



Analiza și Prognoza Seriilor de Timp

Capitolul 0: Fundamente



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 sfârșitul acestui capitol, veți putea 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-ciclu, sezonalitate și reziduuri
3. Aplicați metodele de netezire exponențială (SES, Holt, Holt-Winters, ETS)
4. Evaluați prognozele folosind MAE, RMSE, MAPE, sMAPE
5. Implementați separarea train/validare/test și validarea încrucișată
6. Modelați sezonalitatea folosind variabile dummy sau termeni Fourier
7. Eliminați trendul și sezonalitatea prin metode adecvate
8. Distingeți între trendurile deterministe și stochastice



Surse de date și instrumente software

Surse de date

- **Yahoo Finance**
 - ▶ Prețuri acțiuni, criptomonede, valute
- **FRED (Federal Reserve)**
 - ▶ PIB, șomaj, rate dobânzi
- **Eurostat / INS / BNR**
 - ▶ Date economice europene și românești
- **Seturi clasice**
 - ▶ AirPassengers, Sunspots, CO₂

Python

- `yfinance` — date Yahoo Finance
- `pandas_datareader` — FRED, Eurostat
- `statsmodels` — modele statistice
- `pandas` — manipulare date
- `matplotlib` — vizualizare

R

- `quantmod` — date financiare
- `tseries` — teste serii de timp
- `forecast` — modele de prognoză
- `fredr` — acces FRED API



Structura capitolului

- Motivație
- Ce Este o Serie de timp?
- Descompunerea seriilor de timp
- Metode de Netezire Exponențială
- Evaluarea prognozei
- Modelarea sezonalității
- Gestionarea Trendului și Sezonalității
- Utilizare IA
- Rezumat



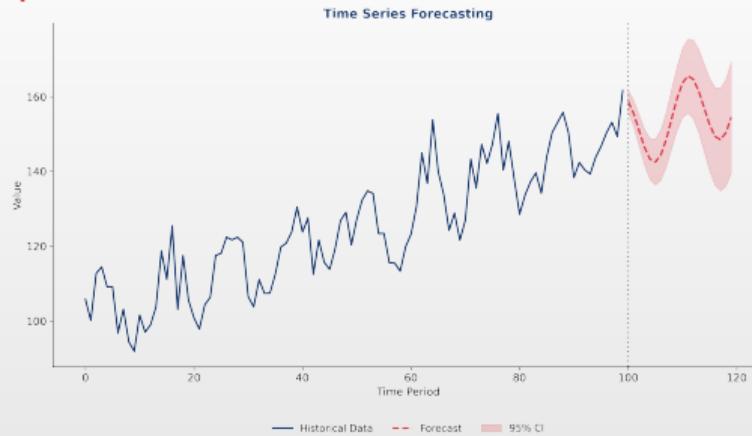
Seriile de timp sunt prețutindeni



TSA_ch0_real_data

- Finanțe:** Prețuri acțiuni, cursuri valutare, volume
- Economie:** PIB, șomaj, rate ale inflației
- Business:** Vânzări, trafic website, cerere
- Știință:** Temperatură, poluare, semne vitale

De ce studiem seriile de timp?



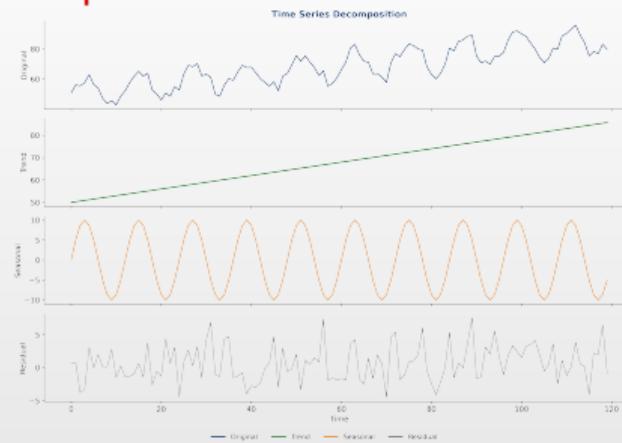
○ TSA cu real data

Obiectiv principal: prognoza

- Folosim tiparele istorice pentru a prezice valori viitoare → esențial pentru planificarea afacerilor, managementul riscului și deciziile de politică



Înțelegerea structurii seriilor de timp



Q TSA cho real data

Descompunere

- Orice serie de timp poate fi descompusă în: **trend-ciclu + sezonialitate + zgomot**



Definiția unei serii de timp

Definiție 1 (Serie de timp)

- **Serie de timp:** secvență de observații $\{X_t\}$ indexate după timp: $\{X_t : t \in \mathcal{T}\}$ unde \mathcal{T} este o mulțime de indici reprezentând momente de timp

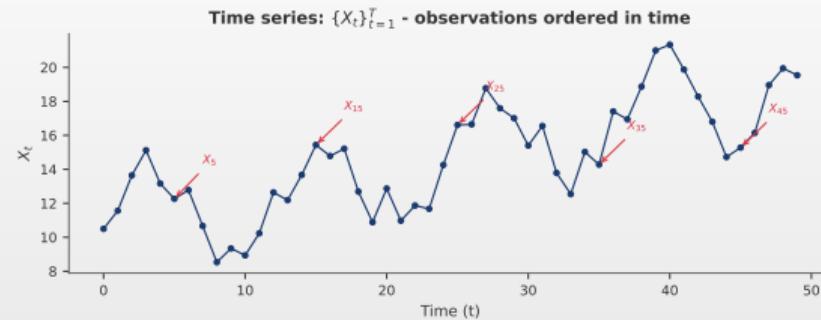
Caracteristici cheie

- **Ordonate:** ordine temporală naturală
- **Dependente:** observațiile consecutive sunt corelate
- **Discrete/Continue:** $t = 1, 2, 3, \dots$

Notație

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

Serie de timp: ilustrație conceptuală

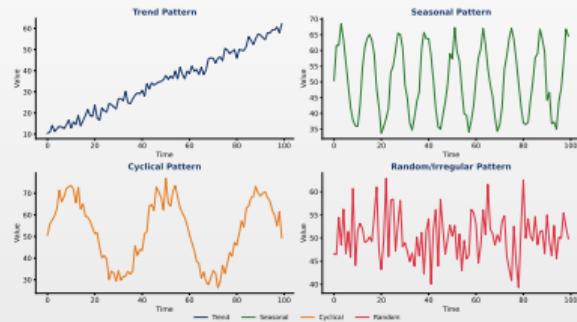


Q TSA_cho_definition

Elemente fundamentale

- Notăție formală:** X_t = valoarea la momentul t , $t \in \{1, 2, \dots, T\}$
- Autocorelație:** $\rho_k = \text{Corr}(X_t, X_{t-k})$ — măsoară dependența temporală

Tipare comune în seriile de timp



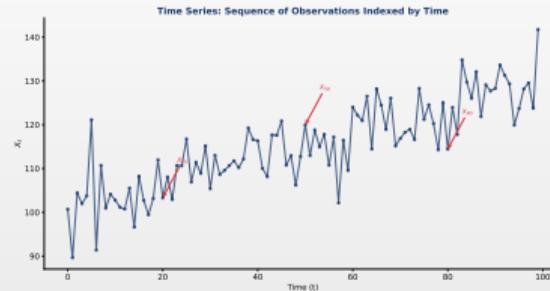
Q TSA_ch0_definition

Tipuri de tipare

- Trend:** creștere sau scădere pe termen lung
- Sezonier:** tipare periodice regulate
- Ciclic:** fluctuații pe termen mediu (2–10 ani)
- Aleatoriu:** fluctuații imprevizibile



