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

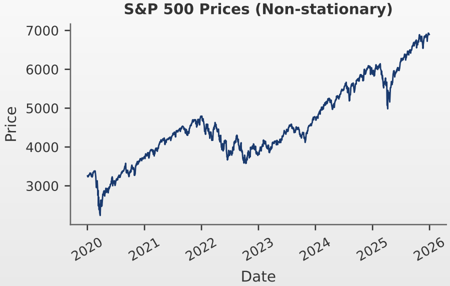
Obiective de învățare

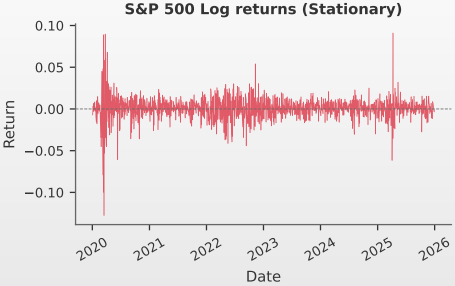
La finalul acestui capitol, veți fi capabili să:

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 și PACF
5. **Aplicați** operatorul lag și diferențierea
6. **Efectuați** teste de staționaritate (ADF, KPSS)
7. **Analizați** date financiare de tip serie de timp
8. **Distingeți** între procesele cu rădăcină unitate și cele staționare în trend

Cuprins

Exemple: serii staționare vs. nestaționare



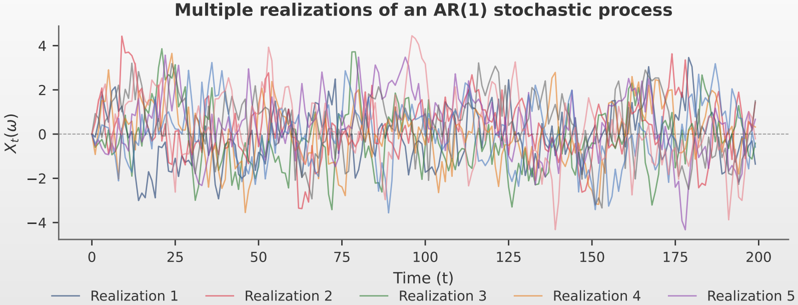


Observații

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

TSA_ch1_stationarity

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

TSA_ch1_stationarity

Proces stochastic: definiție

Definiție 1 (Proces Stochastic)

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

Două perspective

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

Observație cheie

- 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**: $\mu_t = \mathbb{E}[X_t]$
- **Autocovarianța (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)

Punct cheie

- **General**: μ_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

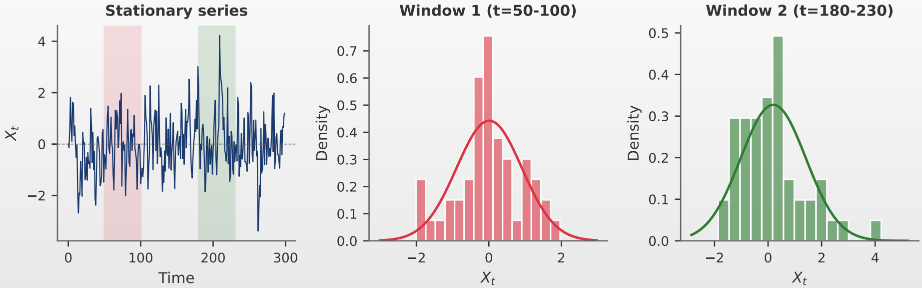
Cu staționaritate

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

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)

TSA_ch1_stationarity

Staționaritatea strictă

Definiție 2 (Staționaritate Strictă (Puternică))

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

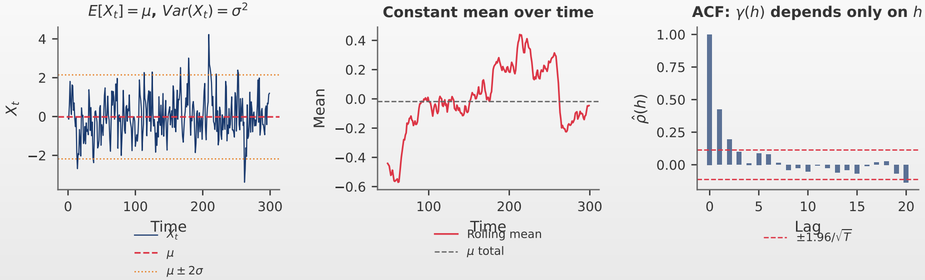
Implicații

- **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)
- **Dependența 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

