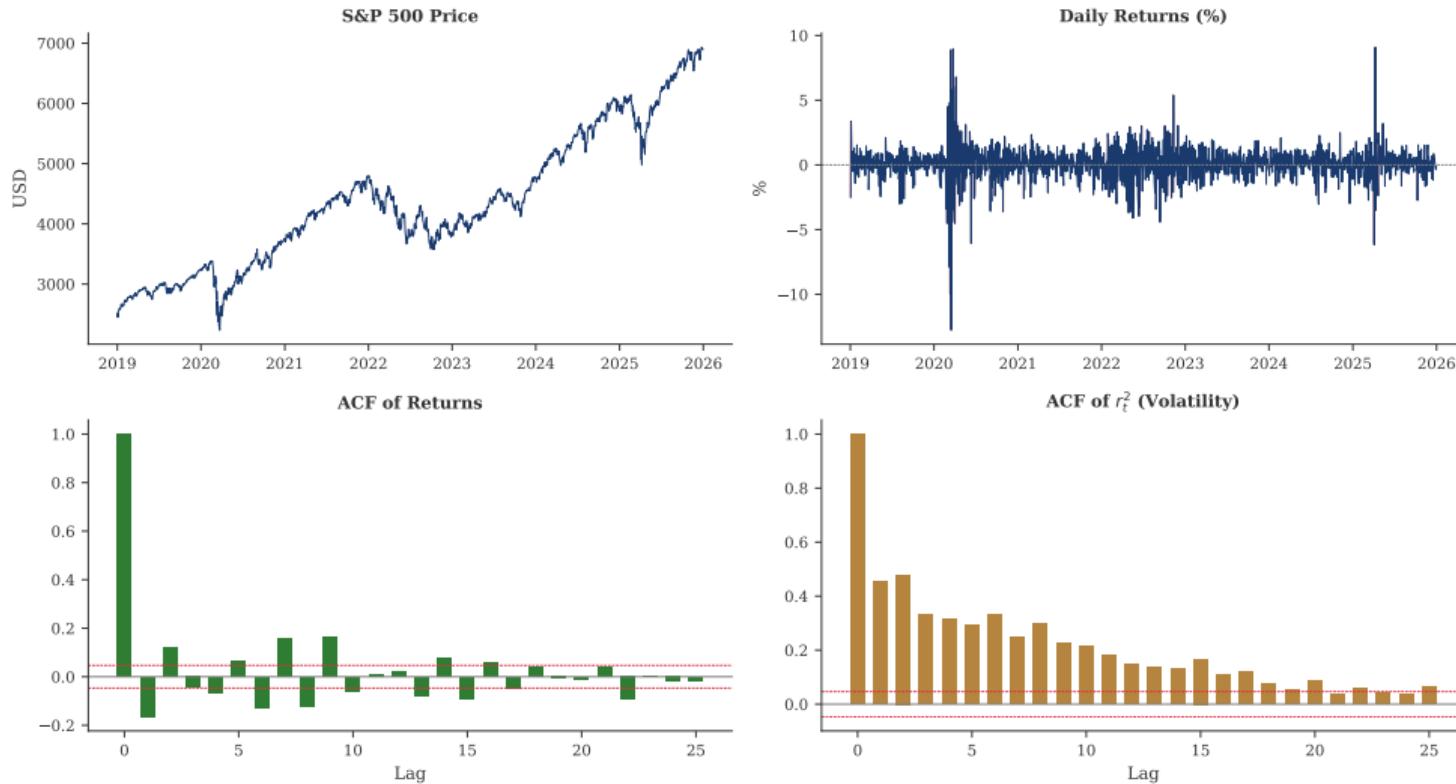
Flux de lucru recomandat:

- ① Rulați testul ADF (nulă = rădăcină unitate)
- ② Rulați testul KPSS (nulă = staționar)
- ③ Dacă rezultatele coincid, procedați cu încredere
- ④ Dacă neconcludent, considerați teste alternative (PP, DF-GLS)

