








Modelarea pieţelor financiare

Seminarul 2: Eficienţa pieţei — aplicaţii



Daniel Traian PELE
Drd. Antoaneta Amza, antoaneta.amza@csie.ase.ro

Academia de Studii Economice din Bucureşti
IDA Institute Digital Assets
Blockchain Research Center
AI4EFin Artificial Intelligence for Energy Finance
Academia Română, Institutul de Prognoză Economică
MSCA Digital Finance

De la randamente la concluzii

- 1. A1: calculează raportul variațelor și interpretează dependența.
- 2. B1: compară estimarea cu incertitudinea clasică și robustă.
- 3. C1: verifică dacă rezultatele codului susțin concluzia economică.
- La fiecare rezultat, distinge efectul estimat, incertitudinea statistică și sensul economic.
- După A1, B1 și C1: aplicații pe date noi, apoi metode pentru aprofundare.

Criterii de evaluare

- derivări corecte
- inferență robustă la heteroscedasticitate și dependență, cu decizia la un prag precizat
- evaluarea în afara eșantionului (*out-of-sample*) folosește doar informația disponibilă la data prognozei
- critica AI (*Artificial Intelligence*, inteligență artificială) documentată în `AI_ERRORS.md`

A1: de la randamente la dependență

- Începe cu notația randamentelor și ipotezele mersului aleator (*random walk*) de pe paginile următoare.
- Rezultatul este o fișă de calcul reproductibilă, urmată de o frază despre semnul dependenței.
- O identitate estimată poate diferi de identitatea teoretică deoarece estimatorii folosesc termenii de la capetele selecției și numitori diferiți.

Randamente și momente de selecție

- **Randamentul logaritmic zilnic** al unei serii cu prețurile de închidere P_t

$$r_t = \ln P_t - \ln P_{t-\mathbf{1}}, \qquad t = 1, \ldots, T$$

  - T: numărul de randamente
  - q: un orizont în zile; $r_t + r_{t-\mathbf{1}} + \cdots + r_{t-q+\mathbf{1}}$ este randamentul logaritmic pe q zile
- Media de selecție, abaterile și dispersia

$$\bar{r} = \hat{\mu} = \frac{1}{T}\sum_{t=\mathbf{1}}^{T} r_t, \qquad e_t = r_t - \bar{r}, \qquad \hat{\sigma}^{\mathbf{2}} = \frac{1}{T-1}\sum_{t=\mathbf{1}}^{T} e_t^{\mathbf{2}}$$

  - $\hat{}$ (accentul circumflex): o estimație calculată din selecție; fără accent: valoarea din populație
  - $\mathbb{E}[\cdot]$, $\mathrm{Var}(\cdot)$, $\mathrm{Cov}(\cdot,\cdot)$: media, dispersia și covarianța în populație
- **i.i.d.**: independent and identically distributed — independente și identic distribuite
  - reperul testelor clasice
  - randamentele zilnice nu sînt i.i.d.: anii calmi alternează cu episoade agitate, ca 2008–2009 și 2020 (gruparea volatilității, *volatility clustering*, Capitolul 1; vezi graficul Descrierea datelor)

Autocorelația

SE: Standard Error — eroare standard.

- **Autocorelația de ordinul** k: corelația dintre randamentul de azi și cel de acum k zile

$$\rho_k = \frac{\text{Cov}(r_t, r_{t-k})}{\text{Var}(r_t)}, \qquad \hat{\rho}_k = \frac{\sum_{t=k+1}^{T} e_t e_{t-k}}{\sum_{t=1}^{T} e_t^2}$$

  - k: lagul, în zile; $\rho_k \in [-1, 1]$; dacă $\rho_k = 0$ pentru orice $k \geq 1$, trecutul nu oferă predictibilitate liniară

- Eroarea standard a lui $\hat{\rho}_k$ cînd $\rho_k = 0$
  - randamente i.i.d.: $\text{SE}(\hat{\rho}_k) = 1/\sqrt{T}$, de aici banda i.i.d. sub ipoteza nulă (*null band*) $\pm 1{,}96/\sqrt{T}$
  - eroare standard robustă, care permite gruparea volatilității (Capitolul 1): $\text{SE}(\hat{\rho}_k) = \sqrt{\hat{\tau}_k}$

- Dispersia robustă a lui $\hat{\rho}_k$ și statistica t robustă

$$\hat{\tau}_k = \frac{\sum_{t=k+1}^{T} e_t^2 e_{t-k}^2}{\left(\sum_{t=1}^{T} e_t^2\right)^2}, \qquad \hat{\delta}_k = T\hat{\tau}_k, \qquad t_k = \frac{\hat{\rho}_k}{\sqrt{\hat{\tau}_k}}$$

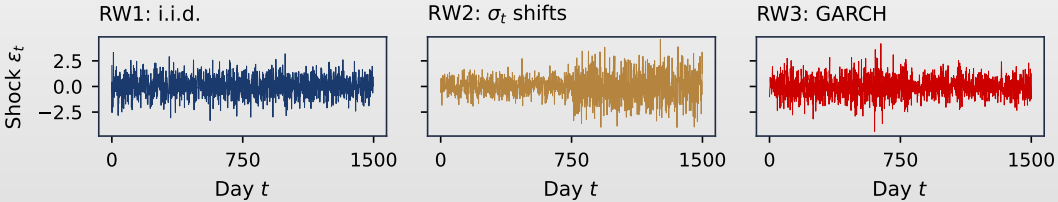
  - $\hat{\delta}_k$: raportul dintre dispersia robustă și cea calculată sub ipoteza i.i.d.
  - $\hat{\delta}_k \approx 1$ pentru randamente i.i.d.; $\hat{\delta}_k > 1$ cînd variațiile mari urmează după variații mari