- $\mathbb{E}[X_t] = \mu$ 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

TSA_ch1_stationarity

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

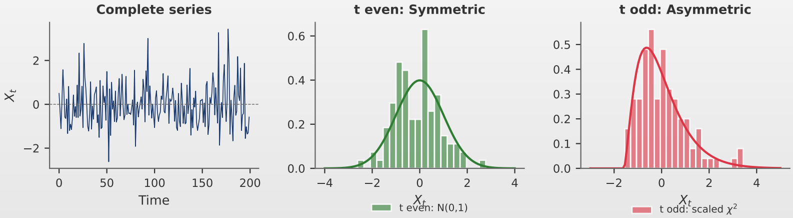
Proprietăți cheie

- **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)

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 $(0 \text{ vs } > 0) \rightarrow X_1 \overset{d}{\neq} X_2$

TSA_ch1_stationarity

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

$\square$

Atenție: Reciproca NU este adevărată!

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

Proprietățile funcției de autocovarianță

Propoziție 1

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

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

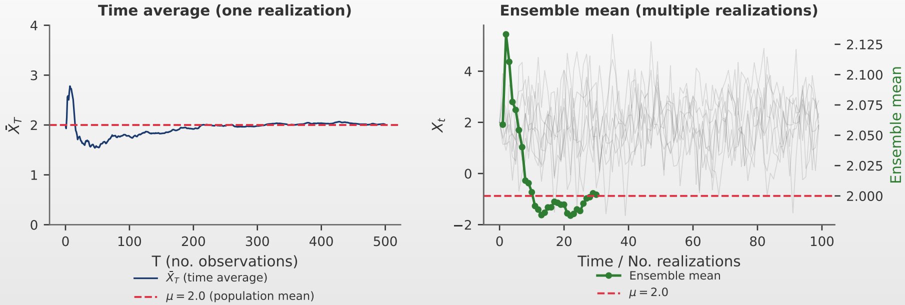
Demonstrație (prop. 3)

- $\mathrm{Var}\left(\sum_{i=1}^{n} a_i X_{t+i}\right) = \sum_{i,j} a_i a_j \gamma(i-j) \geq 0 \quad$ (varianța ≥ 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 μ
- Ergodicitatea garantează că putem estima μ dintr-o **singură serie temporală** suficient de lungă

TSA_ch1_stationarity

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 $T \to \infty$

De ce contează ergodicitatea?

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

Teoremă 2 (Condiție Suficientă)

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

Teorema de descompunere Wold

Teoremă 3 (Wold, 1938)

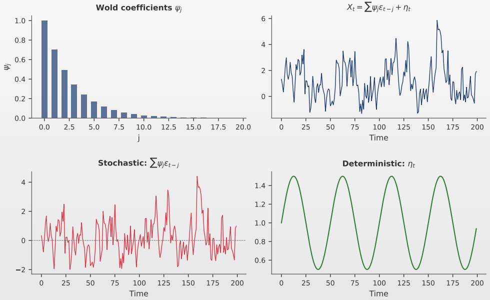
Orice proces **staționar în covarianță** $\{X_t\}$ poate fi scris ca: $X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} + \eta_t$

- $\varepsilon_t \sim WN(0, \sigma^2) \rightarrow$ zgomot alb
  - $\psi_0 = 1$, $\sum \psi_j^2 < \infty$
- $\eta_t \rightarrow$ componenta deterministă (perfect predictibilă)

Semnificația teoremei Wold

- **Descompunere**: Orice proces staționar = **MA(∞)** + componentă deterministă
  - Justifică teoretic modelele MA(q) și ARMA(p, q)
  - Coeficienții ψ_j măsoară impactul șocurilor trecute