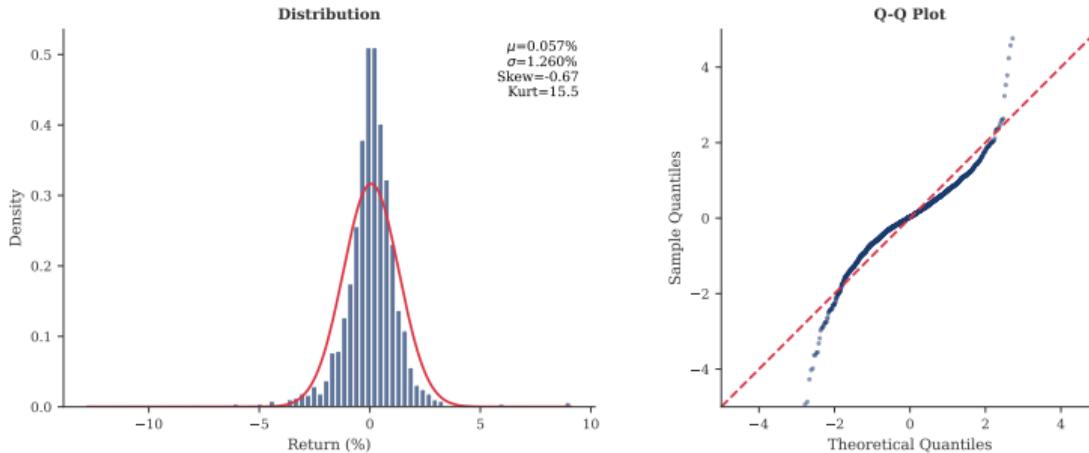
# Testul ADF: Vizualizare cu S&P 500



# Analiza S&P 500: Prezentare Generală



# Fapte Stilizate ale Randamentelor Financiare



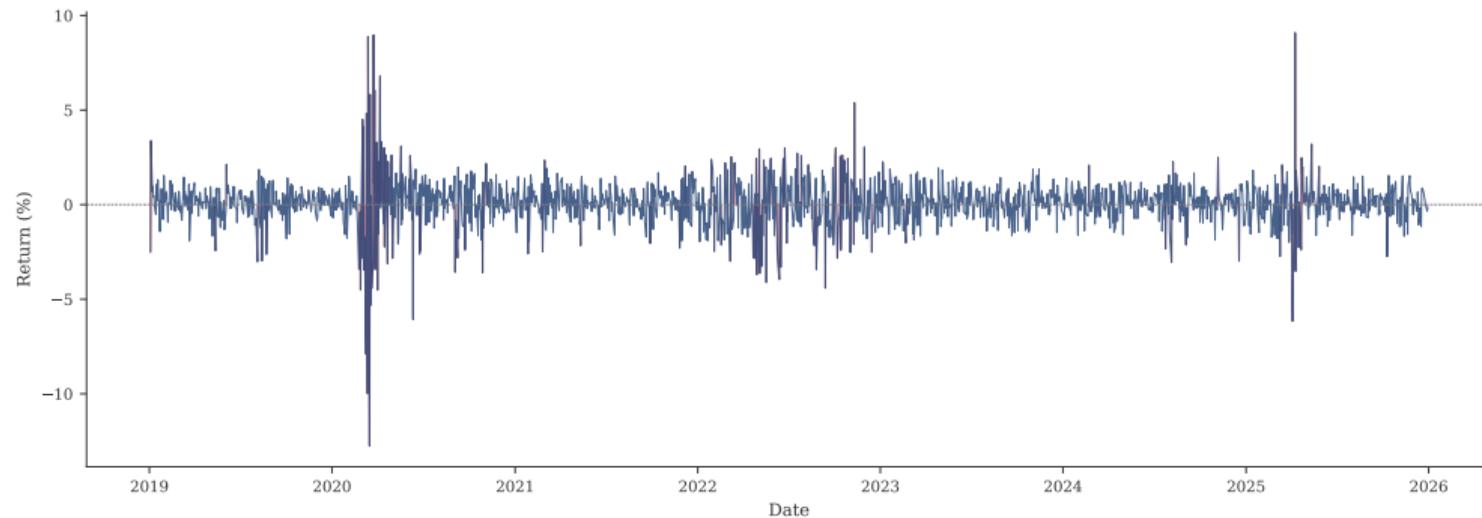
## Proprietăți observate:

- Asimetrie negativă (coadă stângă)
- Kurtoză excesivă ( $\gg 3$ )
- Cozi groase (heavy tails)

## Implicații:

- Distribuția normală inadecvată
- Evenimente extreme mai probabile
- Necesită distribuție Student-t sau similară

## Gruparea Volatilității



### Fapt Stilizat

Randamentele mari (pozitive sau negative) tend să fie urmate de randamente mari. Această **grupare a volatilității** motivează modelele ARCH/GARCH (capitolele viitoare).