Exemplu practic: date financiare reale

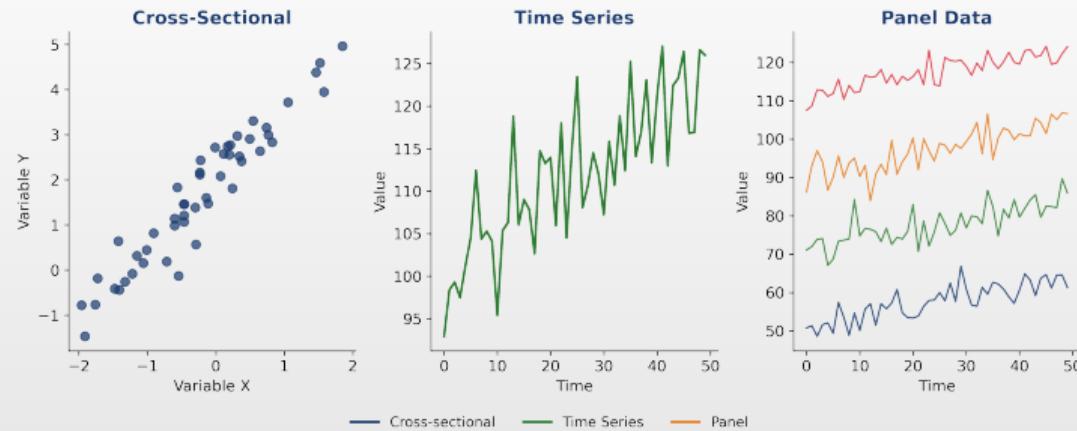


Q TSA_cho_definition

S&P 500 (2024)

- Frecvență zilnică:** ≈ 252 zile de tranzacționare/an
- Caracteristici observate:** trend ascendent, volatilitate în clustere, persistență (momentum)

Tipuri de date: comparație



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-ul a 50 țări, 20 ani



Exemple de date de tip serie de timp



Q TSA_ch0_real_data

Date financiare reale

- Sursă:** Yahoo Finance (2019–2025), normalizează la baza 100
- Bitcoin:** cel mai volatil
- Aur:** cel mai stabil



De ce descompunem o serie de timp?

Obiective

- Înțelegerea tiparelor subiacente
- Eliminarea sezonalității pentru modelare
- Identificarea direcției trendului
- Izolare fluctuațiilor neregulate
- Îmbunătățirea acurateții prognozei

Componente

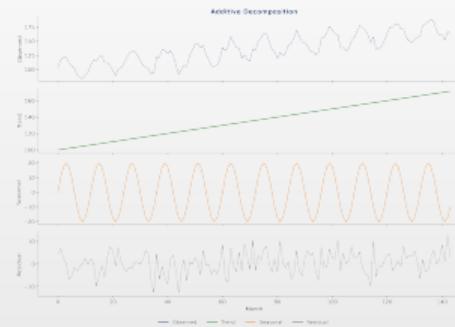
- T_t : Trend-Ciclu
 - ▶ Mișcare pe termen lung
- S_t : Sezonier
 - ▶ Tipar periodic regulat
- ε_t : Reziduu
 - ▶ Zgomot aleatoriu

Modele clasice de descompunere

- Aditiv:** $X_t = T_t + S_t + \varepsilon_t$
 - ▶ Amplitudine sezonieră constantă
- Multiplicativ:** $X_t = T_t \times S_t \times \varepsilon_t$
 - ▶ Amplitudine sezonieră crește cu nivelul



Descompunerea seriilor de timp: exemplu vizual



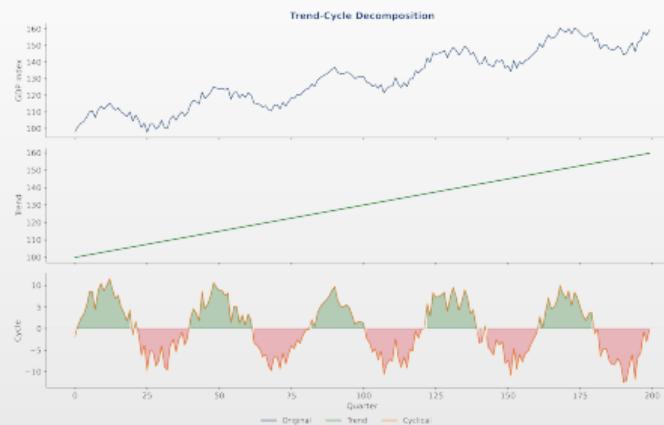
TSA_ch0_decomposition

Componente explicate

- Original:** seria observată
- Trend-Ciclu:** mișcare pe termen lung
- Sezonier:** tipar periodic
- Reziduu:** zgomot aleatoriu



Componenta ciclică



Q TSA_ch0_decomposition

Caracteristici

- **Durată:** fluctuații pe termen mediu (2–10 ani)
- **Aperiodic:** fără perioadă fixă (vs sezonalitate)
- **Origine:** reflectă ciclurile economice

În practică

- **Combinare:** ciclul combinat cu trendul
- **Dificultate:** greu de identificat în serii scurte
- **Soluție:** de obicei absorbit în trend-ciclu



Modelul de descompunere aditivă

Model

- Ecuatie:** $X_t = T_t + S_t + \varepsilon_t$
 - ▶ Componentele se adună pentru a forma seria observată

Când să folosim

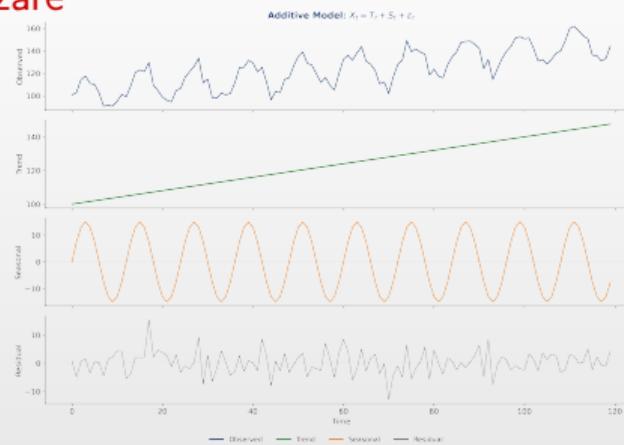
- Fluctuații sezoniere constante**
 - ▶ Amplitudinea nu depinde de nivel
- Varianța seriei stabilă**
 - ▶ Măsoară dispersia în jurul mediei
 - ▶ Estimator: $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$

Proprietăți

- Eroare:** $\mathbb{E}[\varepsilon_t] = 0$ (medie zero)
- Sezonier:** $\sum_{j=1}^s S_j = 0$ (suma sezonala e zero)
- Unități:** S_t sunt aceleasi ca X_t



Descompunere aditivă: vizualizare



Q TSA_ch0_decomposition

Interpretare

- **Descompunere:** Original = Trend + Sezonier + Reziduu
- **Proprietate:** amplitudine sezonieră constantă, nu depinde de nivel

Modelul de descompunere multiplicativă

Model

- **Ecuatie:** $X_t = T_t \times S_t \times \varepsilon_t \rightarrow$ componentele se înmulțesc

Când să folosim

- **Fluctuații crescătoare:** sezonalitatea crește cu nivelul
- **Heteroscedasticitate:** varianța crește în timp
- **Exemple:** date economice/financiare

Proprietăți

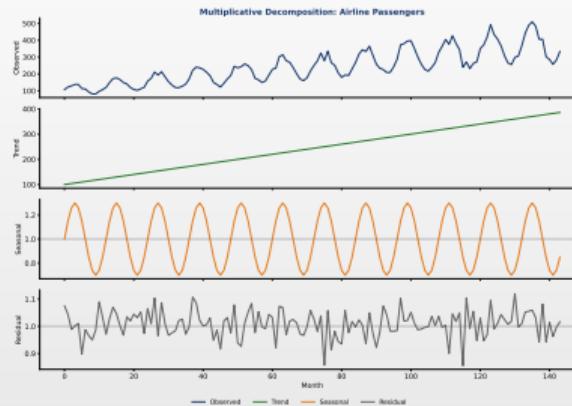
- **Eroare:** $\mathbb{E}[\varepsilon_t] = 1$ (centrat la 1)
- **Sezonier:** $\frac{1}{s} \sum_{j=1}^s S_j = 1$ (media e 1)
- **Unități:** S_t este raport adimensional