Teorema Wold: ilustrare vizuală

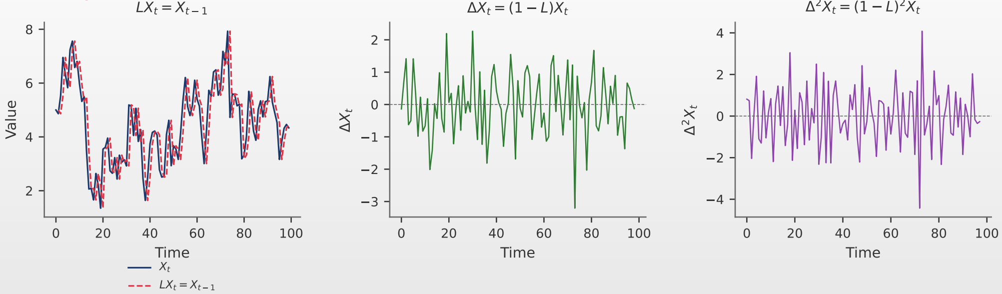


Interpretare

- X_t se descompune în componentă **stochastică** (MA(∞)) și componentă **deterministă** (η_t)
- Coeficienții ψ_j descresc $\to$ șocurile recente au impact mai mare decât cele îndepărtate

TSA_ch1_operators

Operatorul lag: ilustrare vizuală



Proprietăți

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

TSA_ch1_operators

Operatorul lag

Definiție 5 (Operatorul lag)

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

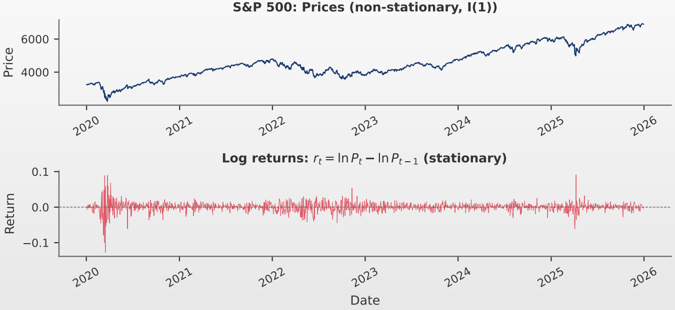
Proprietăți

- **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 ≈ 0, staționar

TSA_ch1_operators

Diferențierea

De ce diferențiem?

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

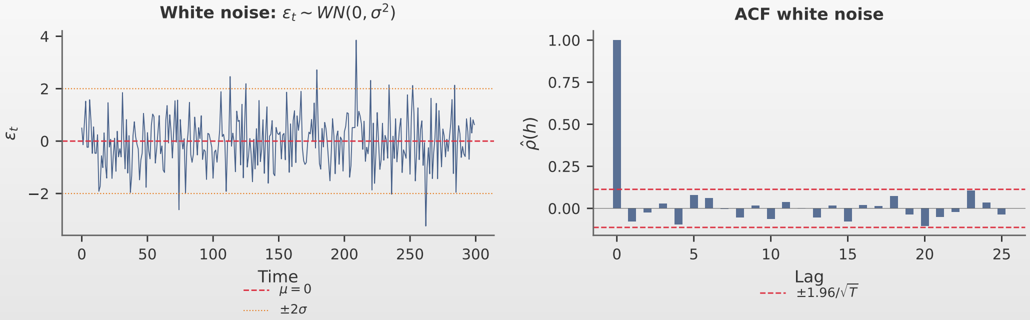
Definiție 6 (Proces Integrat de Ordin d)

- 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 aleatoriu $\rightarrow \Delta X_t = \varepsilon_t$ este staționar
- $I(2)$: Necesită două diferențieri pentru staționaritate

Zgomot alb: ilustrare vizuală



TSA_ch1_white_noise

Procesul de zgomot alb

Definiție 7 (Zgomot Alb)

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

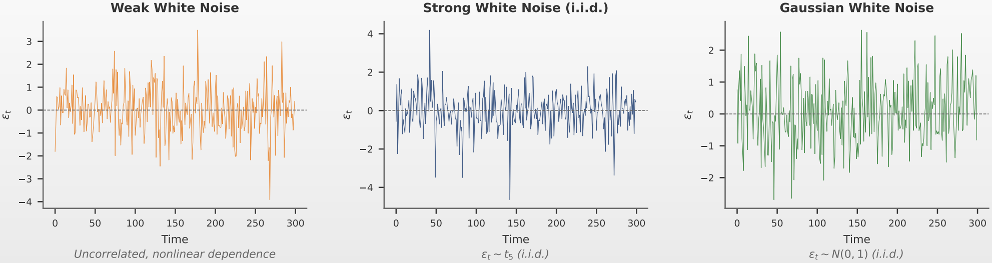
ACF al zgomotului alb

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

Tipuri de zgomot alb (în ordine crescătoare a restricțiilor)

