



# Analiza și Prognoza Seriilor de Timp

Capitolul 1: Procese Stochastice și Staționaritate



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

## Obiective de învățare

La finalul acestui capitol, veți fi capabili să:

1. Definiți procesele stochastice și să înțelegeți proprietățile acestora
2. Distingeți între staționaritatea strictă și slabă (covarianță)
3. Identificați procesele de zgomot alb și mers aleatoriu
4. Calculați și interpretați ACF (Funcția de Autocorelație) și PACF (Funcția de Autocorelație Parțială)
5. Aplicați operatorul lag și diferențierea
6. Efectuați teste de staționaritate: ADF (Augmented Dickey-Fuller) și KPSS (Kwiatkowski-Phillips-Schmidt-Shin)
7. Analizați date financiare de tip serie de timp
8. Distingeți între procesele cu rădăcină unitate și cele staționare în trend

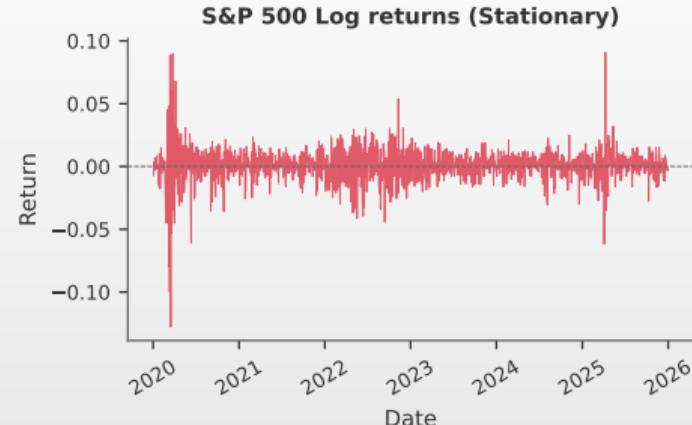


## Cuprins

- Motivație
- Procese Stochastice
- Staționaritate
- Operatorul lag și Diferențierea
- Zgomot Alb și Mers Aleatoriu
- Funcții de Autocorelație
- Testarea Staționarității
- Aplicație pe Date Financiare
- Studiu de Caz: Testarea Staționarității
- Utilizare IA
- Rezumat
- Quiz
- Bibliografie



## Exemple: serii staționare vs. nestaționare

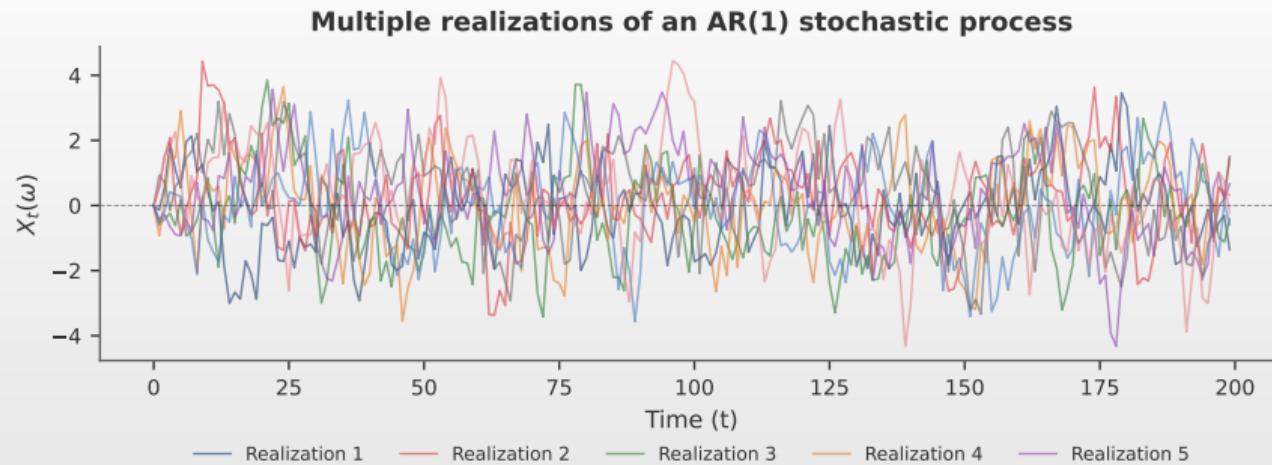


### Observații

- Prețurile (stânga) sunt nestaționare: trend, media se schimbă în timp
- Randamentele (dreapta) sunt staționare: medie  $\approx 0$ , varianță aprox. constantă
- Randamente log:  $r_t = \ln P_t - \ln P_{t-1} \Rightarrow$  nestaționar  $\rightarrow$  staționar



## Proces stochastic: ilustrare vizuală



### Interpretare

- Fiecare linie este o **realizare diferită** din același proces stochastic subiacent
- Observăm doar o **singură realizare**, dar vrem să înțelegem proprietățile procesului



## Proces stochastic: definiție

### Definiție 1 (Proces Stochastic)

- Un **proces stochastic** este o colecție de variabile aleatoare indexate după timp
  - ▶  $\{X_t(\omega) : t \in \mathcal{T}, \omega \in \Omega\}$
  - ▶  $\Omega$  este spațiul eșantion al rezultatelor posibile

### Două perspective

- **$\omega$  fixat:** O realizare  $\{X_t(\omega)\}_{t \in \mathcal{T}}$
- **$t$  fixat:** O variabilă aleatoare  $X_t$

### De reținut

- O serie de timp pe care o observăm este o **singură realizare** a procesului stochastic subiacent



## Momentele unui proces stochastic

Primele două momente caracterizează procesul

- **Funcția de Medie:**  $\mu_t = \mathbb{E}[X_t]$
- **Autocovarianță (ACVF):**  $\gamma(t, s) = \text{Cov}(X_t, X_s)$ 
  - ▶  $\gamma(t, s) = \mathbb{E}[(X_t - \mu_t)(X_s - \mu_s)]$
- **Autocorelația (ACF):**
  - ▶  $\rho(t, s) = \gamma(t, s) / \sqrt{\text{Var}(X_t) \cdot \text{Var}(X_s)}$

### Proprietăți ACF

- **Interval:**  $\rho(t, s) \in [-1, 1]$
- **Normalizare:**  $\rho(t, t) = 1$  (corelație perfectă cu sine)

### Important

- **General:**  $\mu_t$  și  $\gamma(t, s)$  pot depinde de  $t$
- **Staționar:** Elimină această dependență



## De ce contează staționaritatea

### Fără staționaritate

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

### Cu staționaritate

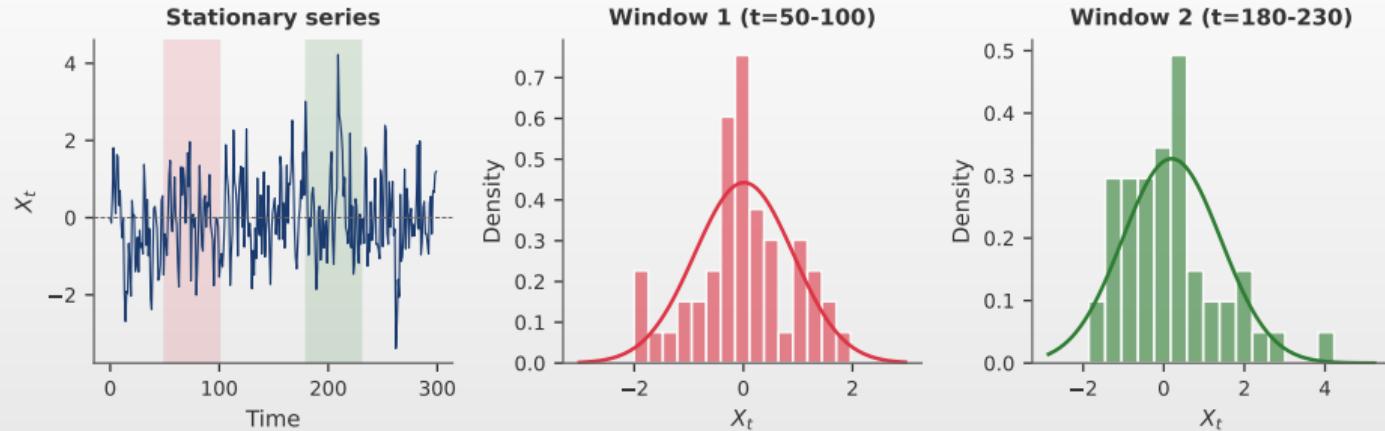
- Proprietăți statistice constante
  - ▶ Ergodicitate justificată
- Putem estima dintr-o singură realizare
- Inferență validă posibilă
- Modelele au sens

### Principiu fundamental

- Majoritatea modelelor de serii de timp necesită staționaritate
- Exemple: ARMA (AutoRegressive Moving Average), ARIMA (AutoRegressive Integrated Moving Average)
- Seriile nestaționare trebuie transformate (de ex., diferențiere) înainte de modelare



## Staționaritatea strictă: ilustrare vizuală



### Interpretare

- Translația în timp nu schimbă distribuția comună a variabilelor
- Oricare două ferestre temporale au aceleași proprietăți statistice
- În practică: verificăm doar primele momente (staționaritate slabă)

## Staționaritatea strictă

### Definiție 2 (Staționaritate Strictă (Puternică))

- Un proces  $\{X_t\}$  este **strict staționar** dacă pentru orice  $k$ , orice  $t_1, \dots, t_k$ , și orice  $h$ :
  - ▶  $(X_{t_1}, \dots, X_{t_k}) \stackrel{d}{=} (X_{t_1+h}, \dots, X_{t_k+h})$
- **Notătie:**  $X \stackrel{d}{=} Y$  înseamnă *egalitate în distribuție*
  - ▶  $P(X \leq x) = P(Y \leq x)$

### Implicații

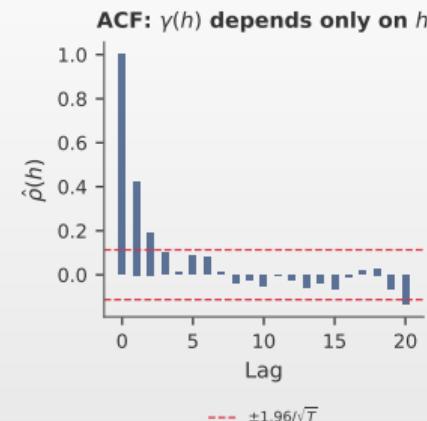
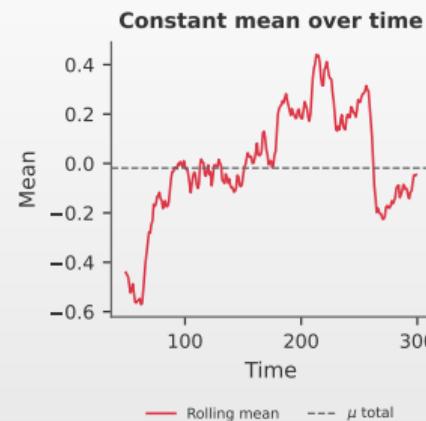
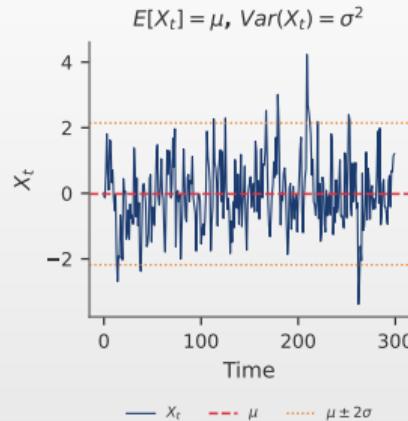
- **Distribuții identice:**  $F_{X_t}(x)$  nu depinde de  $t$ 
  - ▶  $\mathbb{E}[X_t] = \mu$  (medie constantă, dacă există)
  - ▶  $\text{Var}(X_t) = \sigma^2$  (varianță constantă, dacă există)
- **Dependență de lag:** Distribuțiile comune depind doar de lag