- ① **Serie de timp** = observații indexate după timp cu dependență temporală
- ② **Descompunere**: Aditivă  $X_t = T_t + S_t + \varepsilon_t$  sau Multiplicativă
- ③ **Netezire Exponențială**: SES (nivel), Holt (trend), Holt-Winters (sezonier)
- ④ **Evaluare Prognoză**: MAE, RMSE, MAPE; folosiți separări train/validare/test
- ⑤ **Modelarea Sezonalității**: Variabile dummy (orice tipar) sau termeni Fourier (neted)
- ⑥ **Gestionarea Trendului**: Diferențiere (stochastic) sau regresie (determinist)
- ⑦ **Staționaritate**: Media, varianța, autocovarianța constante în timp
- ⑧ **ACF/PACF**: Esențiale pentru identificarea structurii de dependență
- ⑨ **Teste rădăcină unitate**: ADF ( $H_0$ : rădăcină unitate) vs KPSS ( $H_0$ : staționar)

## Formule Importante I

### Descompunere

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

### Netezire Exponențială Simplă (SES)

$$\hat{X}_{t+1|t} = \alpha X_t + (1 - \alpha) \hat{X}_{t|t-1} \quad \text{unde } \alpha \in (0, 1)$$

### Trend Liniar Holt

$$\ell_t = \alpha X_t + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \quad b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1}$$

### Holt-Winters Aditivă

$$\ell_t = \alpha(X_t - S_{t-s}) + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \quad S_t = \gamma(X_t - \ell_t) + (1 - \gamma)S_{t-s}$$

## Formule Importante II

### Media Mobilă (Estimare Trend)

$$\hat{T}_t = \frac{1}{2q+1} \sum_{j=-q}^q X_{t+j}$$

### Autocovarianță și Autocorelație

$$\gamma(h) = \text{Cov}(X_t, X_{t+h}) \quad \rho(h) = \frac{\gamma(h)}{\gamma(0)}$$

### Mers Aleatoriu

$$X_t = X_{t-1} + \varepsilon_t \quad \Rightarrow \quad \text{Var}(X_t) = t\sigma^2 \text{ (nestaționar)}$$

### Diferențiere

$$\Delta X_t = (1 - L)X_t = X_t - X_{t-1}$$

## Capitolul 2: Modele ARMA

- Modele Autoregresive (AR)
- Modele de Medie Mobilă (MA)
- Modele ARMA combinate
- Identificarea modelului folosind ACF/PACF
- Estimarea parametrilor
- Diagnosticarea modelului
- Prognoza

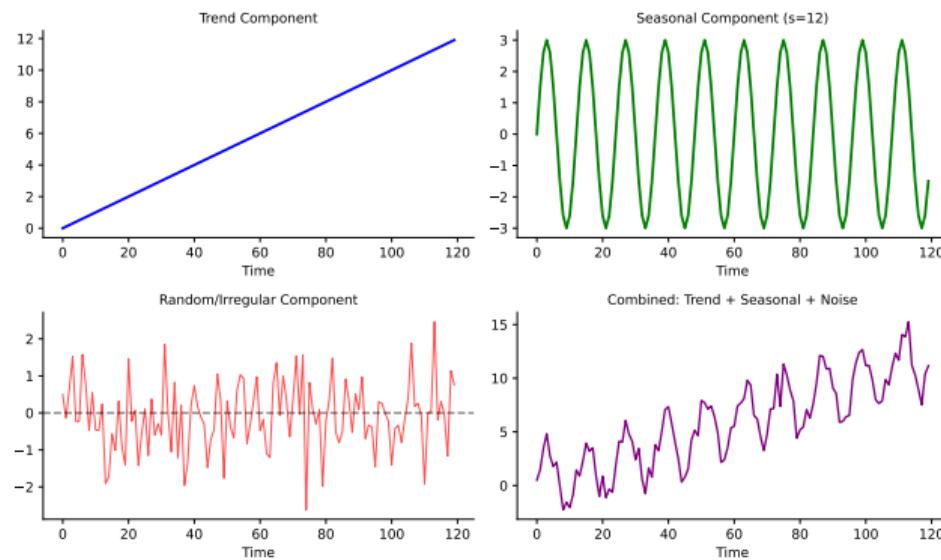
## Întrebarea Quiz 1

### Întrebare

O serie de timp  $Y_t$  arată mișcare ascendentă de-a lungul anilor plus tipare repetitive în fiecare trimestru. Ce componente sunt prezente?