- **Slab**: necorelat, dar pot exista dependențe neliniare
- **Puternic**: ε_t sunt *independente* și identic distribuite (i.i.d.)
- **Gaussian**: $\varepsilon_t \overset{iid}{\sim} N(0, \sigma^2)$ (necorelat $\Rightarrow$ independent)

Cele trei tipuri de zgomot alb

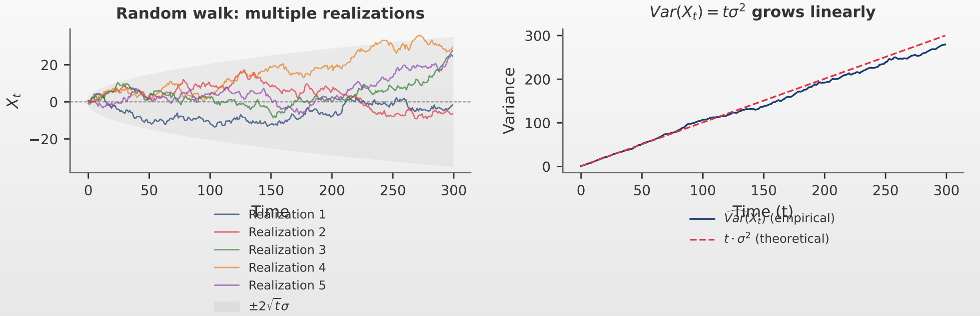


Relația de incluziune: Gaussian $\subset$ puternic (i.i.d.) $\subset$ slab (necorelat)

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

TSA_ch1_white_noise

Mers aleatoriu: vizualizare



Observații

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

TSA_ch1_random_walk

Procesul de mers aleatoriu

Definiție 8 (Mers Aleatoriu)

$X_t = X_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma^2), \ X_0 = 0 \quad \Rightarrow$ **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.

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

$\square$

Nestaționar!

$\text{Var}(X_t) = t\sigma^2$ depinde de t $\rightarrow$ mersul aleatoriu **nu este staționar**

Mers aleatoriu cu drift

Definiție 9 (Mers Aleatoriu cu Drift)

$X_t = c + X_{t-1} + \varepsilon_t$, $c \neq 0$ este **driftul** $\Rightarrow$ **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)

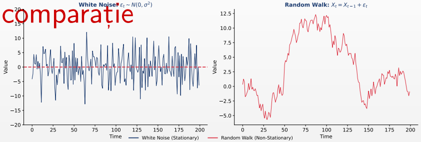
Diferențiere

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

Importanța practică

- PIB nominal, prețuri de acțiuni $\rightarrow$ adesea modele ca RW cu drift
- Testul ADF include variante: fără constantă, cu constantă, cu constantă și trend

Zgomot alb vs mers aleatoriu: comparație



Zgomot alb

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

Mers aleatoriu

- Nestaționar, Var $= t\sigma^2$ (crește), ACF ≈ 1 (lent), șocuri permanente

Legătură

- $\Delta X_t = \varepsilon_t$

TSA_ch1_random_walk

Staționaritate în trend vs. staționaritate în diferențe

Staționaritate în trend (TS)

- **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

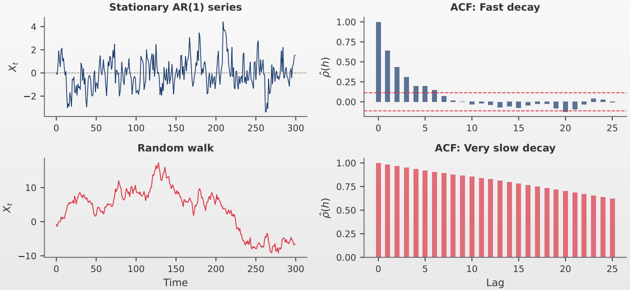
Staționaritate în diferențe (DS)

- **Model**: $Y_t = c + Y_{t-1} + \varepsilon_t$
  - Trend **stochastic**
  - Abaterile de la trend sunt permanente
- **Soluție**: diferențiere (ΔY_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