Eficiența slabă

- Detalii: Capitolul 2 și Franke, Härdle & Hafner (2019)
- **Piață eficientă** Fama (1970): prețurile reflectă informația disponibilă
  - **forma slabă**: informația disponibilă cuprinde prețurile și randamentele trecute
  - **forma semi-tare**: toată informația publică (anunțuri, știri; B9)
  - **forma tare**: și informația privată
- Consecința testabilă a formei slabe
  - randamentele trecute nu permit obținerea de profituri ajustate la risc după costuri
  - în general, restricția privește partea randamentului de peste randamentul cerut: acest exces nu poate fi prevăzut pe baza randamentelor trecute
  - cu un randament cerut constant, eficiența implică faptul că $\mathbb{E}[r_t \mid r_{t-1}, r_{t-2}, \dots]$ nu depinde de trecut: o medie condiționată constantă
- **Problema ipotezei comune** Fama (1991): orice test al eficienței testează simultan și un model al randamentelor așteptate
  - o respingere poate însemna ineficiență *sau* un model greșit al randamentelor așteptate
  - exemplu (B9): un randament anormal negativ după o taxă poate fi o reacție excesivă sau o primă de risc mai mare, pe care modelul de piață nu o surprinde
- Semnul dependenței
  - $\rho_k > 0$: persistență (momentum); $\rho_k < 0$: inversare (mean reversion)

Trei variante ale mersului aleator (1/2)

- **RW1**: Random Walk 1 — mers aleator cu variații independente și identic distribuite.
- **RW2**: Random Walk 2 — mers aleator cu variații independente, fără cerința unei distribuții identice.
- **RW3**: Random Walk 3 — mers aleator cu variații necorelate, care pot fi dependente.
- **GARCH**: Generalized AutoRegressive Conditional Heteroskedasticity — model al dispersiei condiționate, explicată prin șocuri și dispersii trecute.



Trei variante ale mersului aleator (2/2)

- **Mersul aleator** al logaritmului prețului Campbell, Lo & MacKinlay (1997)

$$\ln P_t = \mu + \ln P_{t-\mathbf{1}} + \varepsilon_t, \qquad \text{deci } r_t = \mu + \varepsilon_t$$

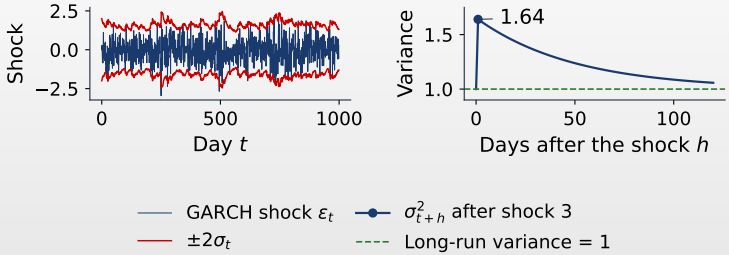
  - μ: driftul, randamentul logaritmic zilnic așteptat; ε_t: șocul zilei t, cu media 0
  - RW1: ε_t i.i.d. (panoul din stînga)
  - RW2: ε_t independente, dispersia se poate schimba (panoul din mijloc)
  - RW3: ε_t doar necorelate, $\text{Cov}(\varepsilon_t, \varepsilon_{t-k}) = 0$, iar ε_t^2 poate fi autocorelat (panoul din dreapta)
  - exemplu de RW3: modelul GARCH(1,1) Bollerslev (1986), cu $\varepsilon_t = \sigma_t z_t$ (slide-ul următor)

- Toate cele trei panouri au $\hat{\rho}_{\mathbf{1}} \approx 0$; cel din dreapta prezintă gruparea volatilității, deci ipoteza nulă relevantă pentru randamentele zilnice este RW3

GARCH(1,1), un exemplu de RW3

- $\sigma_t^2 = \omega + a\,\varepsilon_{t-1}^2 + b\,\sigma_{t-1}^2$ este **dispersia condiționată**: dispersia șocului următor, dată fiind informația disponibilă azi
  - $\omega > 0$, $a, b \geq 0$, $a + b < 1$: dispersia rămîne pozitivă și revine spre nivelul ei de lungă durată, $\omega/(1 - a - b)$
- Exemplu ilustrativ: $\omega = 0{,}02$, $a = 0{,}08$, $b = 0{,}90$
  - dispersia de lungă durată: $0{,}02/(1 - 0{,}98) = 1$
  - după un șoc $\varepsilon_{t-1} = 3$ într-o zi obișnuită ($\sigma_{t-1}^2 = 1$): $\sigma_t^2 = 0{,}02 + 0{,}08 \times 9 + 0{,}90 \times 1 = 1{,}64$
- Detalii despre GARCH: Capitolul 5

GARCH(1,1): ilustrare grafică



- Stînga: un GARCH(1,1) simulat cu $\omega = 0{,}02$, $a = 0{,}08$, $b = 0{,}90$; pe axa orizontală ziua, pe cea verticală șocul ε_t
- Banda $\pm 2\sigma_t$ se lărgește după șocuri mari și se îngustează în perioadele calme: gruparea volatilității
- Dreapta: după un șoc $\varepsilon_t = 3$, dispersia așteptată sare de la 1 la 1,64, apoi revine spre 1 cu rata $a + b = 0{,}98$ pe zi; pe axa orizontală, zilele de după șoc