- A Doar trend
- B Doar sezonalitate
- C Trend și Sezonalitate
- D Doar zgomot aleatoriu

## Întrebarea Quiz 1: Răspuns



Răspuns Corect: (C) Trend și Sezonalitate

Mișcare ascendentă = Trend; Tipare trimestriale = Sezonalitate ( $s=4$ )

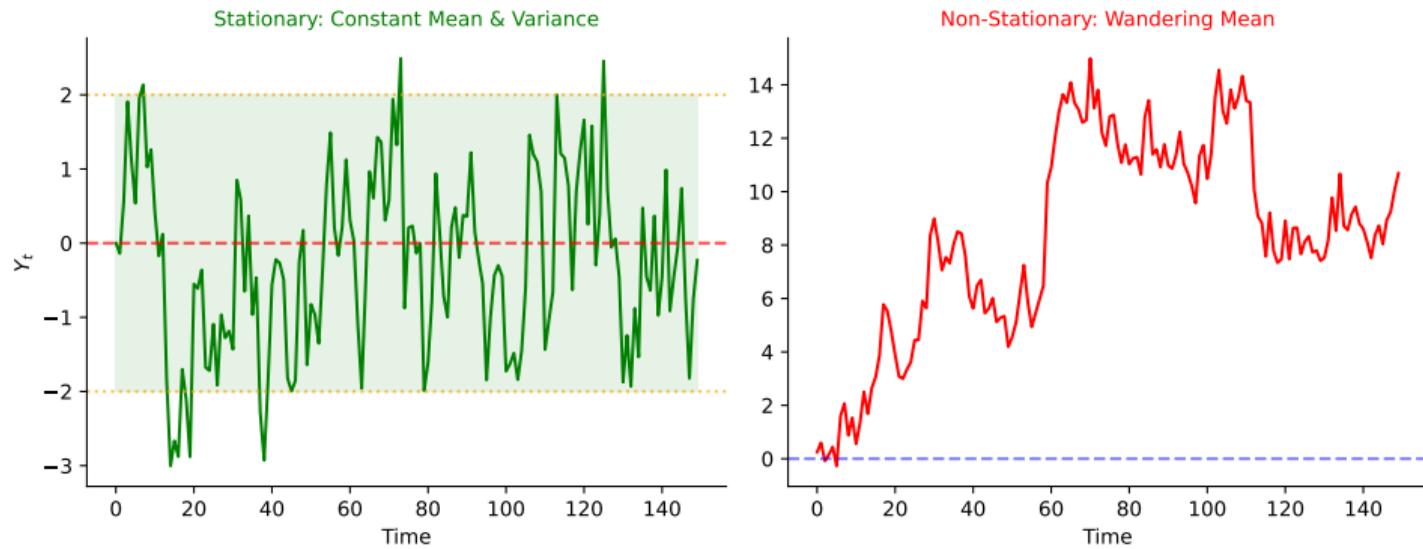
## Întrebarea Quiz 2

### Întrebare

Care dintre următoarele este o caracteristică a unei serii de timp staționare?

- A Media se schimbă în timp
- B Varianța crește în timp
- C Medie și varianță constante în timp
- D Conține o componentă de trend

## Întrebarea Quiz 2: Răspuns



Răspuns Corect: (C) Medie și varianță constantă în timp

Staționaritatea necesită: medie constantă, varianță constantă și autocovarianță depinde doar de lag.