### Notă

- Staționaritatea strictă este o condiție puternică, adesea imposibil de verificat în practică



## Staționaritatea slabă: ilustrare vizuală



### Cele trei condiții

- $E[X_t] = \mu$  constantă  $\Rightarrow$  media nu depinde de timp
- $\text{Var}(X_t) = \sigma^2$  constantă  $\Rightarrow$  varianța nu depinde de timp
- $\text{Cov}(X_t, X_{t+h}) = \gamma(h)$   $\Rightarrow$  autocovarianța depinde doar de lag  $h$



## Staționaritatea slabă (covarianță)

### Definiție 3 (Staționaritate Slabă)

- Un proces  $\{X_t\}$  este **slab staționar** (sau staționar în covarianță) dacă:
  - ▶  $\mathbb{E}[X_t^2] < \infty$  pentru toți  $t$ 
    - Momente finite de ordin 2
  - ▶  $\mathbb{E}[X_t] = \mu$  pentru toți  $t$ 
    - Medie constantă
  - ▶  $\text{Cov}(X_t, X_{t+h}) = \gamma(h)$ 
    - Covarianța depinde doar de lag-ul  $h$ , nu de  $t$

### Proprietăți

- **Autocovarianța** este funcție doar de lag:
  - ▶  $\gamma(h) = \text{Cov}(X_t, X_{t+h}) = \mathbb{E}[(X_t - \mu)(X_{t+h} - \mu)]$
- **Autocorelația:**
  - ▶  $\rho(h) = \gamma(h)/\gamma(0) = \text{Cov}(X_t, X_{t+h})/\text{Var}(X_t)$
- **Notă:**  $\rho(0) = 1$ ,  $|\rho(h)| \leq 1$ ,  $\rho(h) = \rho(-h)$  (simetrie)



## Relația între staționaritate strictă și slabă

### Teoremă 1 (Implicație Fundamentală)

- Dacă  $\{X_t\}$  este **strict staționar** și  $\mathbb{E}[X_t^2] < \infty$ , atunci  $\{X_t\}$  este și **slab staționar**

### Demonstrație.

- Fie  $t_1, t_2$  oarecare și  $h$  deplasare temporală arbitrară
- Din invarianța distribuției comune:  $(X_{t_1}, X_{t_2}) \stackrel{d}{=} (X_{t_1+h}, X_{t_2+h})$
- $\mathbb{E}[X_{t_1}] = \mathbb{E}[X_{t_1+h}] = \mu$  (medie constantă)
- $\text{Cov}(X_{t_1}, X_{t_2}) = \text{Cov}(X_{t_1+h}, X_{t_2+h})$
- Deci autocovarianța depinde doar de diferența  $t_2 - t_1 = h$ , nu de  $t_1$

□

### Atenție: Reciproca NU este adevărată!

- Există procese slab staționare dar **nu** strict staționare



## Exemplu: modelul AR(1) este slab staționar

### Modelul AR(1) (AutoRegressive de ordin 1)

- $X_t = \phi X_{t-1} + \varepsilon_t, \quad |\phi| < 1$
- $\varepsilon_t \sim WN(0, \sigma^2)$  — WN (White Noise) = zgomot alb

### Verificarea celor trei condiții

1. **Medie constantă:**  $\mathbb{E}[X_t] = \phi \mathbb{E}[X_{t-1}] + 0 = \phi \mathbb{E}[X_t] \Rightarrow \mathbb{E}[X_t] = 0$
2. **Varianță constantă:**  $\text{Var}(X_t) = \phi^2 \text{Var}(X_{t-1}) + \sigma^2 \Rightarrow \text{Var}(X_t) = \frac{\sigma^2}{1 - \phi^2}$
3. **Autocovarianță depinde doar de lag:**  $\gamma(h) = \phi^{|h|} \cdot \frac{\sigma^2}{1 - \phi^2}, \quad \rho(h) = \phi^{|h|}$

### Exemplu numeric: $\phi = 0.8, \sigma^2 = 1$

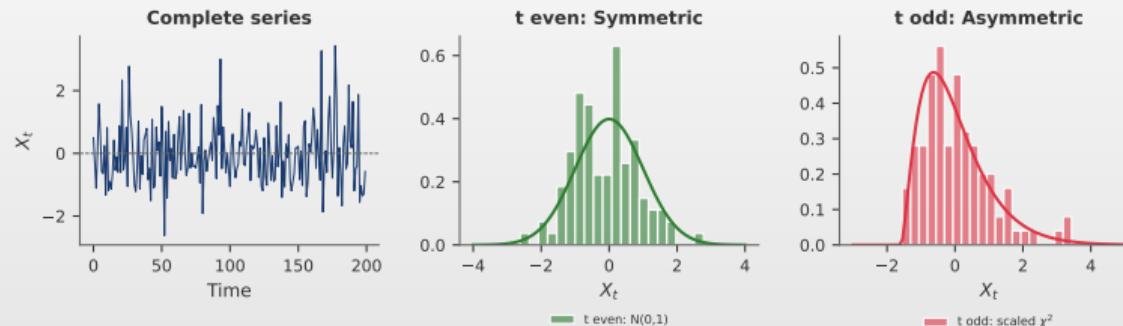
- $\mathbb{E}[X_t] = 0, \quad \text{Var}(X_t) = \frac{1}{1 - 0.64} = 2.78, \quad \rho(1) = 0.8, \quad \rho(2) = 0.64, \quad \rho(5) = 0.33$



## Contraexemplu: slab staționar dar NU strict staționar

### Construcție

- Fie  $\{X_t\}$  variabile aleatoare **independente** cu:  $t$  par:  $X_t \sim N(0, 1)$ ;  $t$  impar:  $X_t \sim \frac{\chi^2(5)-5}{\sqrt{10}}$



### Slab staționar ✓

- $\mathbb{E}[X_t] = 0$ ,  $\text{Var}(X_t) = 1$ ,  $\text{Cov}(X_t, X_{t+h}) = 0$

### NU strict staționar ✗

- Asimetria diferă ( $0$  vs  $> 0$ )  $\Rightarrow X_1 \neq X_2$

Q TSA\_ch1\_stationarity

## Proprietățile funcției de autocovarianță

### Propoziție 1

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

- Simetrie:**  $\gamma(h) = \gamma(-h)$
- Maximum la zero:**  $|\gamma(h)| \leq \gamma(0) = \text{Var}(X_t)$
- Definit nenegativ:**  $\sum_{i,j} a_i a_j \gamma(i-j) \geq 0$  pentru orice  $a_1, \dots, a_n$

### Demonstrație (prop. 3)

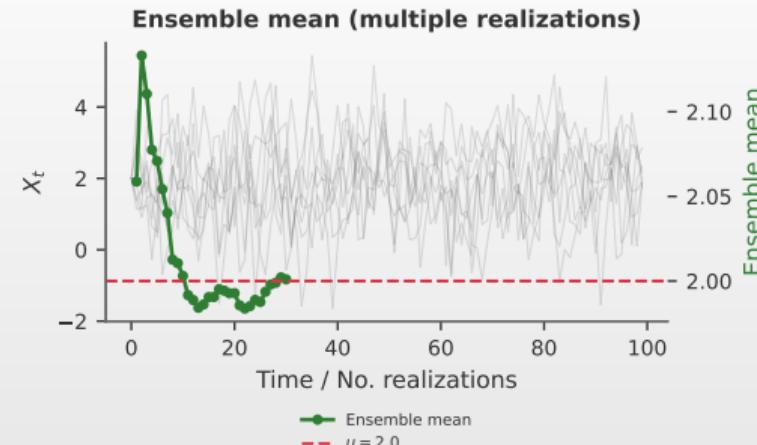
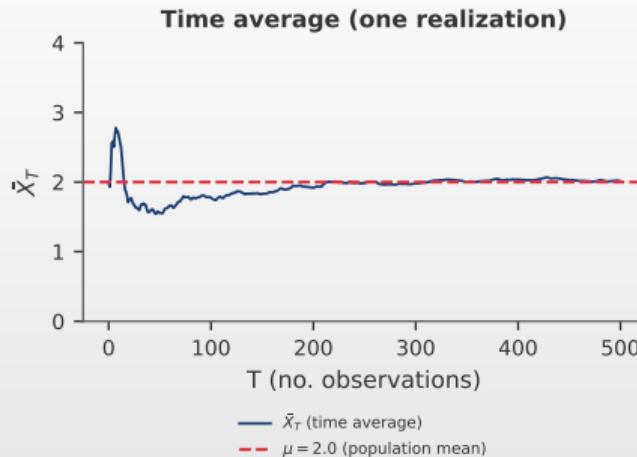
- $\text{Var}(\sum_{i=1}^n a_i X_{t+i}) = \sum_{i,j} a_i a_j \gamma(i-j) \geq 0$  (varianța  $\geq 0$ )

### Implicație

- Nu orice funcție poate fi o funcție de autocovarianță validă



## Ergodicitatea: ilustrare vizuală



- Media temporală** (o singură realizare) și **media ansamblului** (realizări multiple) converg ambele la  $\mu$
- Ergodicitatea garantează că putem estima  $\mu$  dintr-o **singură serie temporală** suficient de lungă



## Ergodicitatea: fundamentul inferenței din date

### Definiție 4 (Ergodicitate pentru Medie)

- Un proces staționar  $\{X_t\}$  este **ergodic pentru medie** dacă:
  - ▶  $\bar{X}_T = \frac{1}{T} \sum_{t=1}^T X_t \xrightarrow{P} \mathbb{E}[X_t] = \mu$  când  $T \rightarrow \infty$

### De ce contează ergodicitatea?

- **Problema:** Avem doar o **singură realizare** a procesului stochastic
- **Soluția:** Ergodicitatea permite estimarea lui  $\mu$  din  $\bar{X}_T$ 
  - ▶ Media temporală convergează la media populației
  - ▶ Fără ergodicitate, inferență statistică nu este posibilă!