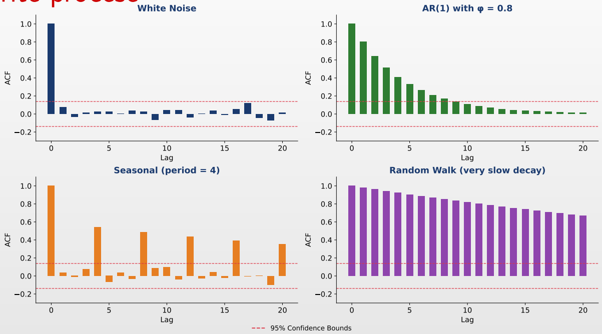


Interpretare

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

TSA_ch1_random_walk

Tipare ACF pentru diferite procese



Interpretare

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

TSA_ch1_acf_patterns

Functia de autocorelatie esantion

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: $\hat{\rho}(h) \approx N(0, 1/T)$

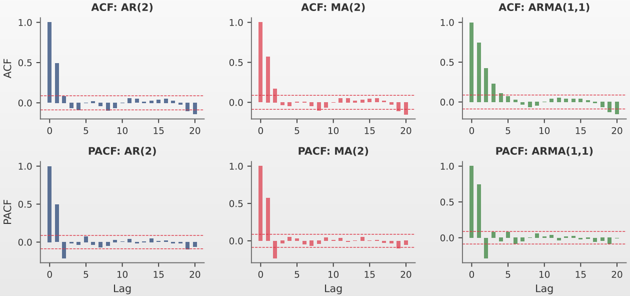
Interval de încredere 95%

- $\pm 1.96/\sqrt{T}$ (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

TSA_ch1_acf_patterns

Funcția de autocorelație parțială (PACF)

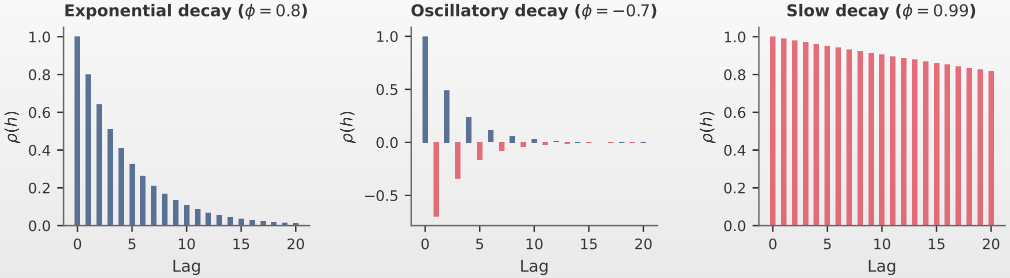
Definiție 10 (Autocorelația Parțială)

- **PACF** la lag-ul h, notat ϕ_{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**: Dependența *directă* la lag-ul h
  - Elimină efectul lag-urilor intermediare

Aplicație cheie: Identificarea ordinului modelului

- **AR(p)**: PACF **se anulează** după lag-ul p
  - ACF scade exponențial sau oscilant
- **MA(q)**: ACF **se anulează** după lag-ul q
  - PACF scade exponențial sau oscilant

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

TSA_ch1_acf_patterns

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 al X_{t-1}
- $SE(\hat{\gamma})$ din regresia OLS

Regula de decizie

- $\tau_{ADF} <$ val. critică $\to$ Respingem H_0 $\to$ Staționar
- $\tau_{ADF} \geq$ val. critică $\to$ 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)

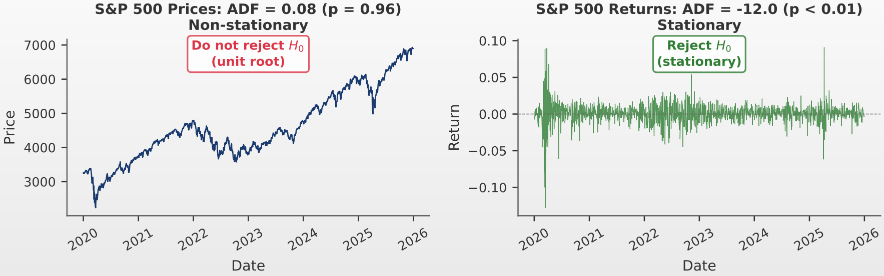
Statistica de Test

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

Regula de decizie

- $LM >$ valoarea critică $\rightarrow$ Respingem H_0 $\rightarrow$ Nestaționar
- $LM \leq$ valoarea critică $\rightarrow$ Staționar