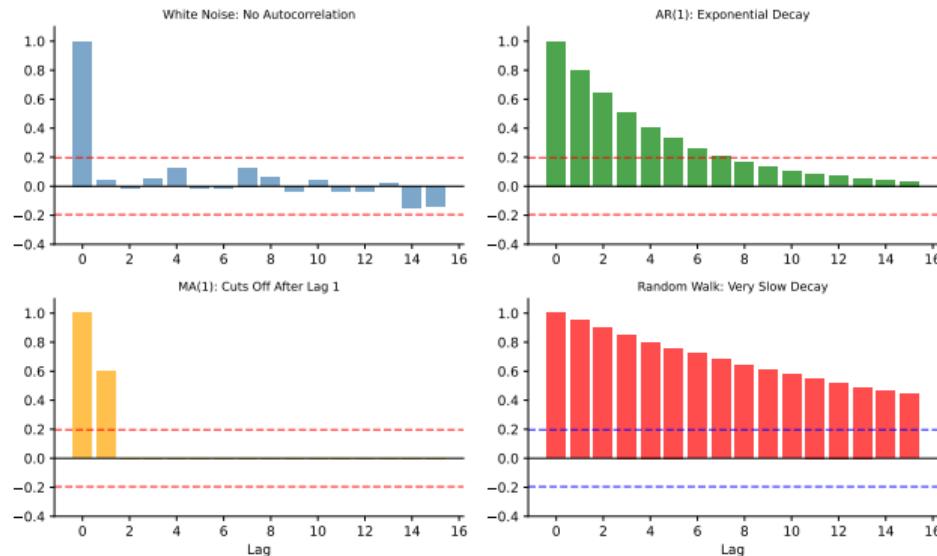
## Întrebarea Quiz 3

### Întrebare

Pentru un proces de zgomot alb, cum arată ACF la lag-uri  $k > 0$ ?

- A Descreștere exponențială
- B Toate valorile semnificative și pozitive
- C Toate valorile aproximativ zero (în interiorul benzilor de încredere)
- D Alternare pozitiv și negativ

## Întrebarea Quiz 3: Răspuns



Răspuns Corect: (C) Aproximativ zero în interiorul benzilor de încredere

Zgomotul alb nu are autocorelație:  $\rho_k = 0$  pentru toți  $k \neq 0$ .

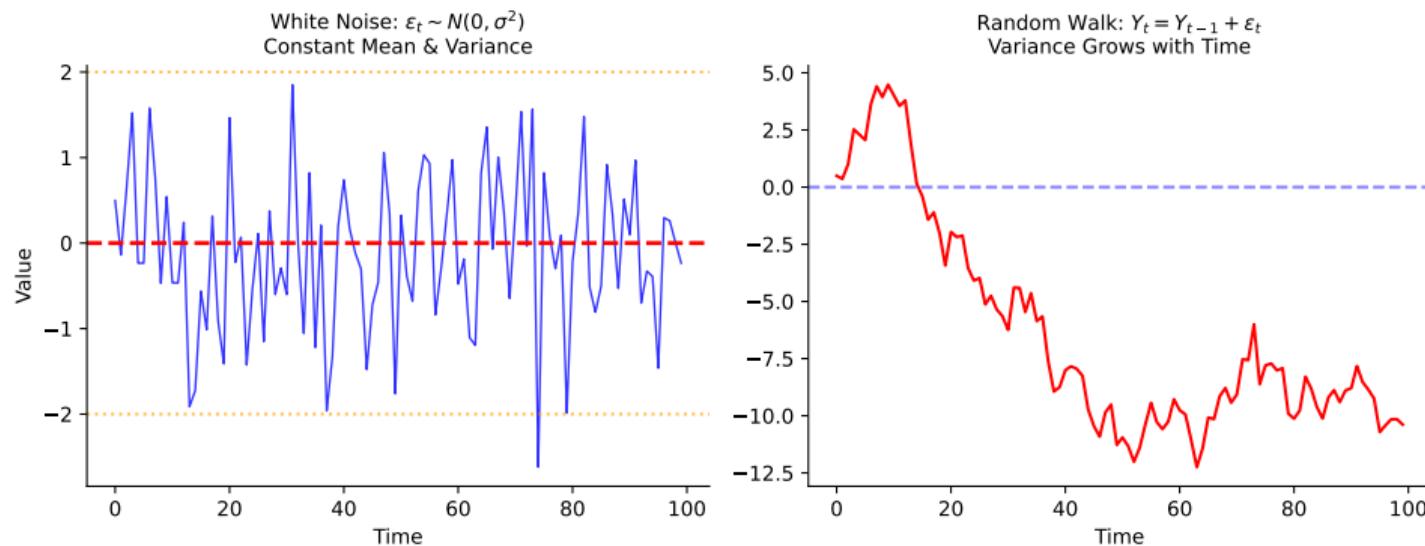
## Întrebarea Quiz 4

### Întrebare

Care este diferența cheie între zgomotul alb și mersul aleatoriu?