### Teoremă 2 (Condiție Suficientă)

- Dacă  $\sum_{h=0}^{\infty} |\gamma(h)| < \infty$  (autocovarianțe absolut sumabile), procesul este ergodic



## Densitatea spectrală: domeniul frecvențelor

### Definiție 5 (Densitatea spectrală de putere)

- Pentru un proces staționar cu  $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$ , **densitatea spectrală** este:

$$S(\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) e^{-i\omega h} = \frac{1}{2\pi} \left[ \gamma(0) + 2 \sum_{h=1}^{\infty} \gamma(h) \cos(\omega h) \right], \quad \omega \in [-\pi, \pi]$$

- $S(\omega)$  descompune varianța procesului pe frecvențe:  $\gamma(0) = \int_{-\pi}^{\pi} S(\omega) d\omega$

### Interpretare

- $S(\omega)$  mare la  $\omega$  mic  $\Rightarrow$  ciclu lung dominant
- Zgomot alb:  $S(\omega) = \frac{\sigma^2}{2\pi}$  (plat)
- AR(1)  $\phi > 0$ : putere concentrată la frecvențe joase
- MA(1)  $\theta > 0$ : putere concentrată la frecvențe înalte

### Conexiuni

- **Perechea Fourier:**  $S(\omega) \leftrightarrow \gamma(h)$  (echivalente)
- Domeniul timp (ACF)  $\equiv$  domeniul frecvență (spectru)
- **Periodograma:** estimator empiric al lui  $S(\omega)$
- Util pentru detectarea sezonalității ascunse



## Teorema de descompunere Wold

### Teoremă 3 (Wold, 1938)

- Orice proces **staționar în covariantă**  $\{X_t\}$  poate fi scris ca:  $X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} + \eta_t$ 
  - ▶  $\varepsilon_t \sim WN(0, \sigma^2) \Rightarrow$  zgomot alb, cu  $\psi_0 = 1, \sum \psi_j^2 < \infty$
  - ▶  $\eta_t \Rightarrow$  componentă deterministă (perfect predictibilă)

### Semnificația teoremei Wold

- **Descompunere:** Orice proces staționar = **MA( $\infty$ )** (Moving Average de ordin infinit) + componentă deterministă
  - ▶ Justifică teoretic modelele MA( $q$ ) și ARMA( $p, q$ )
  - ▶ Coeficienții  $\psi_j$  măsoară impactul șocurilor trecute



## Demonstrația teoremei Wold (schiță)

Schiță de demonstrație.

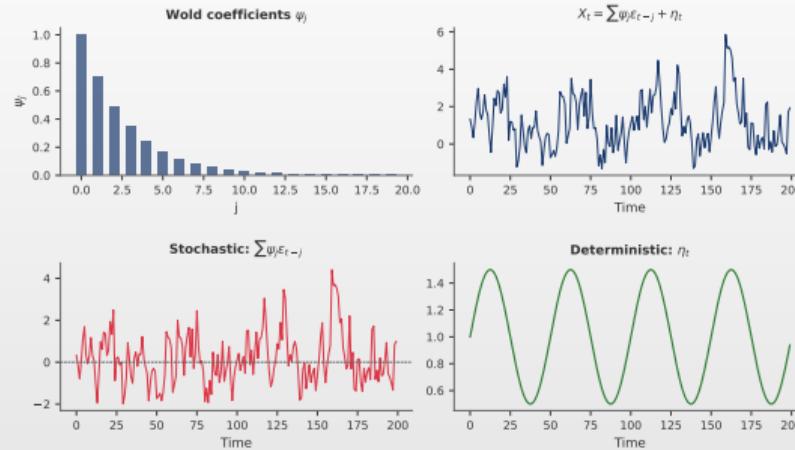
- Spațiul Hilbert al trecutului:** Definim  $\mathcal{H}_t = \overline{\text{sp}}\{X_s : s \leq t\}$  — spațiul închis generat de valorile trecute și prezente, cu produsul scalar  $\langle X, Y \rangle = \text{Cov}(X, Y)$ .
- Inovația:** Definim  $\varepsilon_t = X_t - \hat{X}_t$ , unde  $\hat{X}_t = \text{Proj}_{\mathcal{H}_{t-1}}(X_t)$  este proiecția ortogonală. Prin construcție,  $\varepsilon_t \perp \mathcal{H}_{t-1}$ , deci  $\varepsilon_t \perp \varepsilon_s$  pentru  $t \neq s \Rightarrow \{\varepsilon_t\}$  este zgomot alb.
- Reprezentarea iterativă:** Aplicând recursiv proiecția:

$$X_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \cdots + \eta_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} + \eta_t$$

unde  $\psi_j$  rezultă din proiecțiile succesive, iar  $\eta_t \in \mathcal{H}_{-\infty} = \bigcap_t \mathcal{H}_t$ .

- Convergența:**  $\sum_{j=0}^{\infty} \psi_j^2 < \infty$  deoarece  $\text{Var}(X_t) < \infty$  (staționaritate).
- Componenta deterministă:**  $\eta_t \in \mathcal{H}_{-\infty} \Rightarrow \eta_t$  este *perfect predictibilă* din trecutul infinit. □

## Teorema Wold: ilustrare vizuală

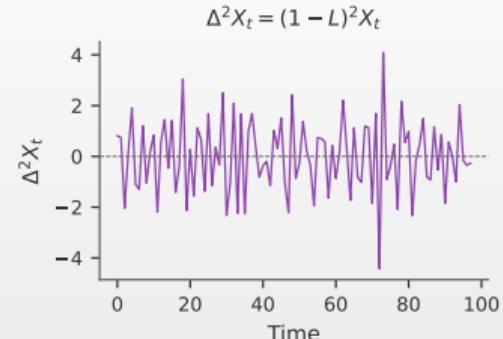
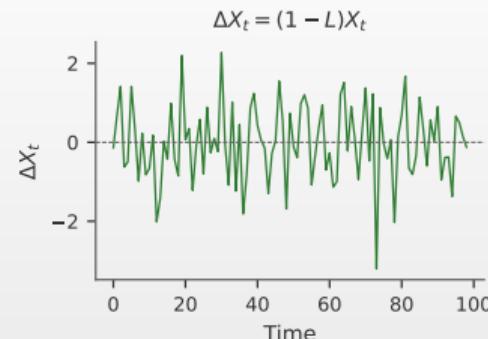
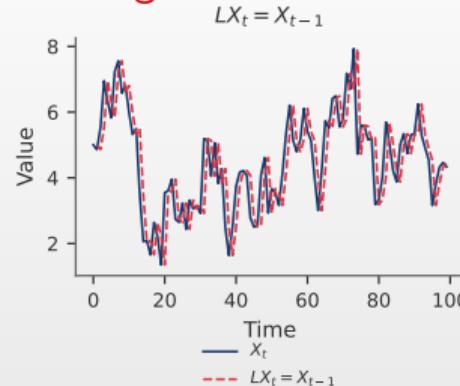


### Interpretare

- $X_t$  se descompune în componentă **stochastică** ( $MA(\infty)$ ) și componentă **deterministă** ( $\eta_t$ )
- Coeficienții  $\psi_j$  descresc  $\Rightarrow$  șocurile recente au impact mai mare decât cele îndepărtate



## Operatorul lag: ilustrare vizuală



### Proprietăți

- $LX_t = X_{t-1} \Rightarrow$  operatorul lag deplasează seria cu o perioadă în trecut
- $L^k X_t = X_{t-k} \Rightarrow$  deplasare cu  $k$  perioade;  $L^0 = I$  (identitate)
- **Operatorul diferență:**  $\Delta = (1 - L)$ , astfel  $\Delta X_t = X_t - X_{t-1}$



## Operatorul lag

### Definiție 6 (Operatorul lag)

- **Operatorul lag** (sau operatorul de întârziere)  $L$  este definit prin:  $LX_t = X_{t-1}$

### Proprietăți

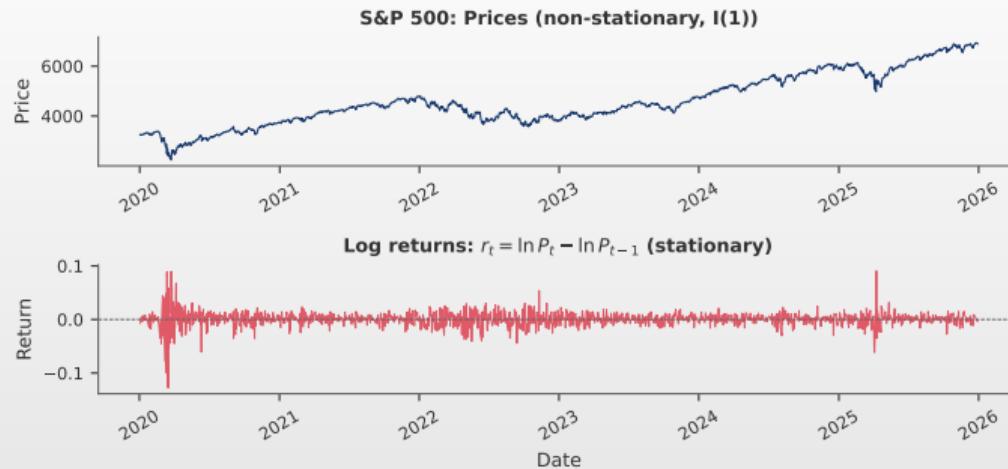
- **Puteri:**  $L^k X_t = X_{t-k}$  (întârzie cu  $k$  perioade)
  - ▶ Notație compactă pentru modele
- **Identitate:**  $L^0 = I$
- **Polinom:**  $(1 - \phi L)X_t = X_t - \phi X_{t-1}$

### Exemple

- **Prima diferență:**  $(1 - L)X_t = X_t - X_{t-1}$
- **A doua diferență:**  $(1 - L)^2 X_t = \Delta^2 X_t$
- **Sezonieră:**  $(1 - L^{12})X_t$



## Efectul diferențierii: S&P 500



### Interpretare

- Sus:** Prețuri S&P 500  $\Rightarrow$  trend clar, nestaționar ( $I(1)$ )
- Jos:** Randamente log  $r_t = \ln P_t - \ln P_{t-1} \Rightarrow$  fluctuează în jurul mediei  $\approx 0$ , staționar



## Diferențierea

### De ce diferențiem?

- Prima Diferență:**  $\Delta X_t = X_t - X_{t-1} = (1 - L)X_t$ 
  - ▶ Elimină trendul și rădăcina unitate
  - ▶ Mers aleator:  $\Delta X_t = \varepsilon_t$