Testul ADF: vizualizare cu S&P 500



TSA_ch1_unit_root_tests

Interpretarea testului ADF

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

Folosirea ADF și KPSS împreună

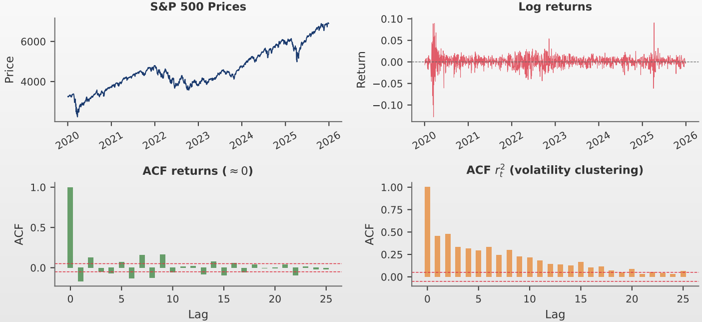
Testare confirmatorie

- **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, DF-GLS)

Flux de lucru

- **Pasul 1**: Test ADF (H_0: rădăcină)
- **Pasul 2**: Test KPSS (H_0: staționar)
- **Pasul 3**: Rezultate concordante → OK
  - Altfel: teste PP, DF-GLS

Analiza S&P 500: prezentare generală

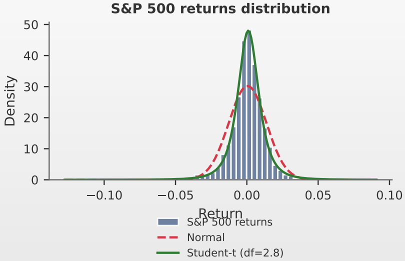


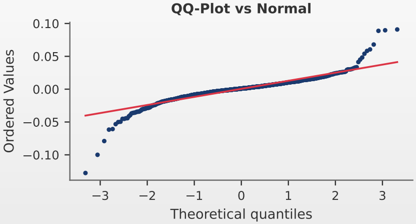
Observații

- **Prețuri**: Trend ascendent, nestaționar; **Randamente**: Medie ≈ 0, staționar
- **ACF randamente**: ≈ 0 (eficient); **ACF** r_t^2: Semnificativ (volatility clustering)

TSA_ch1_ts_basics

Fapte stilizate ale randamentelor financiare





Proprietăți observate

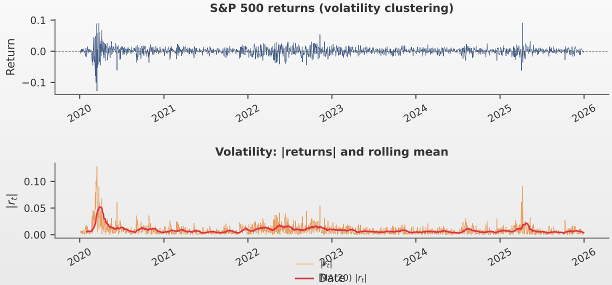
- Asimetrie negativă (coadă stângă)
- Kurtosis excesiv ($\gg 3$)
- Cozi groase (heavy tails)

Implicații

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

TSA_ch1_ts_basics

Volatility clustering

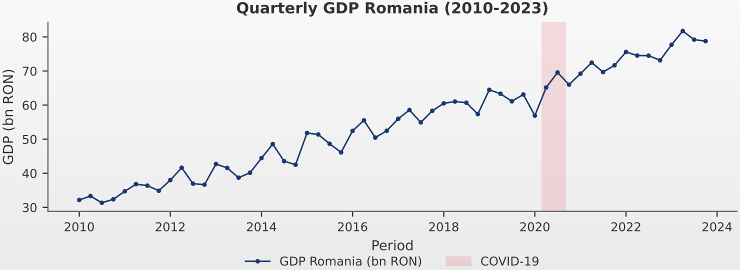


Observații

- Randamente mari (în valoare absolută) urmate de randamente mari
- Perioade de calm urmate de perioade de volatilitate ridicată
- **Volatilitate variabilă în timp** $\rightarrow$ modele ARCH/GARCH (Cap. 5)

TSA_ch1_ts_basics

Studiu de caz: PIB-ul trimestrial al României



TSA_ch1_case_gdp

Analiza inițială

- **Date**: PIB trimestrial România 2010–2023 (56 obs., INS/Eurostat)
- **Observații**: Trend ascendent, posibil sezonier
  - Șoc 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
- **ΔPIB**: staționar
  - Folosim ΔPIB_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 prognozei."