- A Zgomotul alb are trend, mersul aleatoriu nu
- B Mersul aleatoriu este suma cumulativă a zgomotului alb
- C Ambele sunt procese staționare
- D Zgomotul alb are varianță mai mare

## Întrebarea Quiz 4: Răspuns



Răspuns Corect: (B) Mers aleatoriu = suma cumulativă a zgomotului alb

$$Y_t = Y_{t-1} + \varepsilon_t = \sum_{i=1}^t \varepsilon_i \text{ unde } \varepsilon_t \text{ este zgomot alb.}$$

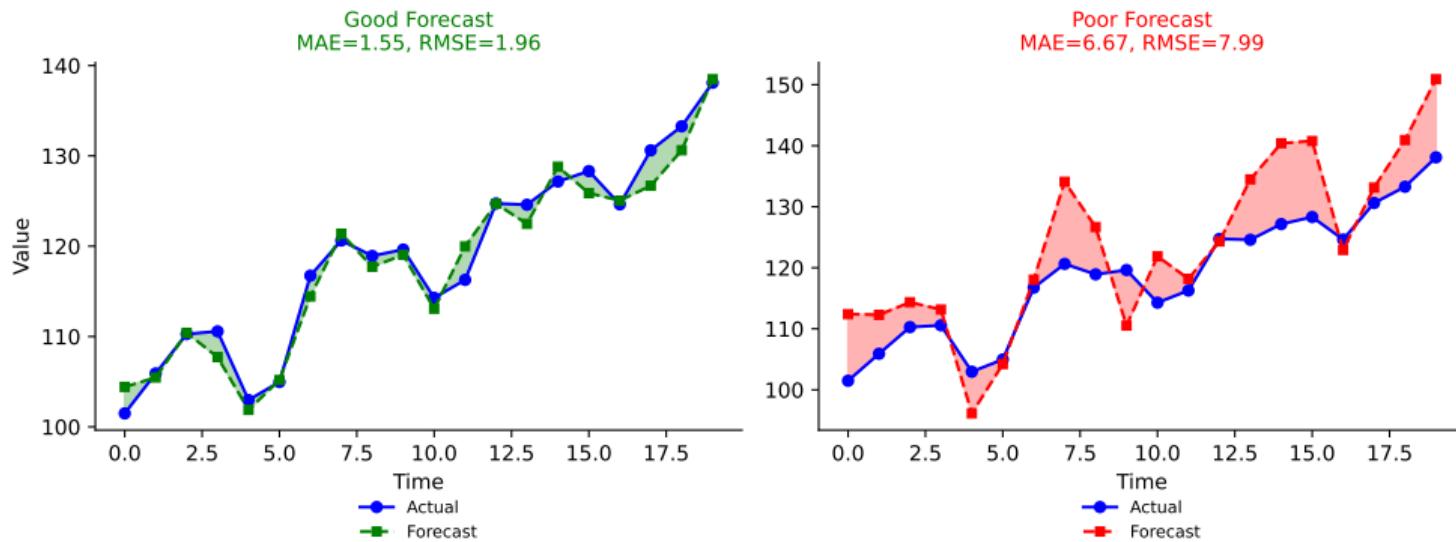
## Întrebarea Quiz 5

### Întrebare

Care metrică de eroare a proguzei este cea mai sensibilă la erori mari (valori aberante)?