### Definiție 7 (Proces Integrat de Ordin $d$ )

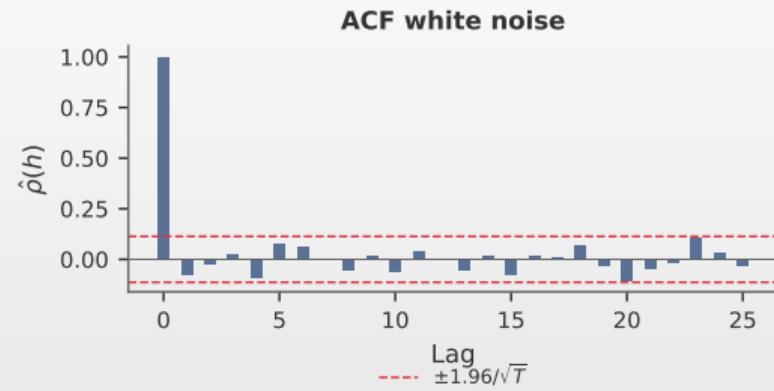
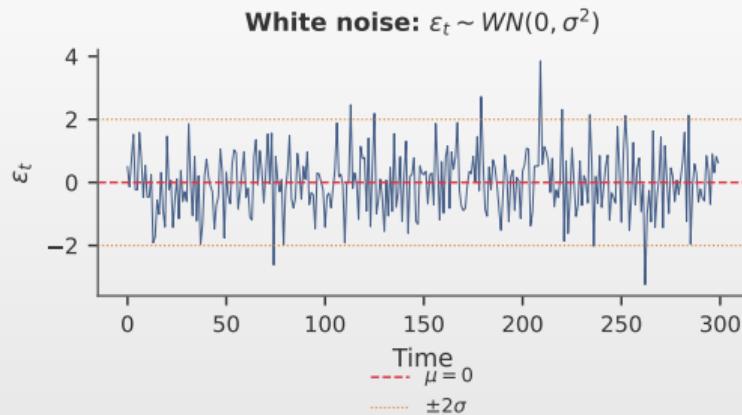
- Un proces  $\{X_t\}$  este **integrat de ordin  $d$** , notat  $X_t \sim I(d)$ , dacă:
  - ▶  $\Delta^d X_t = (1 - L)^d X_t$  este staționar ( $I(0)$  proces)
  - ▶  $\Delta^{d-1} X_t$  nu este staționar

### Exemple

- $I(0)$ : Proces staționar (zgomot alb, AR staționar)
- $I(1)$ : Mers aleator  $\Rightarrow \Delta X_t = \varepsilon_t$  este staționar
- $I(2)$ : Necesită două diferențieri pentru staționaritate



## Zgomot alb: ilustrare vizuală



Q TSA\_ch1\_white\_noise



## Procesul de zgomot alb

### Definiție 8 (Zgomot Alb)

- Un proces  $\{\varepsilon_t\}$  este **zgomot alb**, notat  $\varepsilon_t \sim WN(0, \sigma^2)$ , dacă:
  - ▶  $E[\varepsilon_t] = 0$  pentru orice  $t$  (medie zero)
  - ▶  $Var(\varepsilon_t) = \sigma^2$  pentru orice  $t$  (varianță constantă)
  - ▶  $Cov(\varepsilon_t, \varepsilon_s) = 0$  pentru  $t \neq s$  (necorelat)

### ACF al zgomotului alb

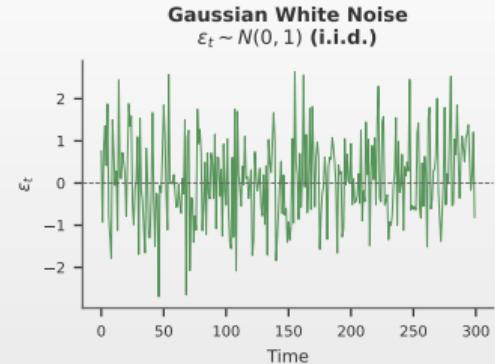
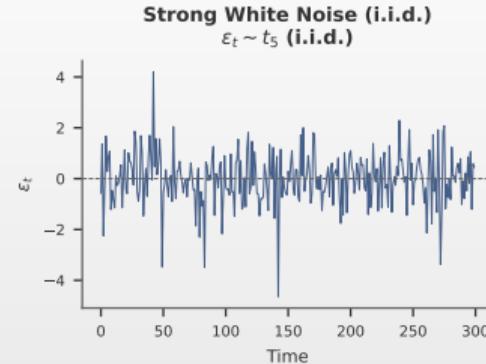
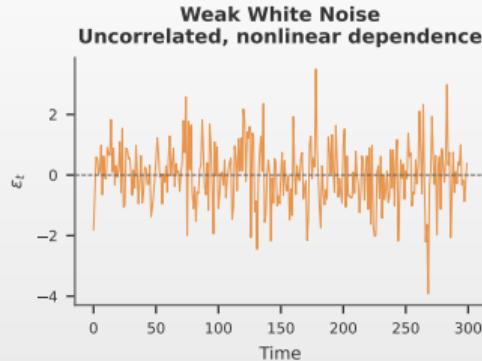
- Din definiție:  $\gamma(0) = \sigma^2$  și  $\gamma(h) = 0$  pentru  $h \neq 0$ ;  $\rho(h) = \begin{cases} 1 & h = 0 \\ 0 & h \neq 0 \end{cases}$

### Observație

- Zgomotul alb este cel mai simplu proces staționar — element fundamental al modelelor ARMA
- Există trei tipuri: slab, puternic (i.i.d.) și Gaussian (vezi diapozitivul următor)



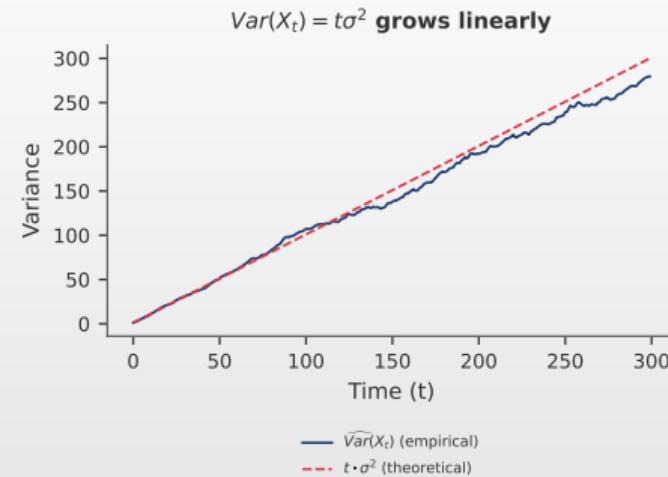
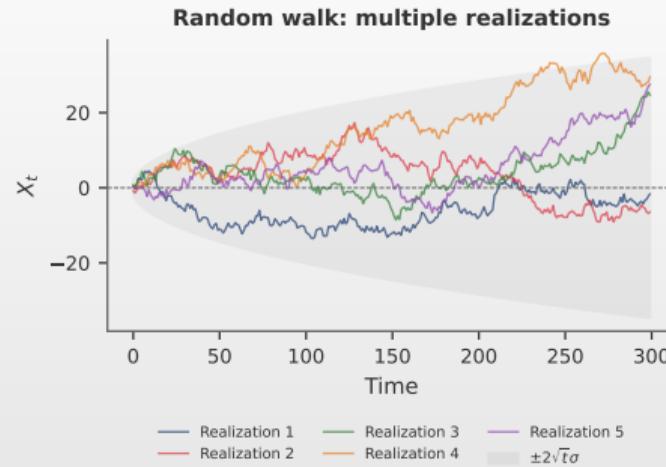
## Cele trei tipuri de zgomot alb



Relația de incluziune: Gaussian  $\subset$  puternic (i.i.d.)  $\subset$  slab (necorelat)

- Slab:**  $\text{Cov}(\varepsilon_t, \varepsilon_s) = 0$ , dar pot exista dependențe neliniare (ex. GARCH)
- Puternic:**  $\varepsilon_t$  sunt i.i.d. (independente și identic distribuite) — orice distribuție (ex. Student- $t$ )
- Gaussian:**  $\varepsilon_t \stackrel{iid}{\sim} N(0, \sigma^2)$  — necorelat  $\Leftrightarrow$  independent

## Mers aleator: vizualizare



### Observații

- Fiecare șoc are **efect permanent**;  $\text{Var}(X_t) = t\sigma^2$  crește liniar cu timpul
- Soluție** — diferențierea transformă în zgomot alb,  $\Delta X_t = \varepsilon_t$



## Procesul de mers aleatoriu

### Definiție 9 (Mers Aleatoriu)

- $X_t = X_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma^2), \quad X_0 = 0$
- Forma explicită:**  $X_t = \sum_{i=1}^t \varepsilon_i$

### Propoziție 2 (Proprietăți)

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

### Demonstrații.

- $\mathbb{E}[X_t] = \mathbb{E}\left[\sum_{i=1}^t \varepsilon_i\right] = 0$
- $\text{Var}(X_t) = \text{Var}\left(\sum_{i=1}^t \varepsilon_i\right) = \sum_{i=1}^t \text{Var}(\varepsilon_i) = t\sigma^2 \quad (\text{independentă})$
- $\text{Cov}(X_t, X_s) = \min(t, s) \sigma^2 \quad (\text{pentru } s \leq t)$

Nestaționar!

Analiza  $\text{Var}(X_t)$  arată că este dependență de  $t \Rightarrow$  mersul aleatoriu **nu este staționar**



## Mers aleator cu drift

### Definiție 10 (Mers Aleatoriu cu Drift)

- $X_t = c + X_{t-1} + \varepsilon_t, \quad c \neq 0$  este **driftul**
- Forma explicită:**  $X_t = ct + \sum_{i=1}^t \varepsilon_i$

### Propoziție 3 (Proprietăți)

- $\mathbb{E}[X_t] = ct$  (trend liniar)
- $\text{Var}(X_t) = t\sigma^2$  (crește cu timpul)

### Diferențiere

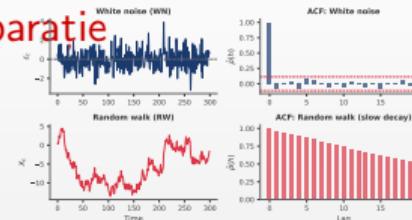
- $\Delta X_t = c + \varepsilon_t$  — constantă plus zgomot alb  $\Rightarrow$  seria diferențiată este staționară

### Importanța practică

- PIB nominal, prețuri de acțiuni  $\Rightarrow$  adesea modele ca RW (Random Walk — mers aleatoriu) cu drift
- Testul ADF include variante: fără constantă, cu constantă, cu constantă și trend



## Zgomot alb vs mers aleatoriu: comparație



### Zgomot alb

- Staționar,  $\text{Var} = \sigma^2$  (const.),  $\text{ACF} = 0$  pentru  $h \neq 0$ , fără memorie

### Mers aleator

- Nestaționar,  $\text{Var} = t\sigma^2$  (crește),  $\text{ACF} \approx 1$  (lent), șocuri permanente

### Legătură

- $\Delta X_t = \varepsilon_t$



## Staționaritate în trend vs. staționaritate în diferențe

### Staționaritate în trend — TS (Trend Stationary)

- Model:**  $Y_t = \alpha + \beta t + \varepsilon_t$ 
  - ▶ Trend **determinist**
  - ▶ Abaterile de la trend sunt temporare
- Soluție:** regresie pe  $t$ , se extrag reziduurile
- Efect:** Șocurile NU au efect permanent