Sfat

- **Transformare logaritmică:** multiplicativ \rightarrow aditiv: $\log X_t = \log T_t + \log S_t + \log \varepsilon_t$



Descompunere multiplicativă: date reale



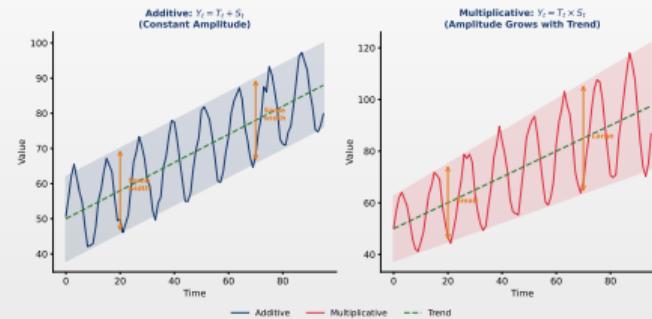
Exemplu

- Date Box-Jenkins: pasageri lunari (1949–1960). Amplitudinea sezonieră crește cu nivelul

[TSA_ch0_decomposition](#)



Aditivă vs multiplicativă: comparație



Q TSA_ch0_decomposition

Diferența cheie

- Multiplicativ:** componenta sezonieră este un *raport*, centrată la valoarea 1
- Aditiv:** componenta sezonieră în *unități absolute*, centrată la valoarea 0



Estimarea trendului: media mobilă

Definiție 2 (media mobilă centrată)

- **Media mobilă centrată** de ordin $2q + 1$:

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

Pentru date sezoniere

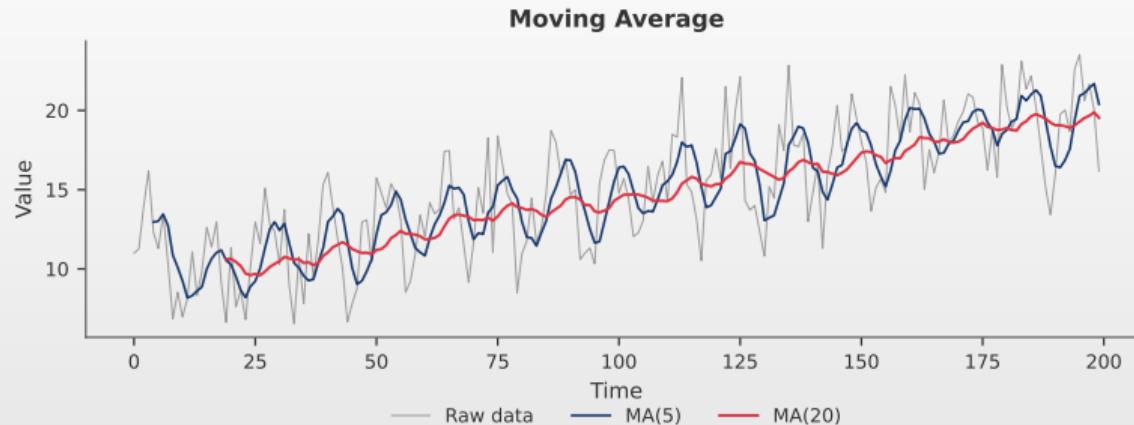
- **Perioada s impară**
 - ▶ Se folosește medie simplă
- **Perioada s pară**
 - ▶ $2 \times s$ MA cu ponderi jumătate

Proprietăți

- **Netezire:** elimină sezonierul & aleatoriu
- **Fereastră mare** → estimare mai netedă
- **Dezavantaj:** pierdere de date la extremități



Media mobilă centrată: ilustrație vizuală



Interpretare

- Netezire:** elimină fluctuațiile pe termen scurt
- Rezultat:** dezvăluie trendul subiacent



Algoritmul descompunerii clasice

Pași pentru descompunerea multiplicativă

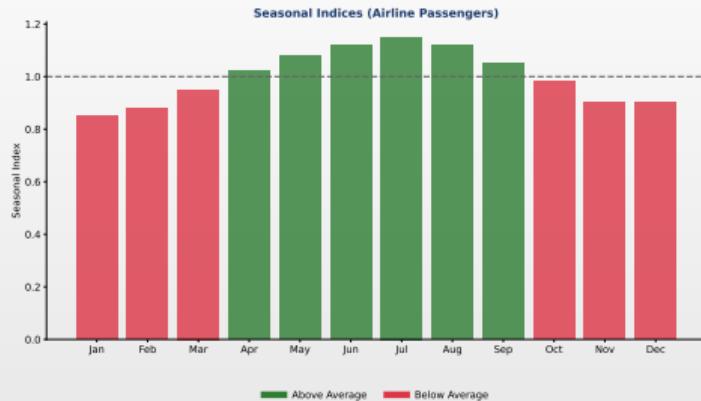
- **Pasul 1 → Estimare Trend:** $\hat{T}_t = MA_s(X_t)$
 - ▶ Medie mobilă centrată de ordinul perioadei sezoniere
- **Pasul 2 → Eliminarea trendului:** $D_t = X_t / \hat{T}_t$
- **Pasul 3 → Estimare Sezonier:** $\hat{S}_j = \text{media}(D_t \text{ pentru sezonul } j)$
- **Pasul 4 → Normalizare:** scalare astfel încât $\frac{1}{s} \sum_{j=1}^s \hat{S}_j = 1$
- **Pasul 5 → Calcul Reziduuri:** $\hat{\varepsilon}_t = X_t / (\hat{T}_t \times \hat{S}_t)$

Notă

- **Pentru descompunere aditivă:** operațiile se modifică
 - ▶ Împărțire → scădere
 - ▶ Înmulțire → adunare



Indici sezonieri: interpretare



Interpretare

- $S_t > 1$: activitate peste medie; $S_t < 1$: sub medie. Vârf de călătorii în iulie–august

Q TSA_cho_seasonal



Descompunerea STL: o abordare modernă

Definiție 3 (STL - Descompunere Sezonier-Trend folosind LOESS)

- **STL:** folosește regresie locală ponderată (LOESS): $X_t = T_t + S_t + R_t$

Avantaje

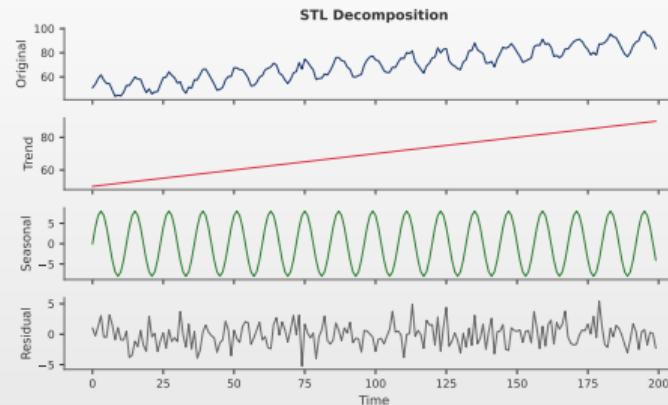
- **Flexibilitate:** orice perioadă sezonieră
- **Variabilitate:** sezonalitatea poate evoluă în timp
- **Robustete:** rezistentă la valori extreme
- **Netezire:** estimări netede ale trendului

Parametri cheie

- **period:** perioada sezonieră
 - ▶ Ex: 12 pentru date lunare, 4 pentru trimestriale
- **seasonal:** fereastra de netezire
- **robust:** ponderare redusă pentru outlieri



Descompunerea STL: ilustrație vizuală



Idee cheie

- STL (Seasonal-Trend-Loess): separă trend + sezonier + rest folosind regresie LOESS

[TSA_cho_decomposition](#)



Netezirea exponențială: prezentare generală

Definiție

- **Netezirea exponențială:** medii ponderate ale observațiilor trecute
 - ▶ Ponderile scad exponențial în timp
 - ▶ Observațiile recente primesc ponderi mai mari

De ce netezire exponențială?