Exercițiu:

1. Rulați prompt-ul într-un LLM 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 prognoza 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: RMSE mic și coeficienți semnificativi *nu garantează* o prognoză utilă.

Concluzii principale

Rezumat

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

Formule importante

Staționaritate 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 aleatoriu (RW)

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

Previzualizare capitolul următor

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
- **Selecție**: AIC, BIC

Întrebarea 1

Întrebare

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

Variante de răspuns

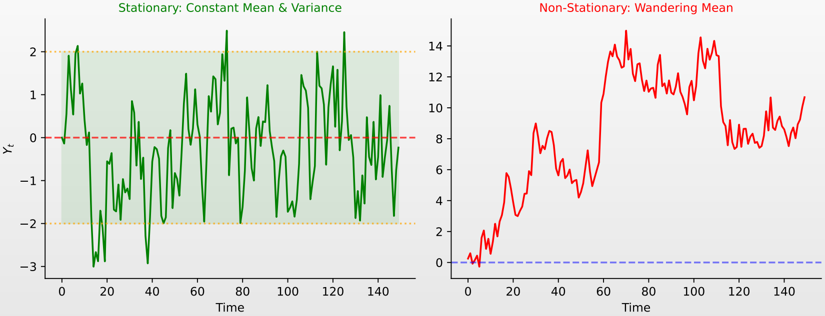
(A) Media zero, varianța infinită, covarianță dependentă de timp

(B) Media constantă, varianța constantă, autocovarianța depinde doar de lag

(C) Distribuție normală, independență, varianță unitară

(D) Trend liniar, sezonalitate constantă, reziduuri albe

Întrebarea 1: Răspuns



Răspuns: (B)

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

TSA_ch1_stationarity

Întrebarea 2

Întrebare

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

Variante de răspuns

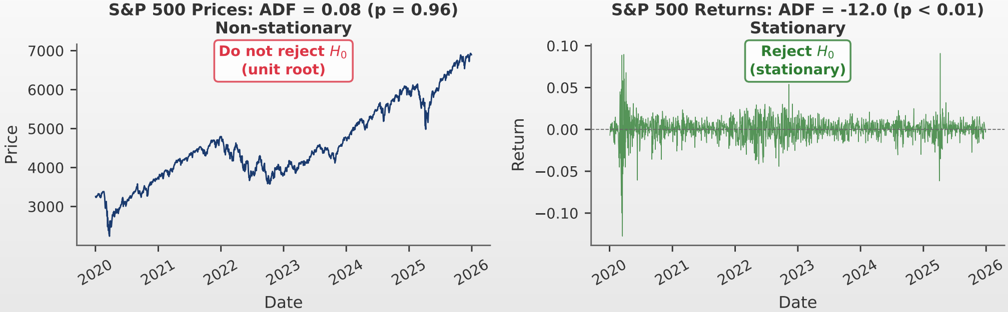
(A) Seria este staționară

(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 $\to$ staționară

TSA_ch1_unit_root_tests

Întrebarea 3

Întrebare

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

Variante de răspuns

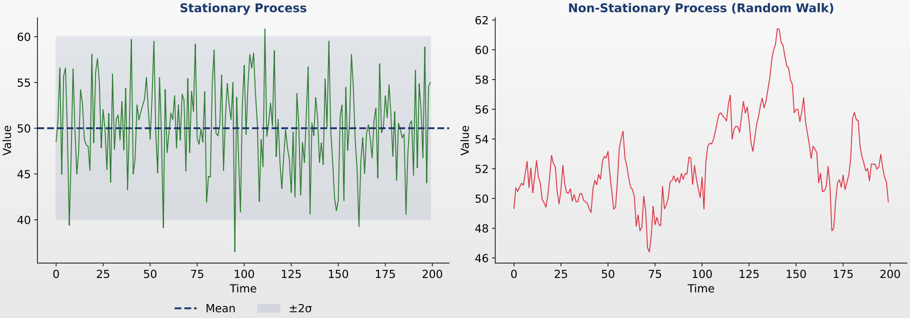
(A) Seria are rădăcină unitate (nestaționară)

(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!