### Staționaritate în diferențe — DS (Difference Stationary)

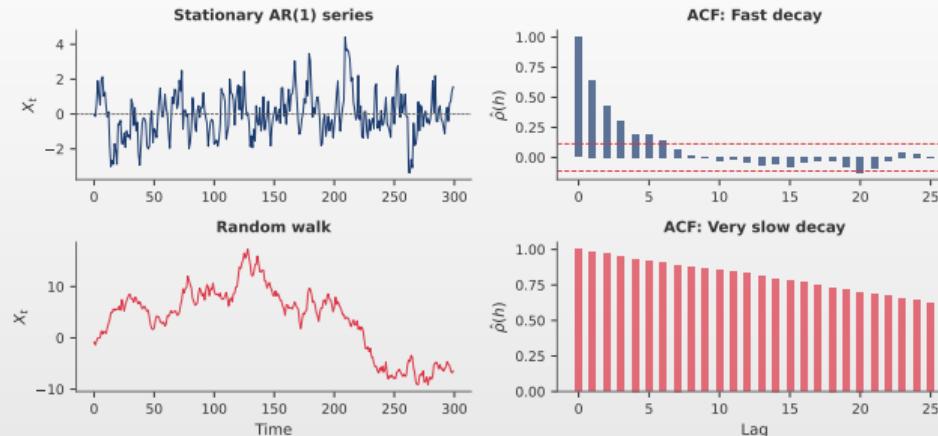
- Model:**  $Y_t = c + Y_{t-1} + \varepsilon_t$ 
  - ▶ Trend **stochastic**
  - ▶ Abaterile de la trend sunt permanente
- Soluție:** diferențiere ( $\Delta Y_t$ )
- Efect:** Șocurile AU efect permanent

### De ce contează distincția?

- Diferențiere pe TS:** introduce rădăcină unitară artificială în MA
- Regresie pe DS:** produce reziduuri tot nestaționare
- Soluție:** Testele ADF și KPSS ajută la distincție



## Comparație ACF: staționar vs mers aleatoriu

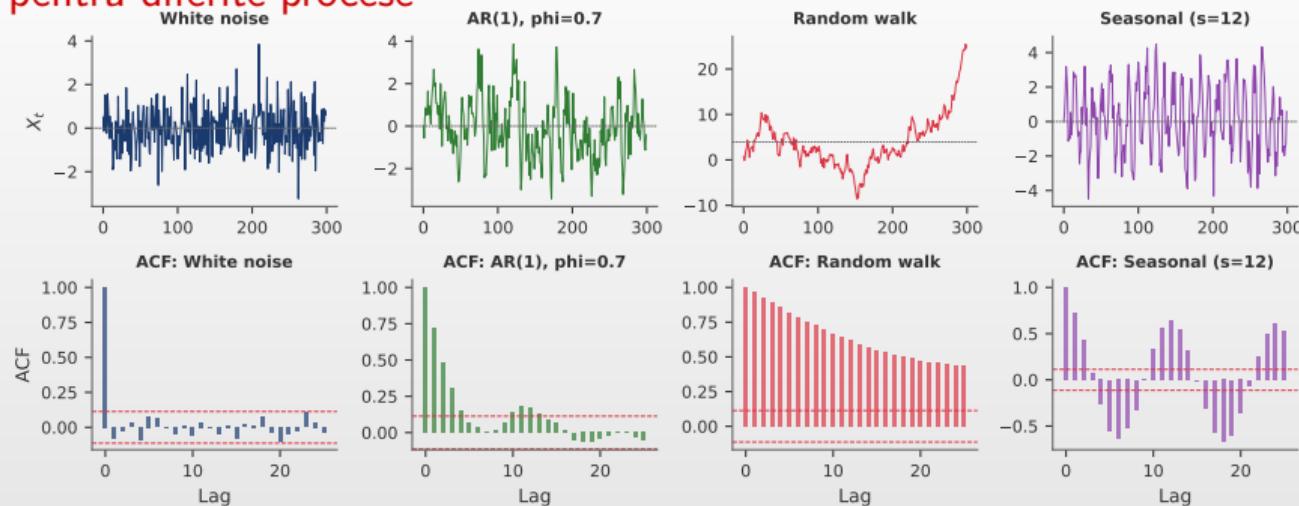


### Interpretare

- Staționar:** ACF scade rapid (exponențial sau oscilant) spre zero
- Mers aleator:** ACF scade foarte lent, rămâne aproape de 1
- Regulă practică:** ACF lent  $\Rightarrow$  suspectăm rădăcină unitate  $\Rightarrow$  test ADF



## Tipare ACF pentru diferite procese



### Interpretare

- Zgomot alb:**  $ACF = 0$ ; **Staționar:** scade rapid; **Nestăționar:** scade lent
- Sezonier:** Vârfuri la lag-uri sezonale (12, 24 pentru date lunare)



## Funcția de autocorelație eșantion

### ACF 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:  $\hat{\rho}(0) = 1$ ,  $|\hat{\rho}(h)| \leq 1$

### Teoremă 4 (Bartlett, 1946)

- Sub  $H_0$ : zgomot alb, pentru  $T$  mare:  $\hat{\rho}(h) \approx N(0, 1/T)$

### Interval de încredere 95%

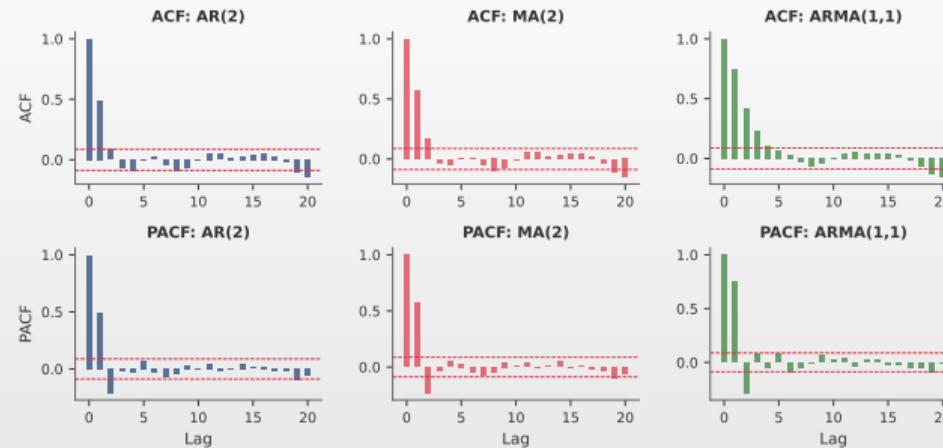
- $\pm 1.96/\sqrt{T}$  (benzile din graficele ACF)

### Atenție

- Formula Bartlett validă doar sub  $H_0$ : zgomot alb
- Pentru AR/MA, varianța asimptotică diferă



## Tipare ACF și PACF



### Reguli de identificare

- AR( $p$ )**: ACF scade exponențial, PACF se anulează după lag  $p$
- MA( $q$ )**: ACF se anulează după lag  $q$ , PACF scade exponențial
- ARMA( $p, q$ )**: Ambele scad exponențial  $\Rightarrow$  identificarea necesită criterii informaționale



## Funcția de autocorelație parțială (PACF)

### Definiție 11 (Autocorelația Parțială)

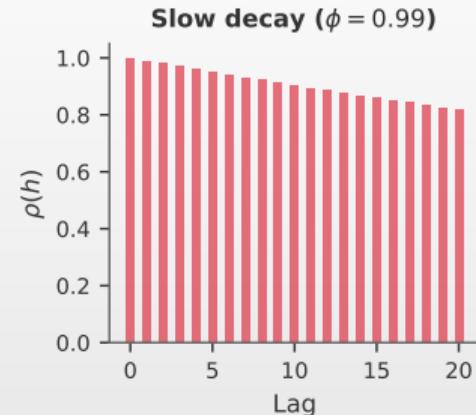
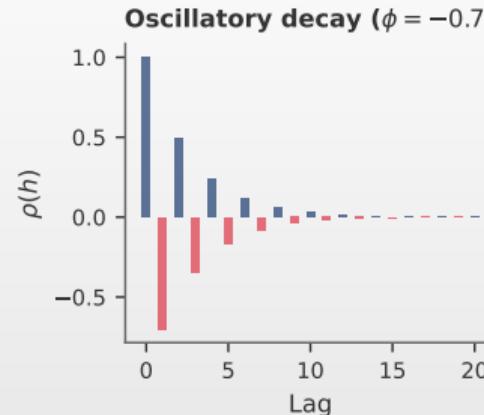
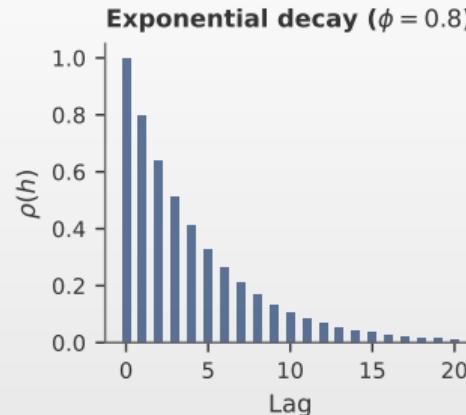
- **PACF** la lag-ul  $h$ , notat  $\phi_{hh}$ : ultimul coeficient din regresia:
  - ▶  $X_t = \phi_{h1}X_{t-1} + \phi_{h2}X_{t-2} + \cdots + \phi_{hh}X_{t-h} + e_t$
- **Alternativ:**
  - ▶  $\phi_{hh} = \text{Corr}(X_t - \hat{X}_t^{(h-1)}, X_{t-h} - \hat{X}_{t-h}^{(h-1)})$
- **Interpretare:** Dependență *directă* la lag-ul  $h$ 
  - ▶ Elimină efectul lag-urilor intermedieare

### Aplicație practică

- PACF identifică direct ordinul  $p$  al unui model AR( $p$ )
- $\phi_{hh} = 0$  pentru  $h > p \Rightarrow$  PACF se anulează după lag-ul  $p$



## Tipare de scădere ACF



### Interpretare

- Scădere exponențială:** Dependență pozitivă persistentă (AR cu  $\phi > 0$ )
- Scădere oscilantă:** Dependență alternantă (AR cu  $\phi < 0$ )
- Viteza de scădere indică puterea memoriei procesului



## Testul Augmented Dickey-Fuller (ADF)

### Modelul ADF

$\Delta X_t = \alpha + \gamma X_{t-1} + \sum_{i=1}^p \delta_i \Delta X_{t-i} + \varepsilon_t, \quad \gamma = \rho - 1, \quad H_0 : \gamma = 0 \Leftrightarrow \rho = 1$