RW1, RW2 și RW3 comparate (1/2)

	Ce trebuie să îndeplinească șocurile ε_t	Ce se relaxează	Mini-exemplu (graficul RW)
RW1	independente și identic distribuite (i.i.d.)	–	șocuri i.i.d. Normale cu dispersie constantă (panoul din stînga); un exemplu discret: $+1\%$ sau -1% cu probabilitatea $1/2$
RW2	independente; distribuția se poate schimba în timp	distribuția identică	$\pm 1\%$ pe zi într-un an calm, $\pm 3\%$ într-unul agitat, fixate dinainte: tot independente, dar dispersia se schimbă (panoul din mijloc)
RW3	doar necorelate: $\mathrm{Cov}(\varepsilon_t, \varepsilon_{t-k}) = 0$	independența	GARCH cu șocuri simetrice: după o zi de $\pm 3\%$, dispersia șocului următor crește, deci o mișcare mare devine mai probabilă, dar semnul ei rămîne imprevizibil (panoul din dreapta)

RW1, RW2 și RW3 comparate (2/2)

- Fiecare treaptă relaxează o cerință: RW1 $\Rightarrow$ RW2 $\Rightarrow$ RW3 (cînd dispersiile există)
- RW2 cu o dispersie fixată dinainte (an calm, an agitat) are randamente independente, dar nestaționare: alternanța perioadelor calme și agitate nu dovedește, singură, dependența
- Pentru teste (Testarea ipotezei $\mathrm{VR}(q) = 1$): $z(q)$ este construit pentru RW1, $z^*(q)$ pentru RW3; RW3 permite totuși predictibilitate neliniară a mediei

Raportul variațelor

- Pentru randamente necorelate, cu dispersia constantă σ^2, dispersia sumei pe q zile este $q\sigma^2$
  - dispersiile se adună doar dacă toate covariațele sînt zero
- **Raportul variațelor** (*variance ratio*, VR) Lo & MacKinlay (1988): împărțim dispersia randamentului pe q zile la q ori dispersia randamentului zilnic

$$\mathrm{VR}(q) = \frac{\mathrm{Var}(r_t + r_{t-1} + \cdots + r_{t-q+1})}{q\,\mathrm{Var}(r_t)}$$

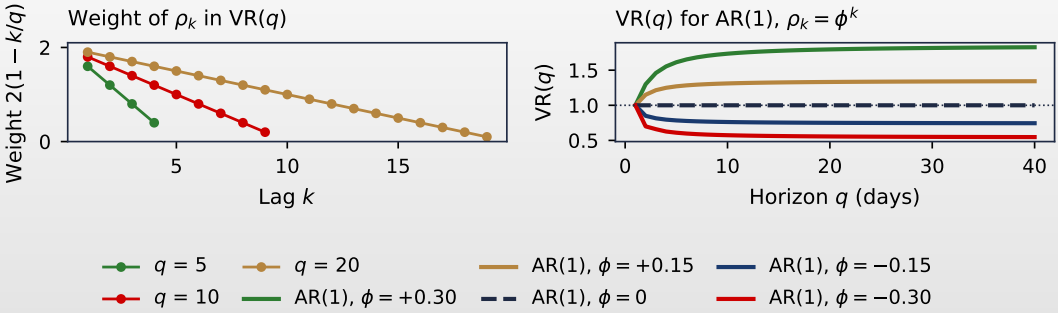
  - identitatea în populație: $\mathrm{VR}(q) = 1$ pentru randamente necorelate, cu dispersia constantă; $\mathrm{VR}(1) = 1$ întotdeauna
  - cu dispersie variabilă în timp (RW2, RW3), identitatea poate să nu mai fie valabilă; în condițiile Lo & MacKinlay (1988), estimatorul $\widehat{\mathrm{VR}}(q)$ tinde totuși la 1, iar statistica robustă $z^*(q)$ (definită mai jos) este asimptotic $N(0,1)$
  - $\mathrm{VR}(q) > 1$: persistență; dispersia randamentului pe q zile depășește reperul construit din q zile independente
  - $\mathrm{VR}(q) < 1$: inversare; variațiile din zile succesive se compensează parțial
- Cazul $q = 2$

$$\mathrm{Var}(r_t + r_{t-1}) = 2\sigma^2 + 2\mathrm{Cov}(r_t, r_{t-1}) \;\Rightarrow\; \mathrm{VR}(2) = 1 + \rho_1$$

  - exemplu: $\rho_1 = -0{,}1$ dă $\mathrm{VR}(2) = 0{,}9$; dispersia sumei pe două zile este cu 10% mai mică decît în cazul a două zile independente

Raportul variațelor ca sumă ponderată (1/2)

AR: AutoRegressive model — model autoregresiv; AR(p) are ordinul p.