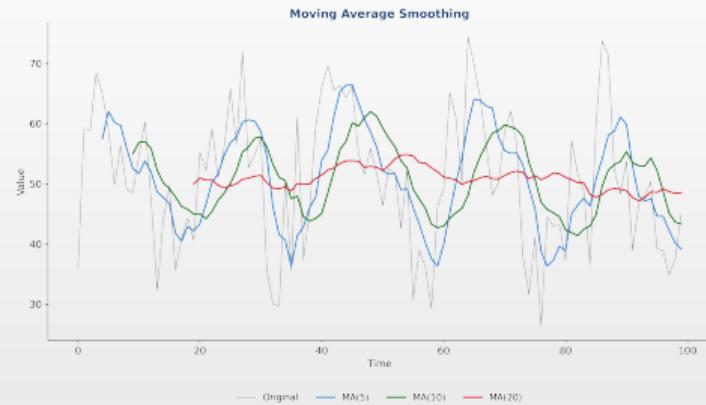
- **Simplă:** ușor de implementat și înțeles
 - ▶ Un singur parametru de netezire
- **Adaptivă:** ponderi mai mari pentru date recente
- **Versatilă:** gestionează trend și sezonalitate

Trei metode principale

- **SES (Simple Exponential Smoothing):** doar nivel
 - ▶ Cea mai simplă metodă exponențială
- **Holt:** nivel + trend
 - ▶ Captează direcția de evoluție
- **Holt-Winters:** + sezonalitate
 - ▶ Model complet cu toate componente



Netezirea cu media mobilă



Compromisul dimensiunii ferestrei

- Fereastră mică:** reactivă dar zgomotoasă
 - ▶ Captează schimbări rapide, dar amplifică zgomotul
- Fereastră mare:** netedă dar cu întârziere
 - ▶ Elimină zgomotul, dar reacționează lent

Q TSA_cho_definition



Netezirea exponențială simplă (SES)

Model

- **Ecuatie:** $\hat{X}_{t+1|t} = \alpha X_t + (1 - \alpha)\hat{X}_{t|t-1}$
 - ▶ $\alpha \in (0, 1)$ este parametrul de netezire

Cum funcționează

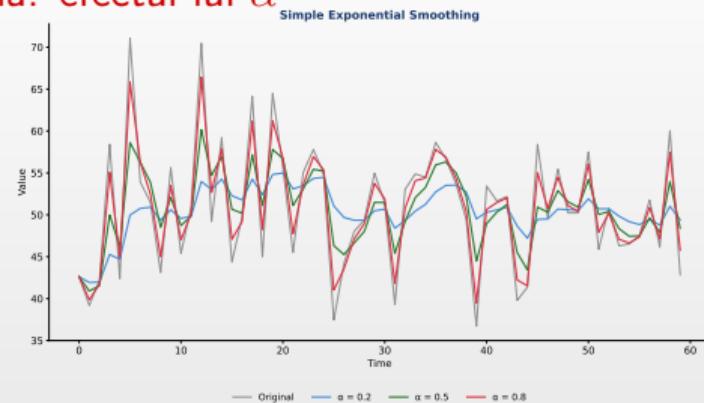
- **Principiu:** ponderile scad exponențial
- α mare
 - ▶ Prognoză reactivă la schimbări
- α mic
 - ▶ Prognoză mai netedă, stabilă

Forma cu nivel

- **Ecuatie:** $\ell_t = \alpha X_t + (1 - \alpha)\ell_{t-1}$
 - ▶ ℓ_t = nivelul estimat la momentul t
 - ▶ Prognoză: $\hat{X}_{t+h|t} = \ell_t$ (constantă)



Netezirea exponențială simplă: efectul lui α



Compromis

- α **mic** → programe netede
 - ▶ Mai multă pondere pe istoria îndepărtată
- α **mare** → urmărește datele
 - ▶ Reacție rapidă la schimbări recente

Q TSA_ch0_smoothing



SES: exemplu numeric pas cu pas

Date: Vânzări lunare (mii EUR)

Date: $X_1 = 100, X_2 = 110, X_3 = 105, X_4 = 115, X_5 = 120 \quad (\alpha = 0.3, \hat{X}_{1|0} = 100)$

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

t	X_t	$\hat{X}_{t t-1}$	e_t	Calcul $\hat{X}_{t+1 t}$
1	100	100.00	0.00	$0.3 \times 100 + 0.7 \times 100 = 100.00$
2	110	100.00	10.00	$0.3 \times 110 + 0.7 \times 100 = 103.00$
3	105	103.00	2.00	$0.3 \times 105 + 0.7 \times 103 = 103.60$
4	115	103.60	11.40	$0.3 \times 115 + 0.7 \times 103.6 = 107.02$
5	120	107.02	12.98	$0.3 \times 120 + 0.7 \times 107.02 = 110.91$

Prognosă și evaluare

$$\hat{X}_{6|5} = 110.91 \quad \text{MAE} = 7.28 \quad \text{RMSE} = 8.97$$



Metoda Holt cu trend liniar

Ecuări

- 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:** $\hat{X}_{t+h|t} = \ell_t + h \cdot b_t$
 - Extrapolează trendul liniar pe h pași

Parametri

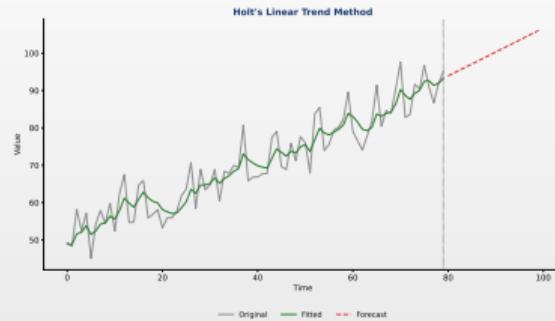
- α : netezire nivel
 - Controlează reactivitatea la schimbări de nivel
- β^* : netezire trend
 - Controlează reactivitatea la schimbări de pantă

Componente

- ℓ_t : nivel estimat
 - Media locală a seriei
- b_t : trend estimat (pantă)
 - Rata de creștere/descrescere



Metoda Holt: vizualizare



Q TSA_ch0_smoothing

Interpretare

- Metoda Holt:** captează nivelul și trendul, le proiectează în orizontul de prognoză
- α : controlează schimbări de nivel
- β^* : controlează schimbări de trend



Metoda sezonieră Holt-Winters

Ecuări (Sezonalitate aditivă)

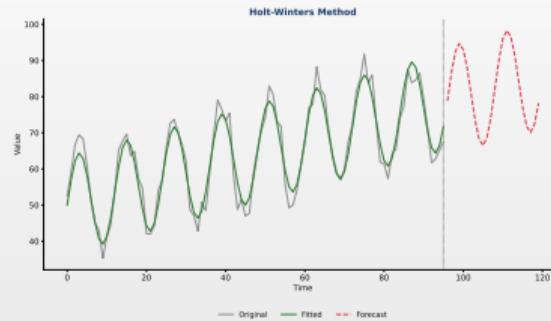
- Nivel:** $\ell_t = \alpha(X_t - S_{t-s}) + (1 - \alpha)(\ell_{t-1} + b_{t-1})$
- Trend:** $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1}$
- Sezonier:** $S_t = \gamma(X_t - \ell_t) + (1 - \gamma)S_{t-s}$
- Prognoză:** $\hat{X}_{t+h|t} = \ell_t + h \cdot b_t + S_{t+h-s(k+1)}$
 - Unde $k = \lfloor (h-1)/s \rfloor$

Parametri

- α — nivel
- β^* — trend
- γ — sezonier
- s — perioadă sezonieră
 - Toți în $(0, 1)$; estimați prin minimizarea erorii