### Ipoteze

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

### Statistica de Test

- $\tau_{ADF} = \hat{\gamma} / SE(\hat{\gamma})$
- $\hat{\gamma}$  = coeficient OLS (Ordinary Least Squares) al  $X_{t-1}$
- $SE(\hat{\gamma})$  din regresia OLS

### Regula de decizie

- $\tau_{ADF} < \text{val. critică} \Rightarrow \text{Respingem } H_0 \Rightarrow \text{Staționar}$
- $\tau_{ADF} \geq \text{val. critică} \Rightarrow \text{Nestaționar (rădăcină unitate)}$
- Valorile critice urmează distribuția Dickey-Fuller (**nu t-Student!**)



## Testul KPSS

### Modelul

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

### Ipoteze (opus ADF)

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

### Statistica de Test

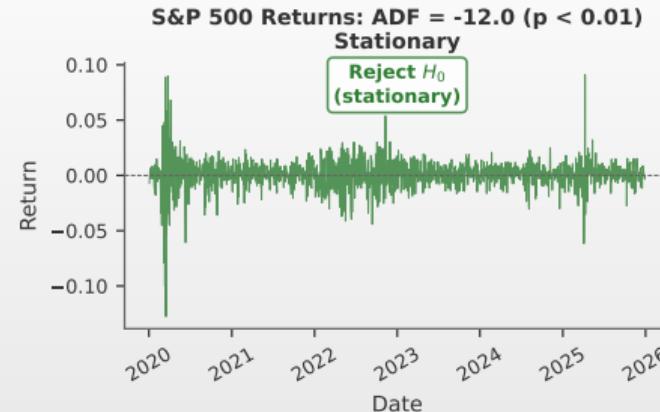
- $LM = \frac{\sum_{t=1}^T S_t^2}{T^2 \hat{\sigma}_{LR}^2}$
- $S_t = \sum_{i=1}^t \hat{e}_i, \quad \hat{\sigma}_{LR}^2 = \text{varianța de lungă durată}$

### Regula de decizie

- $LM >$  valoarea critică  $\Rightarrow$  Respingerem  $H_0 \Rightarrow$  **Nestăționar**
- $LM \leq$  valoarea critică  $\Rightarrow$  **Staționar**



## Testul ADF: vizualizare cu S&P 500



Q TSA\_ch1\_unit\_root\_tests

### Interpretarea testului ADF

- **Ipoteza:**  $H_0$ : Rădăcină unitate
  - ▶ Valori critice:  $-3.43$  (1%),  $-2.86$  (5%),  $-2.57$  (10%)
  - ▶  $\tau < \text{val. critică} \Rightarrow$  respingem  $H_0 \Rightarrow$  serie staționară
- **S&P 500:** Prețuri nestaționare; Randamente staționare



## Folosirea ADF și KPSS împreună

### Testare confirmatorie

- ADF respinge  $H_0$  + KPSS nu respinge: Staționar
- ADF nu respinge + KPSS respinge  $H_0$ : Rădăcină Unitară
- Ambele resping sau ambele nu resping: Neconcludent
  - ▶ Necesită teste suplimentare: PP (Phillips-Perron)
  - ▶ Sau DF-GLS (Dickey-Fuller Generalized Least Squares)

### Flux de lucru

- Pasul 1: Test ADF ( $H_0$ : rădăcină)
- Pasul 2: Test KPSS ( $H_0$ : staționar)
- Pasul 3: Rezultate concordante  $\Rightarrow$  OK
  - ▶ Altfel: teste PP, DF-GLS



## Testul Phillips-Perron (PP)

### Definiție 12 (Phillips-Perron, 1988)

- Testează aceeași ipoteză ca ADF:  $H_0$ : rădăcină unitate ( $\gamma = 0$ )
- Modelul de bază (fără lag-uri augmentate):  $\Delta y_t = \alpha + \gamma y_{t-1} + e_t$
- Corectează autocorelația și heteroscedasticitatea din  $e_t$  prin **corecție neparametrică** (Newey-West) a statisticii  $t$

### Statistica PP

- $Z_t = t_\gamma \cdot \sqrt{\frac{s_e^2}{\hat{\lambda}^2}} - \frac{T(\hat{\lambda}^2 - s_e^2)}{2\hat{\lambda}^2 \cdot SE(\hat{\gamma})}$
- $\hat{\lambda}^2$ : varianța de lungă durată (kernel Newey-West)
- $s_e^2$ : varianța reziduală OLS
- Valorile critice: identice cu ADF (distribuția Dickey-Fuller)

### PP vs ADF

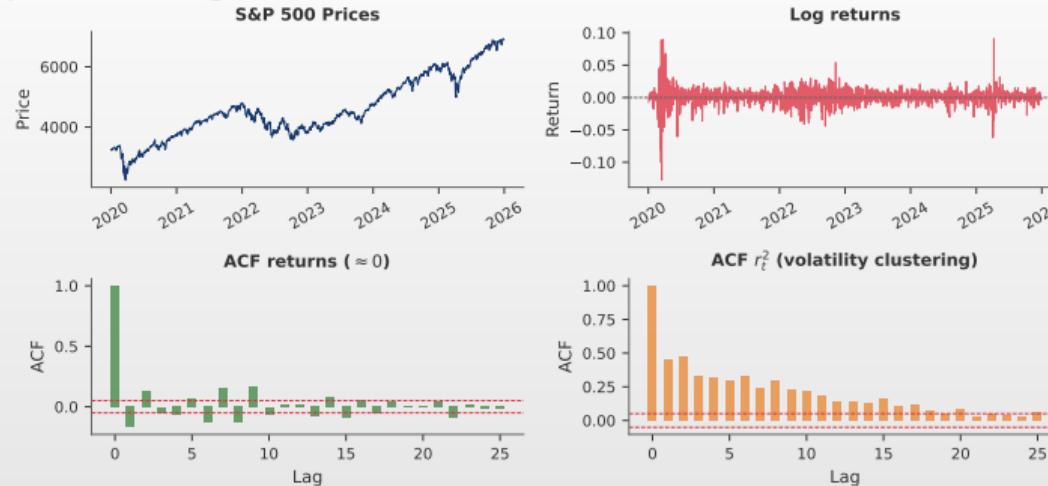
- **ADF**: adaugă lag-uri  $\Delta y_{t-j} \Rightarrow$  parametric
- **PP**: corectează statistica  $t \Rightarrow$  neparametric
- PP mai robust la heteroscedasticitate
- ADF mai robust la MA cu rădăcini aproape de  $-1$

### Python

```
from statsmodels.tsa.stattools import PhillipsPerron
```



## Analiza S&P 500: prezentare generală

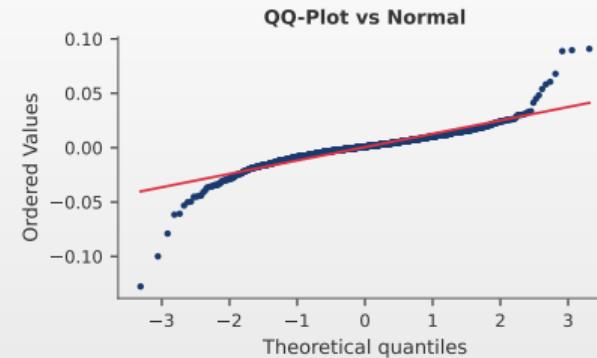
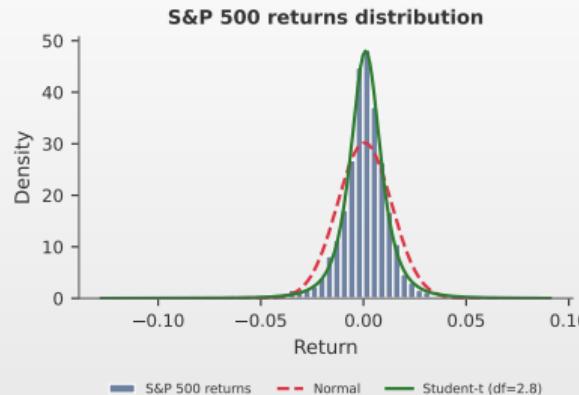


### Observații

- **Prețuri:** Trend ascendent, nestaționar; **Randamente:** Medie  $\approx 0$ , staționar
- **ACF randamente:**  $\approx 0$  (eficient); **ACF  $r_t^2$ :** Semnificativ (aglomerarea volatilității)



## Fapte stilizate ale randamentelor financiare



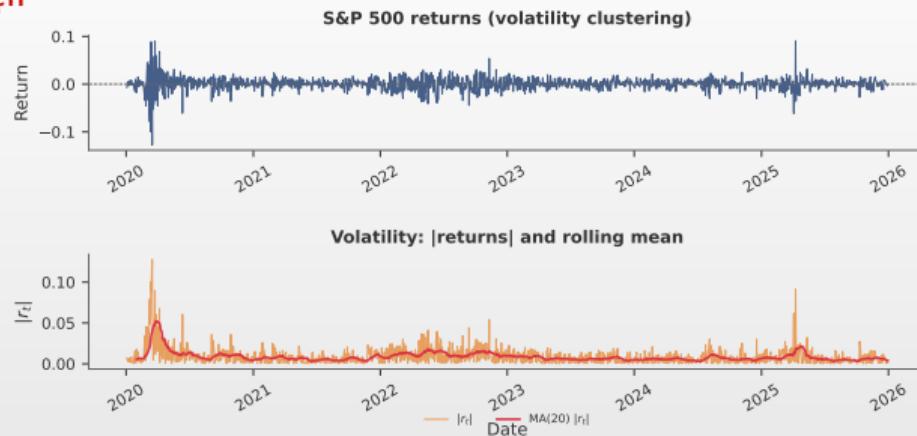
### Proprietăți observate

- Asimetrie negativă (coadă stângă)
- Kurtosis excesiv ( $\gg 3$ )
- Cozi groase (heavy tails)

### Implicații

- Distribuția normală inadecvată
- Evenimente extreme mai probabile
- Necesită distribuții cu cozi groase
- Ex.: Student- $t$ , GED (Generalized Error Distribution)

## Aglomerarea volatilității

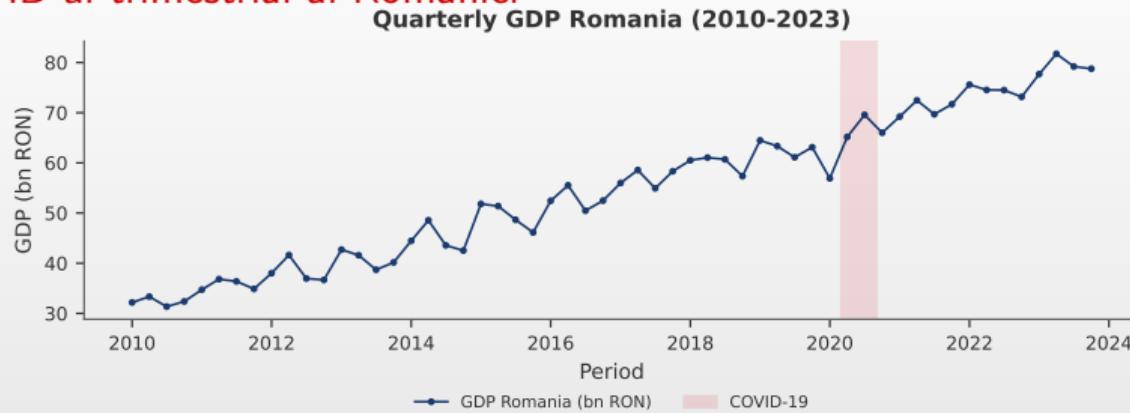


### Observații

- Randamente mari (în valoare absolută) urmate de randamente mari
- Perioade de calm urmate de perioade de volatilitate ridicată
- **Volatilitate variabilă în timp** ⇒ necesită modele specializate (Cap. 5)
- Modele ARCH (AutoRegressive Conditional Heteroskedasticity) / GARCH (Generalized ARCH)