Raportul variațelor ca sumă ponderată (2/2)

- În dezvoltarea dispersiei sumei pe q zile, lagului k îi corespund $2(q-k)$ termeni de covarianță

$$\mathrm{VR}(q) = 1 + 2\sum_{k=1}^{q-1}\left(1-\frac{k}{q}\right)\rho_k$$

- Stînga: ponderea $2(1-k/q)$ scade liniar cu lagul; dreapta: $\mathrm{VR}(q)$ pentru randamente AR(1)
  - AR(1), autoregresia de ordinul 1: $r_t = \phi\, r_{t-1} + u_t$, $|\phi| < 1$, u_t un șoc i.i.d.; atunci $\rho_k = \phi^k$
  - dreapta, curbe teoretice: cînd q crește, $\mathrm{VR}(q)$ se apropie de $(1+\phi)/(1-\phi)$, peste 1 pentru $\phi > 0$ și sub 1 pentru $\phi < 0$
  - exemplu: $\rho_1 = -0{,}1$, toate celelalte $\rho_k = 0$: $\mathrm{VR}(5) = 1 + 2(1-1/5)(-0{,}1) = 0{,}84$
- **Ce credeți**: dacă $\mathrm{VR}(3) = 1$ în populație, sînt randamentele necorelate?
  - **Răspuns**: nu: $\rho_1 = 0{,}05$ și $\rho_2 = -0{,}10$ dau $\mathrm{VR}(3) = 1 + \frac{4}{3}(0{,}05) + \frac{2}{3}(-0{,}10) = 1$
  - $\mathrm{VR}(q)$ testează o singură combinație ponderată a autocorelațiilor, nu fiecare autocorelație

A1 [Rezolvat]: Raportul variațelor, rînd cu rînd — enunț

- **Întrebări**
  - ce calculează $\widehat{\mathrm{VR}}(2)$?
  - cum se leagă de $\hat{\rho}_1$?
- **Date**: opt randamente zilnice în %, în ordine cronologică: $1{,}0; -0{,}5; 0{,}8; 0{,}3; -1{,}2; 0{,}6; 0{,}9; -0{,}4$; nu se folosesc alte date; se aplică estimatorul simplu, cu sumele de pătrate împărțite la numărul de termeni
- **Cerințe**
  1. Calculați media $\hat{\mu}$, abaterile $e_t = r_t - \hat{\mu}$ și dispersia pe o zi $\hat{\sigma}_1^2 = \frac{1}{8}\sum_t e_t^2$.
  2. Calculați autocorelația de ordinul 1 $\hat{\rho}_1 = \sum_{t=2}^{8} e_t e_{t-1} / \sum_{t=1}^{8} e_t^2$.
  3. Formați cele șapte sume suprapuse pe două zile $s_t = r_{t-1} + r_t$ și calculați $\hat{\sigma}_2^2 = \frac{1}{7}\sum_t (s_t - 2\hat{\mu})^2$.
  4. Calculați $\widehat{\mathrm{VR}}(2) = \hat{\sigma}_2^2 / (2\hat{\sigma}_1^2)$ și comparați-l cu $1 + \hat{\rho}_1$.
- **De raportat**: o fișă cu coloanele t, r_t, e_t, e_t^2, $e_t e_{t-1}$, s_t, $(s_t - 2\hat{\mu})^2$ și o frază despre semnul lui $\widehat{\mathrm{VR}}(2) - 1$
- Notebook: secțiunea A1

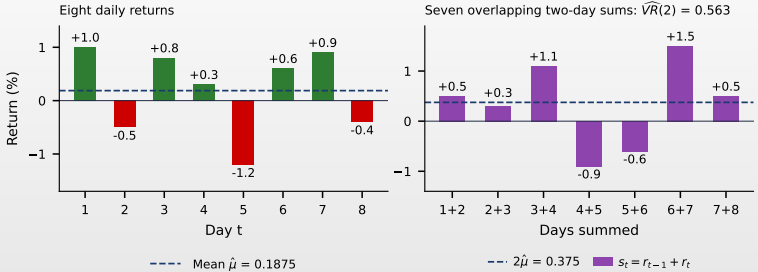
Înapoi Lista exercițiilor

A1 [Rezolvat]: Fișa de lucru

t	r_t	e_t	e_t^2	$e_t e_{t-1}$	s_t	$(s_t - 2\hat{\mu})^2$
1	$+1{,}0$	$+0{,}8125$	0,6602	–	–	–
2	$-0{,}5$	$-0{,}6875$	0,4727	$-0{,}5586$	$+0{,}5$	0,0156
3	$+0{,}8$	$+0{,}6125$	0,3752	$-0{,}4211$	$+0{,}3$	0,0056
4	$+0{,}3$	$+0{,}1125$	0,0127	$+0{,}0689$	$+1{,}1$	0,5256
5	$-1{,}2$	$-1{,}3875$	1,9252	$-0{,}1561$	$-0{,}9$	1,6256
6	$+0{,}6$	$+0{,}4125$	0,1702	$-0{,}5723$	$-0{,}6$	0,9506
7	$+0{,}9$	$+0{,}7125$	0,5077	$+0{,}2939$	$+1{,}5$	1,2656
8	$-0{,}4$	$-0{,}5875$	0,3452	$-0{,}4186$	$+0{,}5$	0,0156
sumă			4,4688	$-1{,}7639$		4,4044

- $\hat{\mu} = 1{,}5/8 = 0{,}1875$; $\hat{\sigma}_1^2 = 4{,}4688/8 = 0{,}5586$; $\hat{\rho}_1 = -1{,}7639/4{,}4688 = -0{,}395$
- $\hat{\sigma}_2^2 = 4{,}4044/7 = 0{,}6292$; $\widehat{\mathrm{VR}}(2) = 0{,}6292/(2 \times 0{,}5586) = 0{,}563$, față de $1 + \hat{\rho}_1 = 0{,}605$