Holt-Winters: captarea sezonalității



Q TSA_ch0_smoothing

Caracteristică cheie

- Descompunere completă:** separă nivel, trend și sezonier
- Prognoze sezoniere:** include atât trend cât și tipar periodic



Cadrul ETS: eroare-trend-sezonalitate

Definiție 4 (Modele ETS)

- **Cadrul ETS:** generalizează netezirea exponențială: $ETS(E, T, S)$

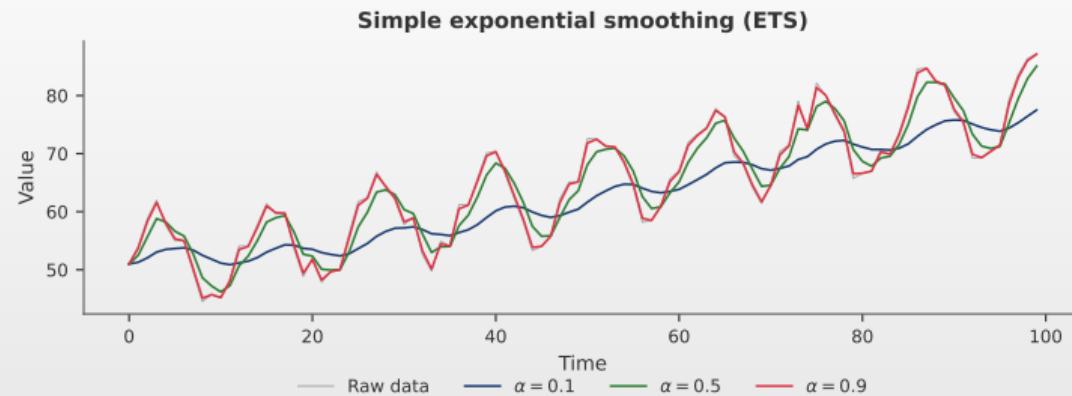
Componentă	N	A	M
Eroare (E)	–	aditivă	multiplicativă
Trend (T)	Niciunul	Aditiv	Multiplicativ
Sezonier (S)	Niciunul	Aditiv	Multiplicativ

Exemple

- **ETS(A,N,N):** Netezire exponențială simplă → doar nivel, fără trend sau sezonalitate
- **ETS(A,A,N):** Metoda Liniară Holt → nivel + trend aditiv
- **ETS(A,A,A):** Holt-Winters aditivă → nivel + trend + sezonalitate aditivă



ETS: ilustrație netezire exponențială



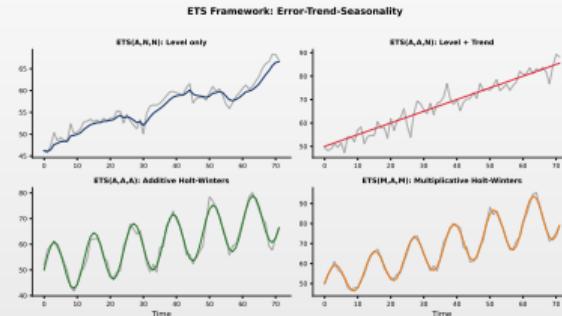
Interpretare

- Observații ponderate exponențial: ponderile scad cu vechimea; observațiile recente au importanță mai mare

TSA_ch0_smoothing



Selectia modelului ETS



TSA_ch0_smoothing

Selectie automată

- Criterii informaționale:** AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion)
- Selectie optimală:** balanță între ajustare și complexitate



Metode cu trend amortizat

Parametrul de amortizare

- Parametru:** $\phi \in (0, 1)$
 - ▶ Previne supra-proiecția trendului
 - ▶ Trendul converge către o constantă

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

Idee cheie

- Asimptotic:** când $h \rightarrow \infty$, prognoza \rightarrow constantă
 - ▶ Previne extrapolare nerealistă pe termen lung
- Avantaj:** adesea mai bună pentru orizonturi lungi



Metrici de acuratețe a prognozei

Eroarea de prognoză

- Definiție:** $e_t = X_t - \hat{X}_t$ (actual minus prezent)
- Pozitivă \Rightarrow subestimează; Negativă \Rightarrow supraestimează

Dependente de scală

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

Independente de scală

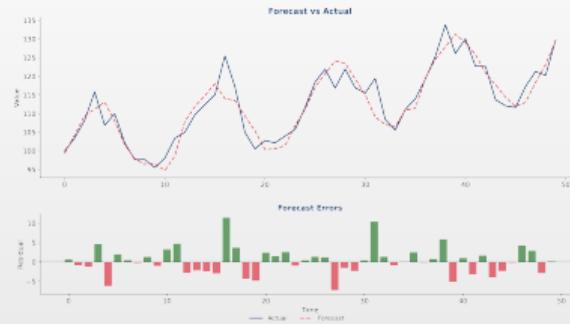
- MAPE:** $\frac{100}{n} \sum \left| \frac{e_t}{X_t} \right|$
- sMAPE:** $\frac{100}{n} \sum \frac{|e_t|}{(|X_t| + |\hat{X}_t|)/2}$

Ce să folosim?

- Aceeași serie:** RMSE, MAE \rightarrow comparare modele pe aceleasi date
- Între serii diferite:** MAPE, sMAPE \rightarrow metrici procentuale, independente de scală



Evaluarea prognozei: exemplu vizual



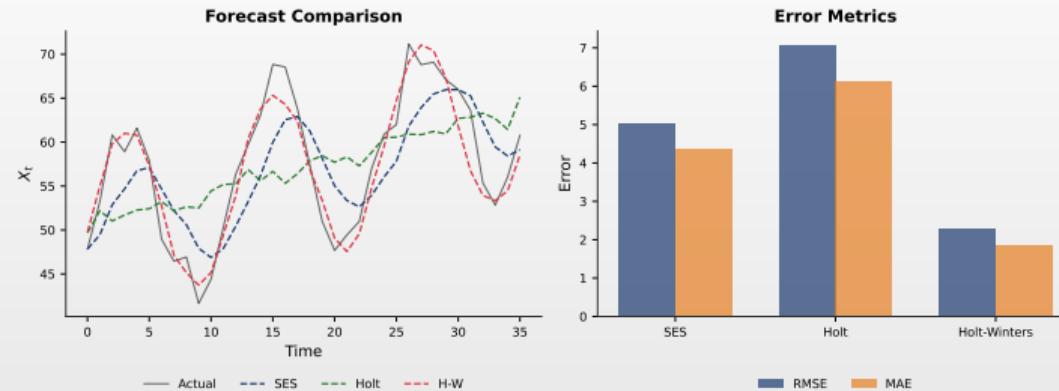
Observații

- Sus: actual vs. prognoză — evaluare vizuală a calității prognozei
- Jos: reziduuri — medie zero, varianță constantă, fără tipar

TSA_ch0_forecast_eval



Compararea metodelor de prognoză



Interpretare

- Stânga: Prognoze SES, Holt, Holt-Winters. Dreapta: Metriici de eroare. Comparație vizuală și cantitativă

Q TSA_ch0_forecast_eval



Diagnosticarea reziduurilor

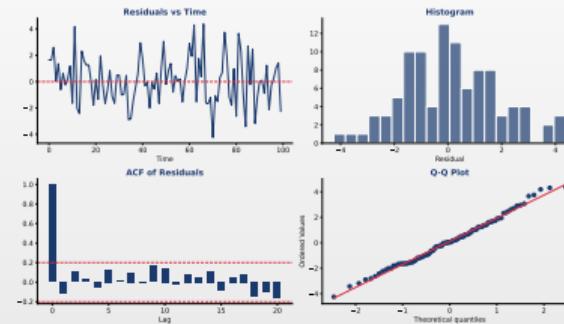
Proprietăți ale reziduurilor

- **Medie zero:** $\mathbb{E}[e_t] = 0$
 - ▶ Prognoza nu are bias sistematic
- **Necorelate:** $\text{Cov}(e_t, e_{t-k}) = 0$
 - ▶ Nu rămâne informație neexploatață
- **Varianță constantă:** $\text{Var}(e_t) = \sigma^2$
- **Normal distribuite:** pentru intervale de încredere