## Studiu de caz: PIB-ul trimestrial al României



TSA\_ch1\_case\_gdp

### Analiza inițială

- **Date:** PIB trimestrial România 2010–2023 (56 obs., INS — Institutul Național de Statistică / Eurostat)
- **Observații:** Trend ascendent, posibil sezonier
  - ▶ řoc structural COVID-19 vizibil
- **Ipoteză:** Serie nestaționară ⇒ testăm cu ADF și KPSS



## Testarea staționarității: ADF și KPSS

### Testul ADF

- Ipoteză:**  $H_0$ : Rădăcină unitate
- Rezultat:** Stat. ADF: -1.23
  - ▶ Val. critică: -2.89
  - ▶ Nu respingem  $H_0$

### Testul KPSS

- Ipoteză:**  $H_0$ : Staționară
- Rezultat:** Stat. KPSS: 0.89
  - ▶ Val. critică: 0.46
  - ▶ Respingem  $H_0$

Concluzie: Ambele teste concordă

- Seria PIB este **nestaționară**  $\Rightarrow$  necesită diferențiere



## Diferențierea: obținerea staționarității

### După diferențiere

- Teste:** Ambele confirmă staționaritate
  - ▶ ADF:  $-4.56$  ( $p < 0.01$ )
  - ▶ KPSS:  $0.21$  ( $p > 0.10$ )

### Concluzie

- PIB nivel:** nestaționar
- $\Delta PIB$ :** staționar
  - ▶ Folosim  $\Delta PIB_t$  pentru modelare

### Rezultat final

- PIB-ul necesită o diferențiere pentru a deveni staționar



## Exercițiu AI: Gândire critică

Prompt de testat în ChatGPT / Claude / Copilot

"Folosind yfinance, descarcă cursul zilnic EUR/RON (EURRON=X) din 2020-01-01 până în 2024-12-31 (aprox. 1.250 observații). Testează dacă seria este staționară folosind testele ADF și KPSS. Ajustează un model adecvat și prognozează cursul pentru următoarele 5 zile lucrătoare. Evaluează fiabilitatea progronei."

**Exercițiu:**

1. Rulați prompt-ul într-un LLM (Large Language Model) la alegere și analizați critic răspunsul.
2. Descărcați date reale EUR/RON și reproduceți analiza. Rezultatele coincid?
3. Testul ADF e specificat corect (trend, lag-uri)? Ce se schimbă dacă modificați opțiunile?
4. Comparați progrona modelului AI cu un benchmark naiv ( $\hat{X}_{t+1} = X_t$ ).
5. Dacă seria e un mers aleatoriu, are sens să ajustăm un model ARMA?

**Atenție:** Un RMSE (Root Mean Squared Error) mic și coeficienți semnificativi *nu garantează* o progrona utilă.



## Concluzii principale

### Rezumat

- **Proces stochastic:** colecție de variabile aleatoare indexate în timp
- **Stationaritate slabă:** medie, variantă, autocovariantă constante
- **Zgomot alb:**  $\varepsilon_t \sim WN(0, \sigma^2)$ 
  - ▶ Staționar,  $ACF = 0$  pentru  $h \neq 0$
- **Mers aleator:**  $X_t = X_{t-1} + \varepsilon_t$ 
  - ▶ Nestaționar,  $Var(X_t) = t\sigma^2$
- **ACF/PACF:** instrumente esențiale pentru identificarea structurii
- **Diferențierea:** transformă serii nestaționare în staționare
- **Teste rădăcină unitate:**
  - ▶ ADF ( $H_0$ : rădăcină unitate) vs KPSS ( $H_0$ : staționar)



## Formule importante

### Stationaritate slabă

- **Momente constante:**
  - ▶  $\mathbb{E}[X_t] = \mu$  (medie constantă)
  - ▶  $\text{Var}(X_t) = \sigma^2$  (varianță constantă)
- **Autocovarianță:**  $\gamma(h) = \text{Cov}(X_t, X_{t+h})$
- **Autocorelație:**  $\rho(h) = \gamma(h)/\gamma(0)$

### Operatorul lag

- **Lag:**  $LX_t = X_{t-1}$
- **Diferență:**  $\Delta X_t = (1 - L)X_t$

### Zgomot alb (WN)

- **Model:**  $\varepsilon_t \sim WN(0, \sigma^2)$
- **ACF:**  $\rho(h) = 0$  pentru  $h \neq 0$

### Mers aleator (RW)

- **Model:**  $X_t = X_{t-1} + \varepsilon_t$
- **Varianță:**  $\text{Var}(X_t) = t\sigma^2$  (crește!)



## Ce urmează

### Capitolul 2: Modele ARMA

- AR( $p$ ):** Modele Autoregresive
- MA( $q$ ):** Modele Medie Mobilă
- ARMA( $p, q$ ):** Modele combinate
- Identificare:** Cu ACF/PACF

### Ce vom învăța

- Estimare:** Parametrii modelului
- Diagnostic:** Verificarea modelului
- Prognoză:** Intervale de încredere
- Selectie:** AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion)



## Întrebarea 1

### Întrebare

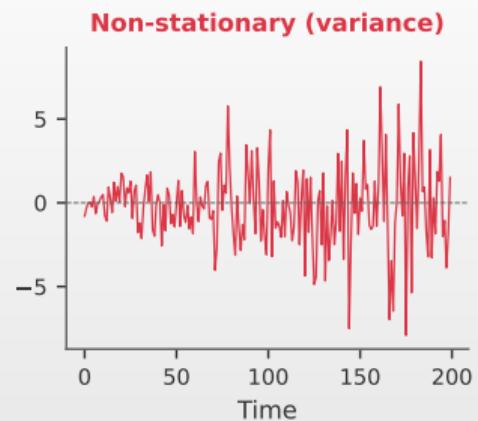
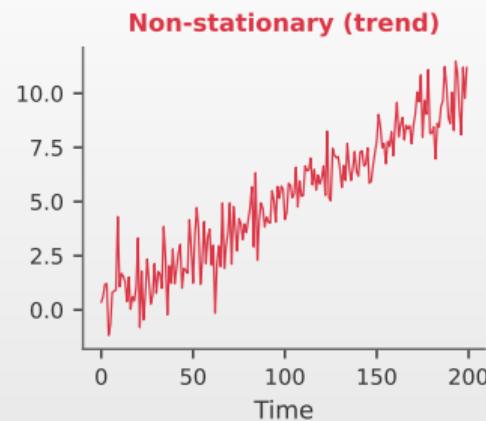
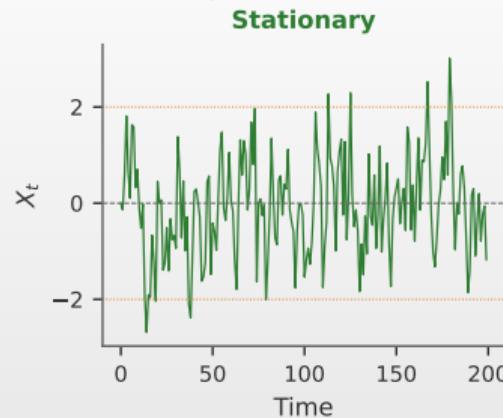
- Care sunt cele trei condiții pentru staționaritatea slabă (în covarianță)?

### Variante de răspuns

- (A)** Media zero, varianța infinită, covarianță dependentă de timp
- (B)** Media constantă, varianța constantă, autocovarianța depinde doar de lag
- (C)** Distribuție normală, independentă, varianță unitară
- (D)** Trend liniar, sezonalitate constantă, reziduuri albe



## Întrebarea 1: Răspuns



Răspuns: (B)

- $\mathbb{E}[X_t] = \mu, \text{Var}(X_t) = \sigma^2, \gamma(t, s) = \gamma(|t - s|)$

Q TSA\_ch1\_stationarity



## Întrebarea 2

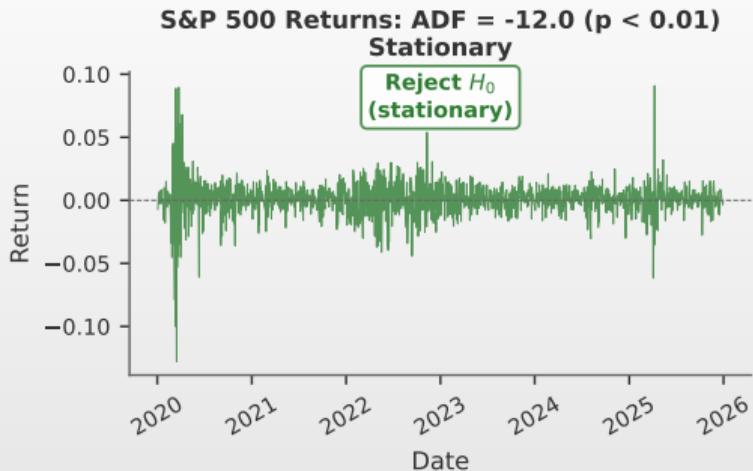
### Întrebare

- Care este ipoteza nulă ( $H_0$ ) în testul ADF (Augmented Dickey-Fuller)?