A1 [Rezolvat]: Rezultate și interpretare



- $\mathrm{VR}(2) = 1 + \rho_1$ este valabil pentru momentele populației; în selecție cele două diferă prin termenii de la capete și prin numitori (7 sume, 8 pătrate): 0,563 față de 0,605 este o diferență de convenție, nu o eroare
- Interpretare: randamentele consecutive se inversează, deci sumele pe două zile variază mai puțin decît două zile separate: inversare (mean reversion)
  - cu $T = 8$, acesta este doar un exercițiu de calcul, nu o dovadă: nu se face niciun test

MFM_ch2_seminar

De ce identitatea VR din selecție are termeni de capăt

- În acest exemplu folosim $e_t = r_t - \bar{r}$, $S = \sum e_t^2$, dispersia pe o zi S/T și media pătratelor sumelor centrate pe două zile împărțită la 2.
- Dezvoltarea pătratelor sumelor dă $\widehat{VR}(2) = \frac{T}{T-1}\left[1 + \hat{\rho}_1 - \frac{e_1^2 + e_T^2}{2S}\right]$.
- Randamentele ipotetice $(2, 0, -2)$ dau $S = 8$, $\hat{\rho}_1 = 0$ și sumele pe două zile $(2, -2)$. Dispersia scalată este $4/2 = 2$, deci $\widehat{VR}(2) = 2/(8/3) = 0{,}75$.
- Identitatea dă $(3/2)[1 - 8/16] = 0{,}75$, iar $1 + \hat{\rho}_1 = 1$. Capetele și numitorii explică diferența; estimatorul Lo–MacKinlay corectează deplasarea prin numitorul m.

A1: verifică înainte să continui

- ⊡ Poate altcineva să refacă fiecare coloană din cele opt randamente, cu aceiași numitori?
- ⊡ Explică de ce două sume consecutive pe două zile au un randament zilnic comun.
- ⊡ Precizează direcția dependenței fără să afirmi că ai obținut o strategie profitabilă.

B1: estimarea are nevoie de incertitudine

- Treci de la opt randamente la selecții din piețe reale.
- Învață eroarea standard robustă și testele raportului variațelor înainte să citești rezultatele.
- Pregătește un tabel care precizează piața, orizontul, statistica, decizia și ipotezele.

De unde luăm datele

- **S&P 500**: Standard & Poor's 500 — indice al acțiunilor a 500 de companii mari din Statele Unite.
- **BET**: Bucharest Exchange Trading (index) — indicele de referință al Bursei de Valori București.
- **BET-TR**: BET Total Return (index) — indicele BET cu dividendele reinvestite.
- Fișiere: data/market; close pentru indici, adjusted_close pentru bănci (dividende și split-uri incluse).
- S&P 500: GSPC.INDX.csv; BET: BET.csv; BET-TR: BETTR.INDX.csv (prima închidere: 23 septembrie 2014).
- Banca Transilvania: TLV.RO.csv; BRD (BRD–Groupe Société Générale): BRD.RO.csv; Bitcoin: BTC-USD.CC.csv, încărcat din 2011 pentru B6.
- Celelalte serii și transformări: secțiunea Data din chapter2_seminar_notebook.ipynb.

Piețe și surse de date

- **BVB**: Bursa de Valori București.
- **WIG20**: Warszawski Indeks Giełdowy 20 — indicele celor mai mari 20 de companii de la Bursa din Varșovia.
- **BUX**: Budapesti Értéktőzsde index — indicele Bursei din Budapesta.
- **PX**: Prague Stock Exchange index (index Burzy cenných papírů Praha) — indicele Bursei din Praga.
- **BIST**: Borsa İstanbul (BIST 100 index) — Bursa din Istanbul (indicele BIST 100).
- **IPC**: Índice de Precios y Cotizaciones — indicele Bursei din Mexic.
- **FRED**: Federal Reserve Economic Data — baza de date economice a Rezervei Federale (St. Louis).

Datele folosite

Seria	Prețul folosit	Calendar	Folosită în
S&P 500	închidere	zilele de tranzacționare din Statele Unite ale Americii	A3, B1–B3, B4 (valori p furnizate), B5, C1
BET	închidere	zilele de tranzacționare BVB	A2, B1, B3, B4 (valori p furnizate), B7, C2
BET-TR și 8 indici emergenți	închidere (WIG20, BUX, PX, BIST 100, Bovespa, IPC, Nifty 50, Shanghai)	fiecare pe calendarul propriu	B7, C2
Euro Stoxx 50	închidere	calendarul propriu	B4 (valori p furnizate)
Bitcoin	închidere (USD)	în fiecare zi	B6
Banca Transilvania, BRD	închidere ajustată (dividende, split-uri)	zilele comune cu BET	B9
S&P 500 lunar	Shiller S&P Composite, închiderile S&P 500, dividendele SPY; titluri de stat americane la 3 luni (FRED, TB3MS)	medii lunare, etichete la sfîrșit de lună	B8
49 de portofolii sectoriale	Kenneth French Data Library, ponderate cu capitalizarea, lunar	sfîrșit de lună	C3