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):**
 - ▶ $JB = \frac{T}{6} \left(S^2 + \frac{(K-3)^2}{4} \right) \sim \chi_2^2$
 - ▶ S = asimetrie, K = curtosis

Diagnosticarea reziduurilor: vizualizare



TSA_ch0_forecast_eval

Ce să verificăm

- Grafic temporal:** fără tipare sistematice
- Histogramă:** verificare normalitate
- ACF:** fără autocorelație semnificativă
- Grafic Q-Q:** confirmare normalitate



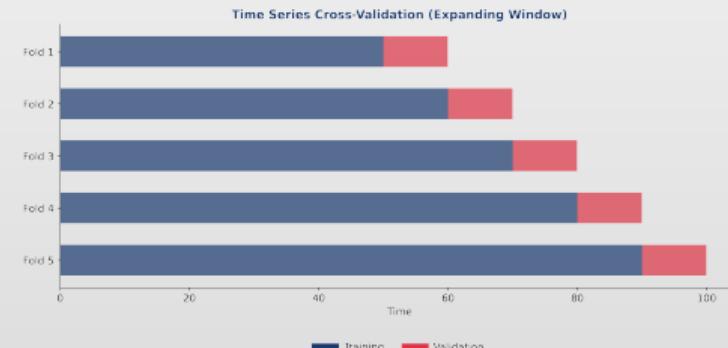
Validarea încrucișată pentru serii de timp

De ce nu CV standard?

- Dependență temporală**: observațiile sunt corelate
- Ordinea contează**: trebuie respectată cronologia
- K-fold standard** → data leakage

CV (cross-validation) cu origine mobilă

- Pasul 1**: antrenare pe $\{X_1, \dots, X_t\}$
- Pasul 2**: prognoză \hat{X}_{t+h}
- Pasul 3**: incrementare t , repetare



Separarea train / validare / test

Set de Antrenare

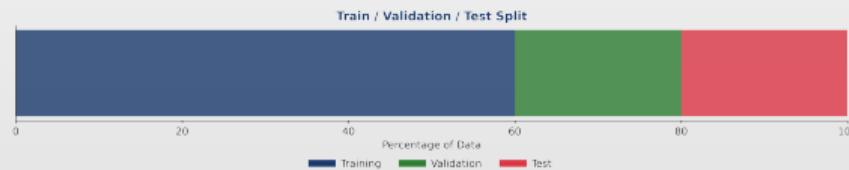
- Potrivirea parametrilor modelului
- Cea mai mare porțiune (60–80%)
- Folosit pentru estimare

Set de Validare

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

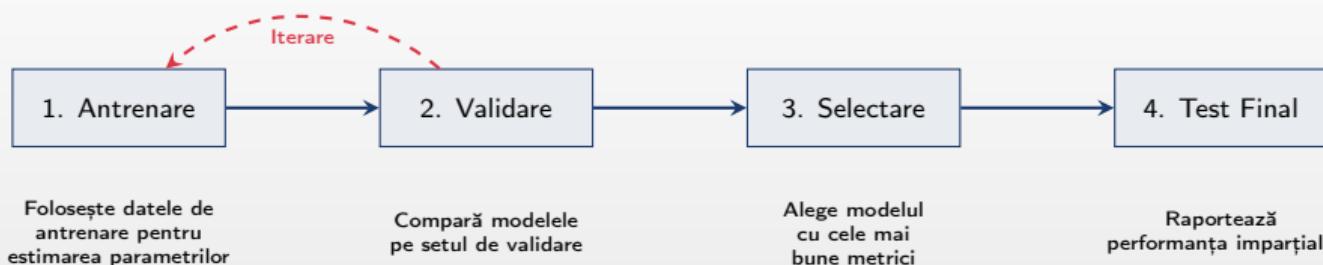
Set de Test

- Doar evaluare finală
- Nu se folosește pentru ajustare
- Performanță imparțială



Q TSA_ch0_forecast_eval

Fluxul de lucru pentru dezvoltarea modelului

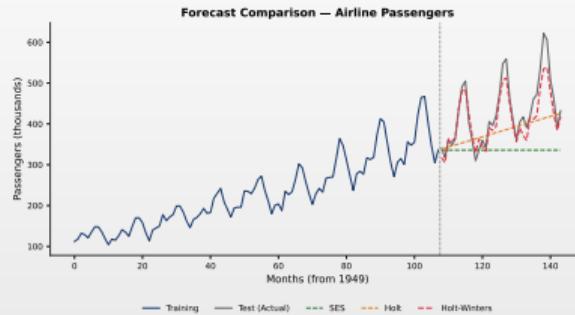


Regulă critică

- Niciodată testul pentru selecție!**
 - ▶ Folosiți doar pentru evaluare finală
- Evitați surgerea de date**
 - ▶ Estimări prea optimiste ale performanței



Date reale: compararea prognozelor



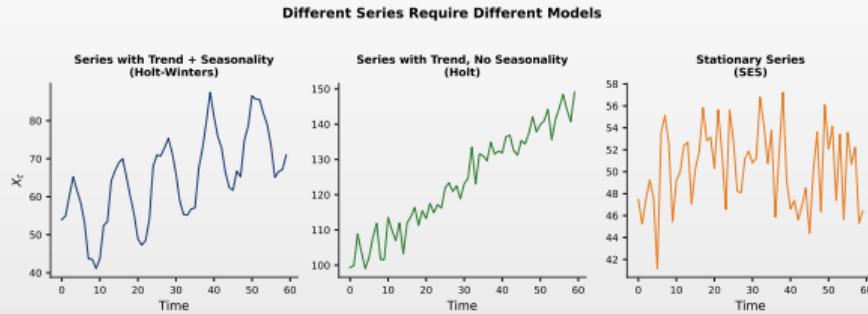
 TSA_ch0_forecast_eval

Interpretare

- Date:** pasageri companii aeriene
- Cel mai bun:** Holt-Winters multiplicativă — ideal pentru date cu sezonalitate crescătoare



Performanța prognozei pe diferite seturi de date



TSA_ch0_forecast_eval

Interpretare

- Serii diferite:** necesită modele diferite
- Date sezoniere:** preferați metode sezoniere
- Nu există model universal:** testați mai multe abordări

Modelarea sezonalității: două abordări

1. Variabile dummy

- Model:** $X_t = \mu + \sum_{j=1}^{s-1} \gamma_j D_{jt} + \varepsilon_t$
- $D_{jt} = 1$ dacă t în sezonul j
- $s - 1$ parametri
- Orice tipar sezonier

2. Termeni Fourier

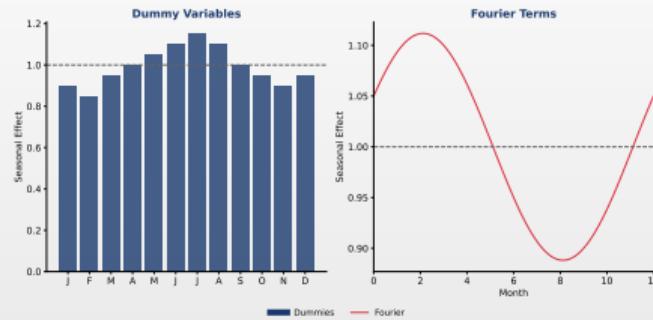
- Model:** $X_t = \mu + \sum_{k=1}^K \left[\alpha_k \sin\left(\frac{2\pi kt}{s}\right) + \beta_k \cos\left(\frac{2\pi kt}{s}\right) \right]$
- Funcții sinusoidale
- $2K$ parametri
- Tipare netede

Compromis

- Variabile dummy**
 - Orice tipar sezonier, dar mai mulți parametri
- Termeni Fourier**
 - Tipare netede, mai puțini parametri



Variabile dummy vs termeni Fourier



Q TSA_cho_seasonal

Comparatie

- **Variabile dummy:** captează orice formă, necesită $s - 1$ parametri
- **Termeni Fourier:** doar $2K$ parametri, tipare netede, sinusoidale



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 (adăugați termeni)

Recomandări

- Folosiți Dummy**
 - ▶ Tipare neregulate, coeficienți interpretabili
- Folosiți Fourier**
 - ▶ Tipare netede, sezonalitate de înaltă frecvență
 - ▶ Utilizat în TBATS și Prophet



De ce eliminăm trendul și sezonalitatea?