- A MAE (Eroarea Medie Absolută)
- B RMSE (Rădăcina Erorii Medii Pătratice)
- C MAPE (Eroarea Medie Absolută Procentuală)
- D Toate sunt la fel de sensibile

## Întrebarea Quiz 5: Răspuns



Răspuns Corect: (B) RMSE

RMSE ridică la pătrat erorile, deci erorile mari au impact disproportional:  $\sqrt{\frac{1}{n} \sum e_t^2}$

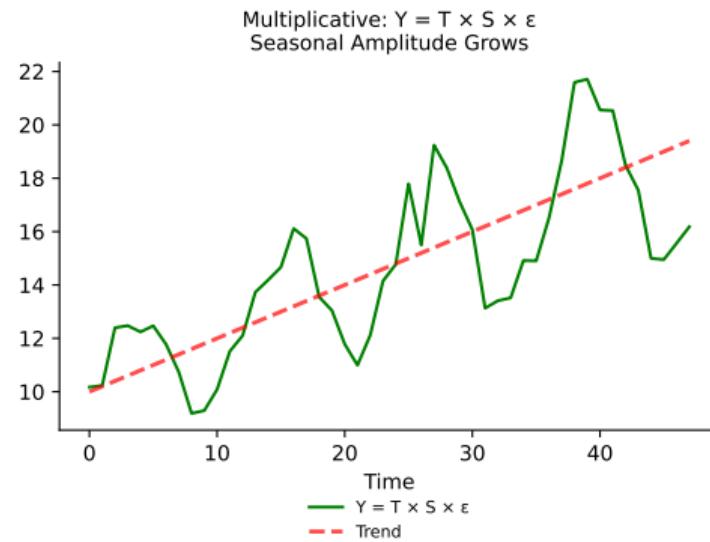
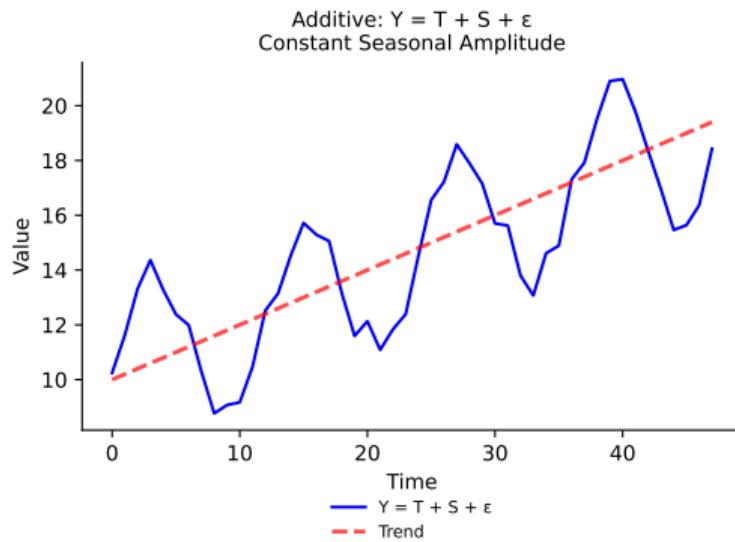
## Întrebarea Quiz 6

### Întrebare

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

- A Când seria nu are trend
- B Când amplitudinea sezonieră este constantă
- C Când amplitudinea sezonieră crește odată cu nivelul seriei
- D Când seria este staționară

## Întrebarea Quiz 6: Răspuns



Răspuns Corect: (C) Amplitudinea sezonieră crește odată cu nivelul

Multiplicativă:  $Y_t = T_t \times S_t \times \epsilon_t$  — oscilațiile sezoniere proporționale cu trendul.

## Referințe

-  Hyndman, R.J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice*. 3rd ed., OTexts.
-  Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press.
-  Box, G.E.P., Jenkins, G.M., Reinsel, G.C., & Ljung, G.M. (2015). *Time Series Analysis: Forecasting and Control*. 5th ed., Wiley.
-  Tsay, R.S. (2010). *Analysis of Financial Time Series*. 3rd ed., Wiley.
-  Cleveland, R.B., Cleveland, W.S., McRae, J.E., & Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition. *Journal of Official Statistics*, 6(1), 3-73.

## Date Reale Utilizate în Acest Capitol

- **Pasageri Aviație:** Set de date clasic Box-Jenkins, 1949–1960
- **S&P 500:** Yahoo Finance (SPY), date istorice
- **Pete Solare:** Set de date Statsmodels, observații lunare

## Software și Instrumente

- **Python:** statsmodels, pandas, matplotlib, yfinance
- **R:** pachetele forecast, tseries
- **Surse de Date:** Yahoo Finance, FRED Economic Data

# Vă Mulțumesc!

Întrebări?

*Grafice generate folosind Python (statsmodels, matplotlib)*

Materiale curs disponibile la: <https://github.com/danpele/Time-Series-Analysis>