- Prețurile zilnice de închidere din datele cursului, cîte o serie pe activ, din 2000 (Bitcoin din 2011 pentru B6; în curs, Bitcoin intră în testele pe întreaga selecție din 2014) pînă la 18 sept. 2026; randamente logaritmice zilnice (Notații: randamente), pe calendarul propriu al fiecărei serii
- Regula de curățare pentru indici: doar zilele lucrătoare; o închidere egală cu cea anterioară este eliminată (de obicei este o zi nelucrătoare completată cu prețul anterior)

Pregătire: rulați și verificați

```
# secțiunile „Data” și
# „Efficiency tests” definesc
#   log_returns(name, start=None)
#   variance_ratio(r, q) -> (VR, z, z*)
#   chow_denning(r), runs_test(r), dfa_hurst(r)
sp = log_returns('sp500')    # S&P 500
bet = log_returns('bet')     # BET
print(len(sp), sp.index[0].date(),
      sp.index[-1].date(), len(bet))
print(sp.head(3).round(5))
```

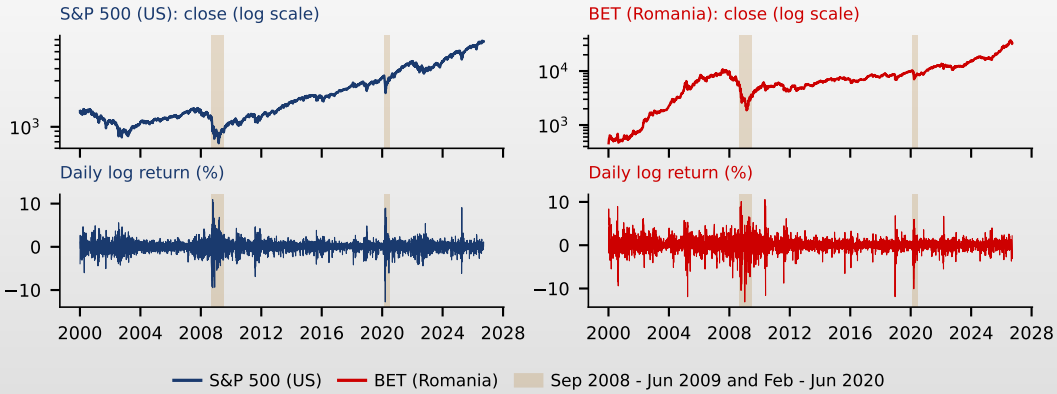
- Rezultat așteptat: 6 717 randamente S&P 500 între 2000-01-04 și 2026-09-18; 6 670 randamente BET

data	**închidere**	r_t
2000-01-03	1 455,22	–
2000-01-04	1 399,42	−0,03910
2000-01-05	1 402,11	+0,00192
2000-01-06	1 403,45	+0,00096

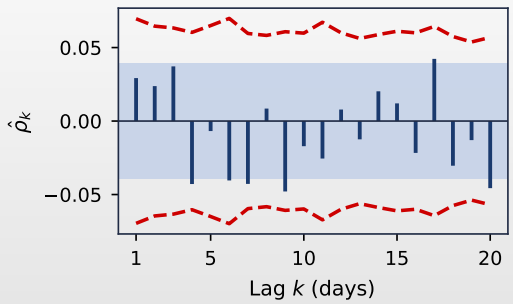
- Dacă numărul sau prima valoare diferă, verificați seria, tipul de preț (închidere ajustată sau neajustată) și regula de curățare

Descrierea datelor

US: United States — în grafice, seria indicelui S&P 500.



Două benzi sub ipoteza nulă pentru autocorelație



- Randamente GARCH simulate, $T = 2\,500$: $\rho_k = 0$ la orice lag
  - banda i.i.d. $\pm 1{,}96/\sqrt{T} = \pm 0{,}039$ (zona colorată)
  - banda robustă $\pm 1{,}96\sqrt{\hat{\tau}_k}$, în medie $\pm 0{,}061$ (linie întreruptă)
- Citirea graficului
  - autocorelațiile estimate la 6 din 20 de laguri ies din banda i.i.d.; acestea sînt respingeri false, deoarece toate valorile adevărate ale lui ρ_k sînt 0
  - nicio autocorelație estimată nu iese din banda robustă
  - aici $\hat{\delta}_k \approx 2{,}5$: eroarea standard robustă este de $\sqrt{2{,}5} \approx 1{,}6$ ori mai mare
- Explicația
  - pentru randamente i.i.d., $\mathrm{Var}(e_t e_{t-k}) = \sigma^4$ (de unde $\mathrm{SE} = 1/\sqrt{T}$); valoarea crește cînd e_t^2 și e_{t-k}^2 sînt corelate pozitiv

Estimatorul Lo–MacKinlay

- Dispersia pe o zi: $\hat{\sigma}_a^2 = \hat{\sigma}^2$, dispersia de selecție de pe slide-ul de notații pentru randamente, cu $\hat{\mu} = \bar{r}$
  - indicele a marchează estimatorul pe o zi
- Dispersia sumelor **suprapuse** pe q zile, împărțită la q

$$\hat{\sigma}_c^2(q) = \frac{1}{m} \sum_{t=q}^{T} \Big(\sum_{j=0}^{q-1} r_{t-j} - q\hat{\mu} \Big)^2, \qquad m = q(T-q+1)\Big(1 - \frac{q}{T}\Big)$$