### Variante de răspuns

- (A)** Seria este staționară
- (B)** Seria are rădăcină unitate (este nestaționară)
- (C)** Seria nu are autocorelație
- (D)** Seria are distribuție normală

## Întrebarea 2: Răspuns



Răspuns: (B)

- $H_0$ : rădăcină unitate;  $\tau <$  val. critică  $\Rightarrow$  staționară



## Întrebarea 3

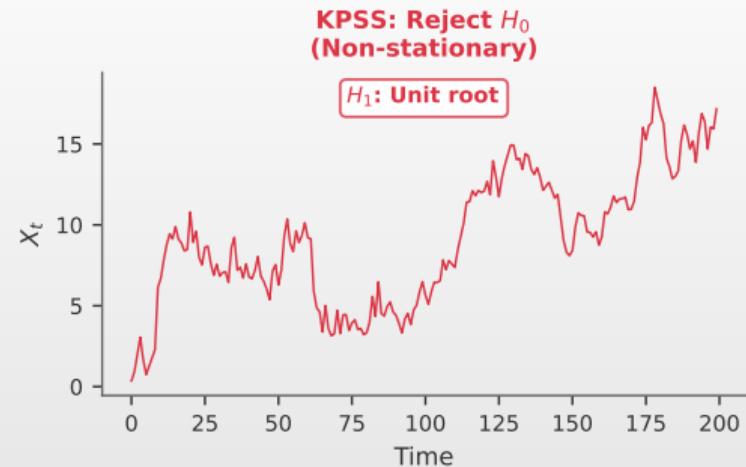
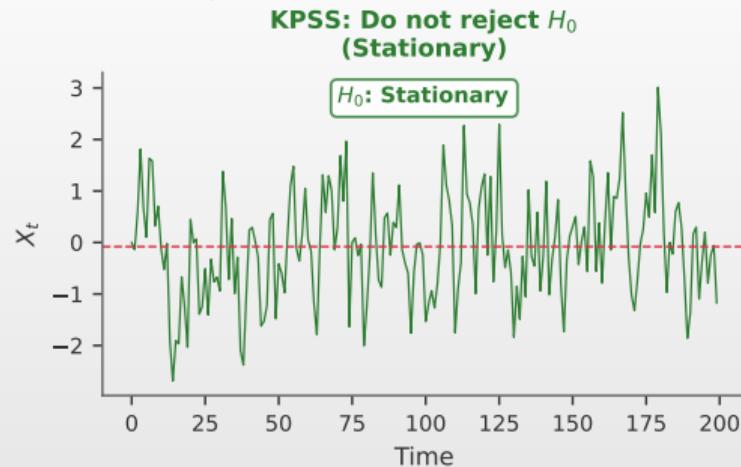
### Întrebare

- Care este ipoteza nulă ( $H_0$ ) în testul KPSS?

### Variante de răspuns

- (A)** Seria are rădăcină unitate (nestaționară)
- (B)** Seria este staționară
- (C)** Seria este un mers aleatoriu
- (D)** Seria are trend determinist

### Întrebarea 3: Răspuns



Răspuns: (B)

- KPSS:  $H_0$  staționară (opus ADF). Folosiți ambele teste!



## Întrebarea 4

### Întrebare

Care este proprietatea fundamentală a varianței unui mers aleatoriu  $X_t = X_{t-1} + \varepsilon_t$ ?

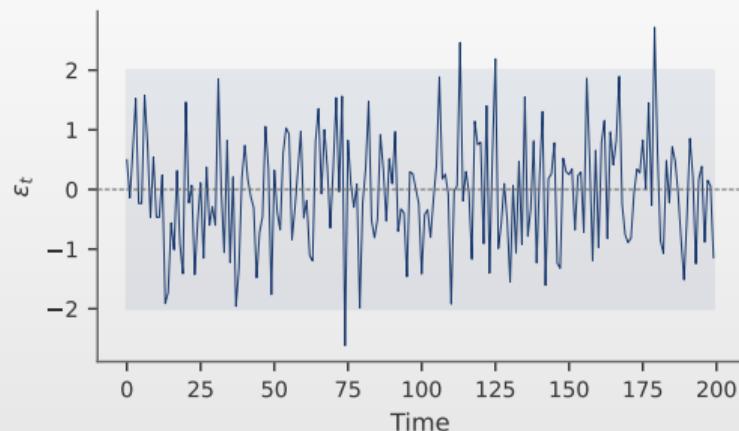
### Variante de răspuns

- (A) Varianța este constantă:  $\text{Var}(X_t) = \sigma^2$
- (B) Varianța crește liniar cu timpul:  $\text{Var}(X_t) = t\sigma^2$
- (C) Varianța scade cu timpul
- (D) Varianța este zero

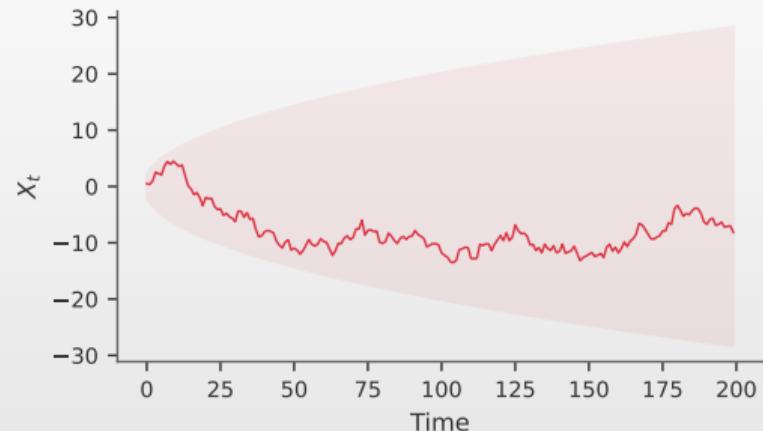


## Întrebarea 4: Răspuns

White noise:  $\text{Var} = \sigma^2$  (const.)



Random walk:  $\text{Var} = t\sigma^2$  (grows!)



Răspuns: (B)

- $\text{Var}(X_t) = t\sigma^2$  crește liniar  $\Rightarrow$  nestaționar

Q TSA\_ch1\_random\_walk



## Întrebarea 5

### Întrebare

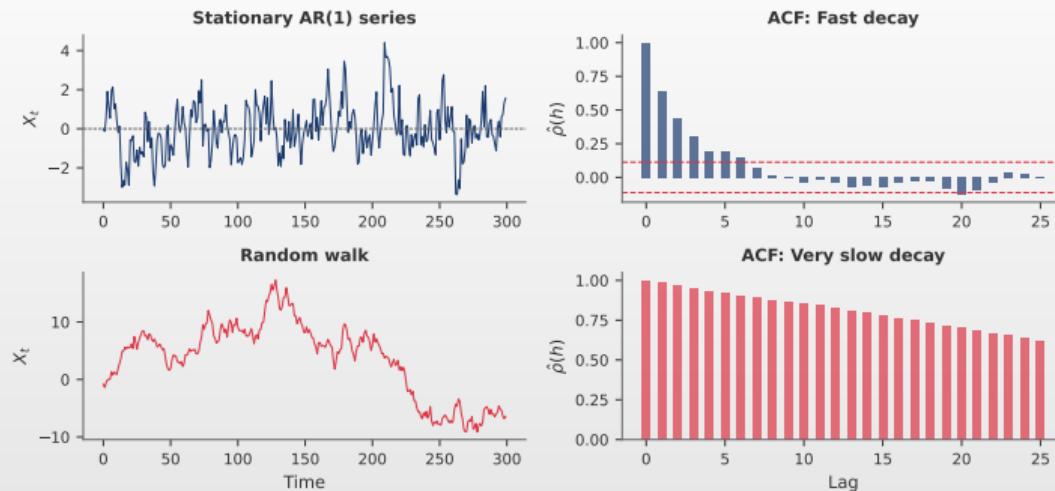
- Cum arată ACF-ul unui mers aleatoriu (serie nestaționară cu rădăcină unitate)?

### Variante de răspuns

- (A)** Toate valorile sunt zero după lag 0
- (B)** Scade exponențial rapid
- (C)** Scade foarte lent (persistență ridicată)
- (D)** Oscilează între pozitiv și negativ



## Întrebarea 5: Răspuns



Răspuns: (C)

- ACF  $\approx 1$  pentru multe lag-uri, scădere lentă  $\Rightarrow$  test ADF



## Întrebarea 6

### Întrebare

Cum obținem randamente staționare dintr-o serie de prețuri financiare  $P_t$ ?

### Variante de răspuns

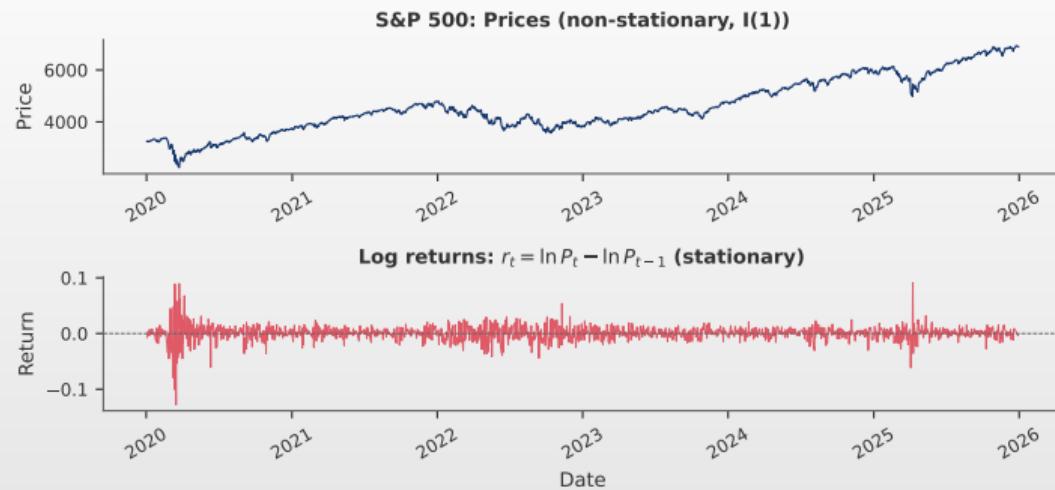
(A) Diferențiere simplă:  $\Delta P_t = P_t - P_{t-1}$

(B) Logaritmare apoi diferențiere:  $r_t = \ln P_t - \ln P_{t-1}$

(C) Doar logaritmare:  $\ln P_t$

(D) Standardizare:  $(P_t - \bar{P})/s_P$

## Întrebarea 6: Răspuns



Răspuns: (B)

- Randamente log:  $r_t = \ln P_t - \ln P_{t-1}$
- Mai întâi  $\ln$  (stabilizează varianța), apoi  $\Delta$  (elimină trendul)  $\Rightarrow$  serie staționară



## Bibliografie I

### Manuale fundamentale

- Hamilton, J.D. (1994). *Time Series Analysis*, Princeton University Press.
- 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.

### Referințe clasice

- Wold, H. (1938). *A Study in the Analysis of Stationary Time Series*, Almqvist & Wiksell.
- Bartlett, M.S. (1946). On the Theoretical Specification and Sampling Properties of Autocorrelated Time-Series, *JRSS Supplement*, 8(1), 27–41.
- Box, G.E.P., & Jenkins, G.M. (1970). *Time Series Analysis: Forecasting and Control*, Holden-Day.



## Bibliografie II

### Teste de staționaritate

- Dickey, D.A., & Fuller, W.A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *JASA*, 74(366), 427–431.
- Kwiatkowski, D., et al. (1992). Testing the Null Hypothesis of Stationarity, *Journal of Econometrics*, 54(1–3), 159–178.

### Resurse online și cod

- **Quantlet:** <https://quantlet.com> – Platformă de cod pentru metode cantitative
- **Quantinar:** <https://quantinar.com> – Platformă de învățare pentru 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 capitol



# Vă Mulțumim!

## Întrebări?

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