Motive pentru eliminarea trendului

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

Motive pentru desezonalizare

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

Important

- Modelăm seria transformată**
 - ▶ Cu trendul și sezonalitatea eliminate
- Inversăm transformarea**
 - ▶ Readucem prognoza la scala originală



Metode de eliminare a trendului

Șase abordări comune de eliminare a trendului

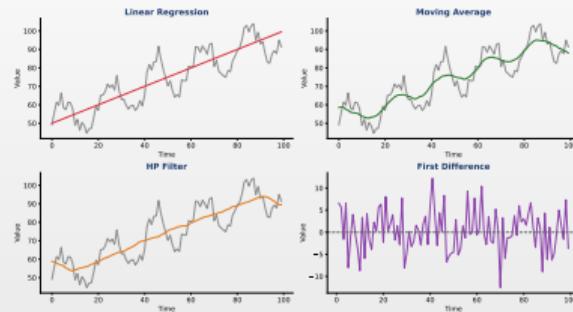
- Diferențiere:** $\Delta X_t = X_t - X_{t-1}$
 - ▶ Cea mai utilizată, elimină trend stochastic
- Regresie liniară:** $\hat{T}_t = \hat{\beta}_0 + \hat{\beta}_1 t$
- Polinomială:** polinom de ordin superior
- Filtru HP:** echilibru estimare vs netezime
- Media mobilă:** $\hat{T}_t = MA_q(X_t)$
- LOESS:** regresie polinomială locală

Alegerea depinde de

- Natura trendului**
 - ▶ Determinist vs stochastic
- Scopul analizei**
 - ▶ Prognoză vs analiză descriptivă



Metode de eliminare a trendului: comparație



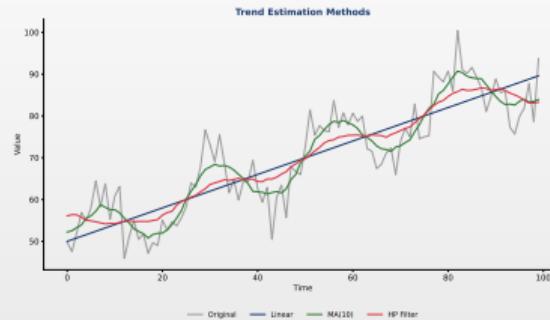
Q TSA_cho_detrending

Idee cheie

- **Metode diferite:** produc reziduuri diferite
- **Alegere după tipul de trend:** considerați obiectivele analizei



Estimarea trendului: abordări multiple



Q TSA_cho_detrending

Comparație metode

- Media mobilă:** simplă dar cu lag
- Regresie polinomială:** flexibilă, parametrică
- HP filter:** standard macroeconomic



Filtrul Hodrick-Prescott (HP)

Definiție 5 (Filtrul HP)

- Filtrul HP:** descompune X_t în trend τ_t și ciclu c_t : $X_t = \tau_t + c_t$

$$\min_{\{\tau_t\}} \left\{ \sum_{t=1}^T (X_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2 \right\}$$

Interpretare

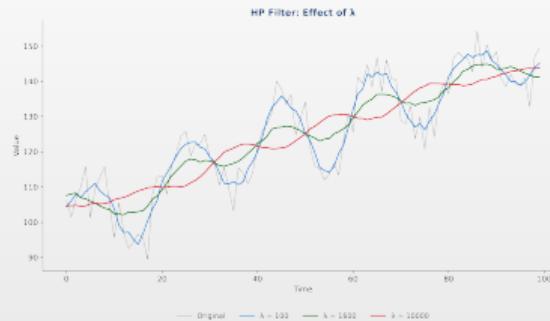
- Primul termen**
 - Ajustare la date
- Al doilea termen**
 - Penalizare netezime
- λ
 - Controlează echilibrul între fidelitate și netezime

Valori standard λ (Ravn-Uhlig)

- Anual**
 - $\lambda = 6.25$
- Trimestrial**
 - $\lambda = 1600$ (standard macroeconomic)
- Lunar**
 - $\lambda = 129600$



Filtrul HP: efectul lui λ



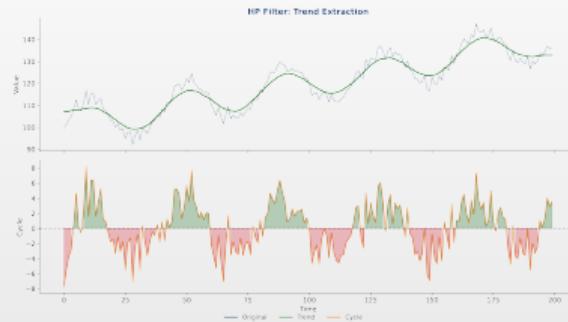
Q TSA_cho_detrending

Compromis

- λ mic:** trend flexibil — urmează datele îndeaproape
- λ mare:** trend neted — se apropie de trend liniar



Filtrul HP: extragerea ciclului de afaceri



Q TSA_cho_detrending

Aplicație

- Macroeconomie:** extragerea ciclurilor de afaceri
- Serii comune:** PIB, șomaj, inflație



Filtrul HP: limitări

Probleme cunoscute

- Instabilitate la extremități**
 - ▶ Estimările trendului nesigure la început și sfârșit
- Cicluri false**
 - ▶ Poate crea dinamici artificiale
- Alegerea λ**
 - ▶ Rezultatele sensibile la parametru

Alternative

- Filtre bandă:** Baxter-King, Christiano-Fitzgerald
 - ▶ Izolează frecvențe specifice
- Filtrul Hamilton:** bazat pe regresie
- Componente neobservate:** modele state-space

Critica lui Hamilton (2018)

- “De ce Nu Ar Trebui Să Folosiți Niciodată Filtrul HP”
 - ▶ Sugerează utilizarea regresiei pe valori întârziate



Metode de eliminare a sezonalității

Patru abordări pentru eliminarea sezonalității

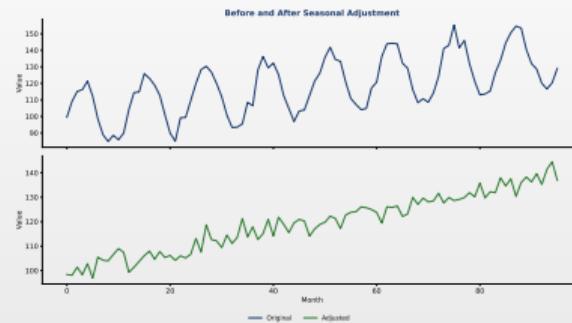
- Diferențiere sezonieră:** $\Delta_s X_t = X_t - X_{t-s}$
 - ▶ Elimină tipar periodic, simplu de aplicat
- Împărțire (multiplicativ):** $X_t^{adj} = X_t / \hat{S}_t$
- Scădere (aditiv):** $X_t^{adj} = X_t - \hat{S}_t$
- X-13ARIMA-SEATS:** standard oficial US Census Bureau
 - ▶ Metodă sofisticată, utilizată de institutele de statistică

Perioada sezonieră s

- Lunar: $s = 12$ | Trimestrial: $s = 4$



Ajustare sezonieră: vizualizare



Q [TSA_ch0_seasonal](#)

Rezultat

- Seria ajustată sezonieră: dezvăluie trendul subiacent, elimină fluctuațiile periodice