  - indicele c marchează estimatorul din sumele suprapuse pe q zile
  - suprapuse: o nouă sumă pe q zile începe în fiecare zi, în total $T - q + 1$ sume
  - $q\hat{\mu}$: randamentul așteptat pe q zile; m: numitorul care corectează deplasarea în selecții mici
- Estimatorul

$$\widehat{\mathrm{VR}}(q) = \frac{\hat{\sigma}_c^2(q)}{\hat{\sigma}_a^2}$$

  - în A1 aplicăm aceeași idee pas cu pas, în varianta simplă: numitorii T și $T - 1$, fără corecție de deplasare

Testarea ipotezei $\mathrm{VR}(q) = 1$

- Sub RW1 (i.i.d.): statistica clasică

$$z(q) = \frac{\widehat{\mathrm{VR}}(q) - 1}{\sqrt{2(2q-1)(q-1)/(3qT)}} \approx N(0,1)$$

  - numitorul este eroarea standard i.i.d. a lui $\widehat{\mathrm{VR}}(q)$; pentru $q = 2$ este egal cu $1/\sqrt{T}$

- Sub RW3 (gruparea volatilității permisă): statistica robustă

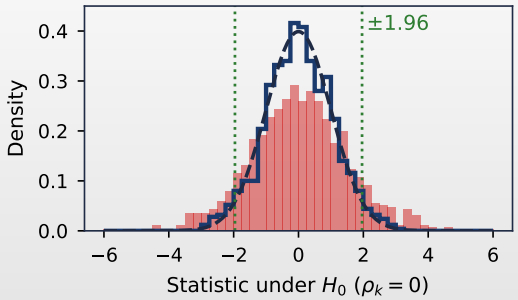
$$z^*(q) = \frac{\sqrt{T}\,(\widehat{\mathrm{VR}}(q) - 1)}{\sqrt{\hat{\theta}(q)}}, \qquad \hat{\theta}(q) = \sum_{k=1}^{q-1} \Big(\frac{2(q-k)}{q}\Big)^2 \hat{\delta}_k$$

  - $\hat{\theta}(q)$: dispersia robustă a lui $\sqrt{T}\,\widehat{\mathrm{VR}}(q)$; ponderile sînt cele ale lui ρ_k în $\mathrm{VR}(q)$, la pătrat
  - $\hat{\delta}_k = T\hat{\tau}_k$ este definit pe slide-urile de notații; dacă $|z^*(q)| > 1{,}96$, respingem RW3 la 5%

Cînd este valabilă $z^*(q)$ și un exemplu

- Cînd este valabilă statistica $z^*(q)$ Lo & MacKinlay (1988): RW3 plus trei condiții
  - momente de ordinul 4 finite
  - dependența dintre zile depărtate în timp se stinge (mixing)
  - $\mathbb{E}[\varepsilon_t^2\,\varepsilon_{t-j}\varepsilon_{t-k}] = 0$ pentru $j \neq k$: atunci $\hat{\rho}_j$ și $\hat{\rho}_k$ sînt asimptotic necorelate, deci dispersiile lor robuste $\hat{\delta}_k$ se adună în $\hat{\theta}(q)$
  - un GARCH(1,1) cu șocuri simetrice satisface condiția momentelor încrucișate; un efect de levier puternic (Capitolul 1) o poate încălca
  - momentele de ordinul 4 finite și mixing-ul cer restricții suplimentare asupra parametrilor: cu șocuri Normale, momentul de ordinul 4 este finit dacă $(a+b)^2 + 2a^2 < 1$
  - GARCH-ul ilustrativ de pe aceste slide-uri le satisface: $(0{,}08 + 0{,}90)^2 + 2 \times 0{,}08^2 = 0{,}973 < 1$
- Exemplu: $T = 2\,500$, $\widehat{\mathrm{VR}}(2) = 0{,}95$
  - $z(2) = -0{,}05 \times 50 = -2{,}5$: ipoteza nulă este respinsă la 5%
  - cu $\hat{\delta}_1 = 4$: $z^*(2) = -2{,}5/\sqrt{4} = -1{,}25$, nu respingem

Distribuțiile lui z și z^* sub ipoteza nulă



- 1 000 de selecții GARCH simulate, $T = 2\,500$: fără autocorelație, deci ipoteza nulă RW3 este adevărată
  - distribuția lui $z(2)$ este prea dispersată: abaterea standard este 1,55 în loc de 1
  - $|z(2)| > 1{,}96$ în 18,4% din selecții, în loc de 5%
  - distribuția lui $z^*(2)$ este apropiată de $N(0,1)$, cu 6,2% respingeri
- Concluzia pentru B1
  - în prezența grupării volatilității, o valoare mare a lui $|z|$ nu este suficientă; comparați $|z^*|$ cu pragul 1,96
  - Pentru $q = 2$, $z^* = z/\sqrt{\hat{\delta}_1}$: gruparea volatilității reduce precizia ca și cum selecția ar fi de $\hat{\delta}_1$ ori mai mică

Mai multe orizonturi deodată: Chow–Denning

- Dacă testăm separat patru orizonturi la 5%, probabilitatea de a respinge fals la cel puțin un orizont depășește 5%
  - cu cît sînt mai multe orizonturi, cu atît este mai mare probabilitatea unei respingeri false (testare multiplă, Testarea multiplă: FWER, Bonferroni și Holm)
- Statistica **Chow–Denning** Chow & Denning (1993): cea mai mare valoare absolută dintre statisticile robuste

$$\mathrm{CD} = \max_{q \in \{\mathbf{2},\mathbf{5},\mathbf{10},\mathbf{20}\}} |z^*(q)|$$