TSA_ch1_unit_root_tests

Întrebarea 4

Întrebare

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

Variante de răspuns

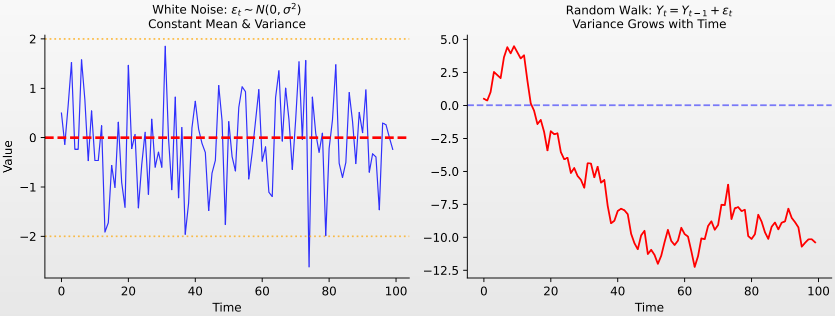
(A) Varianța este constantă: $\text{Var}(X_t) = \sigma^2$

(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



Răspuns: (B)

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

TSA_ch1_random_walk

Întrebarea 5

Întrebare

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

Variante de răspuns

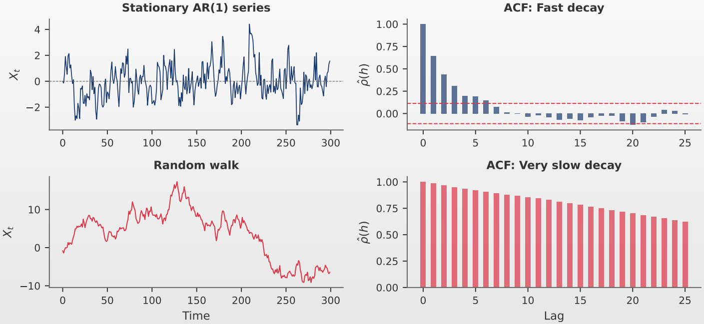
(A) Toate valorile sunt zero după lag 0

(B) Scade exponențial rapid

(C) Scade foarte lent (persistență înaltă)

(D) Oscilează între pozitiv și negativ

Întrebarea 5: Răspuns



Răspuns: (C)

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

TSA_ch1_random_walk

Întrebarea 6

Întrebare

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

Variante de răspuns

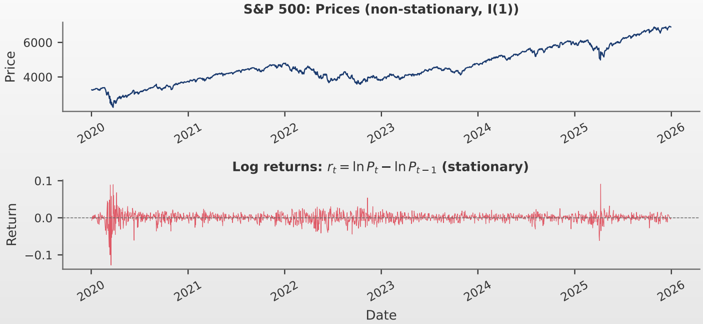
(A) Diferențiere simplă: $\Delta P_t = P_t - P_{t-1}$

(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 Δ (elimină trendul) $\rightarrow$ serie staționară

TSA_ch1_operators

Bibliografie

Manuale fundamentale

- Hyndman & Athanasopoulos (2021). *Forecasting*, OTexts
- Shumway & Stoffer (2017). *Time Series Analysis*, Springer
- Hamilton (1994). *Time Series Analysis*, Princeton

Referințe clasice

- Wold (1938). *Analysis of Stationary Time Series*
- Bartlett (1946). "Sampling Properties", *JRSS*

Vă Mulțumim!

Întrebări?

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

Quantlet Quantinar