Trend determinant vs stochastic

Trend determinant

- Model:** $X_t = \beta_0 + \beta_1 t + \varepsilon_t$
- Caracteristici:**
 - ▶ Trendul este o funcție de timp
 - ▶ ε_t este staționar
- Metodă:** eliminarea trendului prin regresie

Trend stochastic

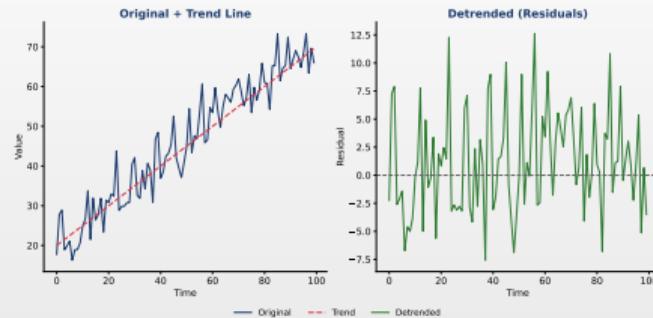
- Model:** $X_t = X_{t-1} + \varepsilon_t$
- Caracteristici:**
 - ▶ Componentă de mers aleatoriu
 - ▶ ΔX_t este staționar
- Metodă:** eliminarea trendului prin diferențiere

Metodă greșită = probleme

- Diferențiere pe trend determinant** → supra-diferențiere
 - ▶ Introduce dependență artificială în serie
- Regresie pe trend stochastic** → regresie falsă
 - ▶ Rezultate statistice invalide



Exemplu: trend determinist



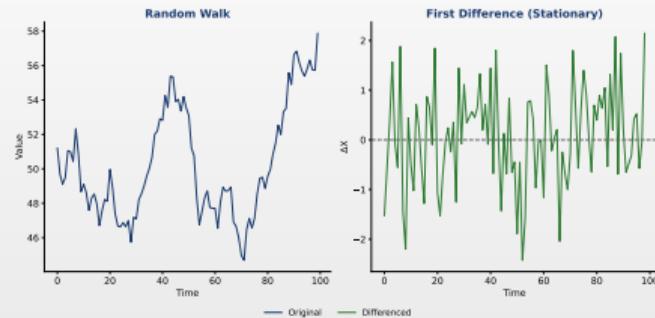
Q TSA_ch0_detrending

Cheie

- Metodă: regresie
- Rezultat: reziduuri staționare, ACF scade rapid



Exemplu: trend stochastic (mers aleatoriu)



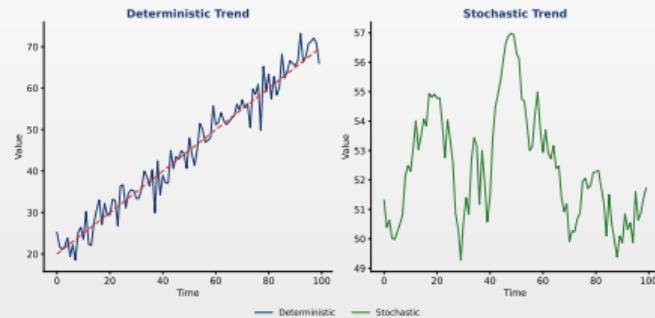
Q TSA_ch0_detrending

Cheie

- Metodă:** diferențiere
- Rezultat:** diferențele sunt staționare (zgomot alb)



Comparație alăturată



Q TSA_cho_detrending

Retineteți

- Trend deterministic:** folosiți regresie — trendul este o funcție predictibilă de timp
- Trend stochastic:** folosiți diferențiere — trendul conține o componentă aleatoare



Exercițiu AI: Gândire critică

Prompt de testat în ChatGPT / Claude / Copilot

"Folosind yfinance, descarcă datele AAPL. Vreau să văd ce tipare există și să prezic unde va fi prețul anul viitor. Fă-l să arate profesional cu grafice."

Exercițiu:

1. Rulați prompt-ul într-un LLM la alegere și analizați critic răspunsul.
2. Ce tip de descompunere alege modelul? E corect? Justificați.
3. Cum evaluează calitatea prognozei? Metrica e calculată corect?
4. Verifică reziduurile — prezintă structură neexplicată?
5. Rescrieți analiza corect și comparați cu un benchmark sezonier naiv.

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



Rezumat

Ce am învățat în acest capitol

- Definiția și Caracteristicile Seriei de Timp
 - ▶ Secvență de observații ordonate temporal cu dependență
- Descompunere (Aditivă vs multiplicativă)
 - ▶ Componente: Trend-Ciclu + Sezonier + Reziduu
- Metode de Netezire Exponențială
 - ▶ SES (nivel), Holt (+ trend), Holt-Winters (+ sezonalitate), ETS
- Evaluarea și Validarea Prognozei
 - ▶ Metrici: MAE, RMSE, MAPE; Cross-Validation cu origine mobilă

Idee cheie

- Înțelegeți Înainte de a Modela:
 - ▶ Vizualizați și descompuneți datele mai întâi
 - ▶ Alegeți aditiv vs multiplicativ în funcție de comportamentul varianței



Quiz rapid

Verificați-vă cunoștințele

1. Care este diferența între descompunerea aditivă și multiplicativă?
2. Când ar trebui să folosiți Holt-Winters în loc de SES?
3. De ce nu putem folosi CV standard k-fold pentru serii de timp?
4. Ce înseamnă $\alpha = 0.9$ în netezirea exponențială?
5. Cum distingeți între trend determinist și stochastic?



Răspunsuri quiz

Răspunsuri

1. **Aditivă vs multiplicativă:** aditivă când amplitudinea sezonieră e constantă; multiplicativă când crește cu nivelul
2. **Holt-Winters:** când datele au trend și sezonalitate; SES gestionează doar nivelul
3. **CV:** K-fold standard ignoră ordinea temporală → data leakage
4. $\alpha = 0.9$: pondere mare pe observații recente, reacționează rapid dar e mai volatilă
5. **Trend:** determinist → funcție de timp (regresie); stochastic → mers aleatoriu (diferențiere)



Ce urmează?

Capitolul 1: Procese stochastice și staționaritate

- Procese Stochastice:** fundament matematic, variabile aleatoare indexate după timp
- Staționaritate:** strictă (distribuție invariantă) vs slabă (momente invariante)
- Procese Fundamentale:** zgomot alb și mers aleatoriu → blocuri pentru ARIMA
- ACF și PACF:** instrumente pentru identificarea modelului

Întrebări?



Bibliografie I

Fundamente ale seriilor de timp

- Wold, H. (1938). *A Study in the Analysis of Stationary Time Series*, Almqvist & Wiksell.
- Hamilton, J.D. (1994). *Time Series Analysis*, Princeton University Press.
- Brockwell, P.J., & Davis, R.A. (2016). *Introduction to Time Series and Forecasting*, 3rd ed., Springer.

Descompunere și analiză exploratorie

- Cleveland, R.B., Cleveland, W.S., McRae, J.E., & Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess, *Journal of Official Statistics*, 6(1), 3–33.
- Hyndman, R.J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice*, 3rd ed., OTexts.



Bibliografie II

Netezire exponențială și fundamente ETS

- Holt, C.C. (1957/2004). Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages, *International Journal of Forecasting*, 20(1), 5–10.
- Winters, P.R. (1960). Forecasting Sales by Exponentially Weighted Moving Averages, *Management Science*, 6(3), 324–342.
- Hyndman, R.J., Koehler, A.B., Ord, J.K., & Snyder, R.D. (2008). *Forecasting with Exponential Smoothing: The State Space Approach*, Springer.

Resurse online și cod

- Quantlet: <https://quantlet.com> → Depozit de cod pentru statistică
- Quantinar: <https://quantinar.com> → Platformă de învățare metode cantitative
- GitHub TSA_ch0: https://github.com/QuantLet/TSA/tree/main/TSA_ch0



Vă Mulțumim!

Întrebări?

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



Quantlet



Quantinat