  - valoarea critică c din $(2\Phi(c) - 1)^{\mathbf{4}} = 0{,}95$, Φ fiind funcția de repartiție $N(0,1)$, adică $c = 2{,}49$ (nu 1,96)
  - valoarea p: $p = 1 - (2\Phi(\mathrm{CD}) - 1)^{\mathbf{4}}$
  - exponentul 4 este numărul orizonturilor: limita le tratează ca independente, deci testul este ușor conservator
- Interpretare
  - $\mathrm{CD} > 2{,}49$: testul comun respinge RW3 la 5%; cea mai mare valoare a lui $|z^*|$ indică orizontul cu abaterea cea mai pronunțată

Ce orizonturi testăm cu raportul variațelor?

- Lo & MacKinlay (1988), tabelul 1a: patru orizonturi, $q = 2, 4, 8, 16$, pentru observații săptămînale.
- Aici testăm patru orizonturi, $q = 2, 5, 10, 20$ zile: două zile, aproximativ o săptămînă, două săptămîni și o lună bursieră pentru acțiuni.
- Profilul din curs se extinde descriptiv pînă la $q = 60$; cele patru orizonturi testate aici rămîn aceleași.
- Cele patru orizonturi sînt o alegere explicită, nu o garanție de suficiență.

B1 [Rezolvat]: Raportul variațelor clasic și robust — enunț

- **Întrebare**: este respinsă ipoteza mersului aleator pentru S&P 500 și BET atunci cînd permitem gruparea volatilității?
- **Date**: S&P 500 și BET, randamente logaritmice zilnice 2000–2026, $T = 6\,717$ și $6\,670$; H_0 pentru $z(q)$: RW1 (i.i.d.); H_0 pentru $z^*(q)$: RW3 plus condițiile din Lo & MacKinlay (1988)
- **Cerințe**
  1. Pentru S&P 500 și $q = 2$, calculați $\hat{\sigma}_a^2$, $\hat{\sigma}_c^2(2)$, $\widehat{\mathrm{VR}}(2)$, $\hat{\delta}_1$, $z(2)$ și $z^*(2)$ cu formulele Lo–MacKinlay (Estimatorul Lo–MacKinlay).
  2. Pentru ambele piețe și $q = 2, 5, 10, 20$, calculați $\widehat{\mathrm{VR}}(q)$, $z(q)$, $z^*(q)$ și valorile p bilaterale cu `variance_ratio(r, q)`.
  3. Calculați statistica Chow–Denning $\max_q |z^*(q)|$ și decideți la 5% cu valoarea critică 2,49; raportați valoarea p a statisticii (Mai multe orizonturi deodată: Chow–Denning).
- **De raportat**: un tabel cu un rînd pentru fiecare piață și q ($\widehat{\mathrm{VR}}$, z, z^*, p), decizia testului Chow–Denning (CD) și o frază despre motivul pentru care $|z| > |z^*|$
- Notebook: secțiunea B1

◂ Înapoi ▸ Lista exercițiilor

B1 [Rezolvat]: S&P 500, $q = 2$, pas cu pas

- Randamente în %: $T = 6\,717$, $\hat{\mu} = 0{,}0247$; dispersia pe o zi $\hat{\sigma}_a^2 = \frac{1}{T-1}\sum_t e_t^2 = 1{,}4708$
- 6 716 sume suprapuse $r_t + r_{t-1} - 2\hat{\mu}$; numitorul $m = q(T - q + 1)(1 - q/T) = 13\,428$; $\hat{\sigma}_c^2(2) = 1{,}3246$
- $\widehat{\mathrm{VR}}(2) = 1{,}3246/1{,}4708 = 0{,}9006$
- Clasic: $\mathrm{SE} = \sqrt{2(2q-1)(q-1)/(3qT)} = \sqrt{1/T} = 0{,}0122$, deci $z(2) = (-0{,}0994)/0{,}0122 = -8{,}15$
- Robust: $\hat{\delta}_1 = T\sum_t e_t^2 e_{t-1}^2/(\sum_t e_t^2)^2 = 4{,}967$ (circa 1 pentru randamente i.i.d.); pentru $q = 2$, $\hat{\theta}(2) = \hat{\delta}_1$
  - $z^*(2) = \sqrt{T}\,(\widehat{\mathrm{VR}}(2) - 1)/\sqrt{\hat{\theta}(2)} = 81{,}96 \times (-0{,}0994)/\sqrt{4{,}967} = -3{,}66$, $p = 0{,}0003$

MFM_ch2_seminar

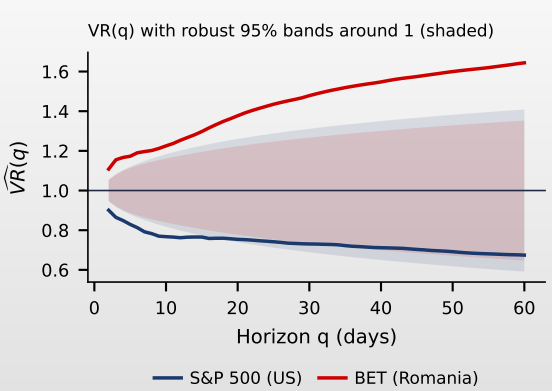
B1 [Rezolvat]: Ambele piețe, patru orizonturi

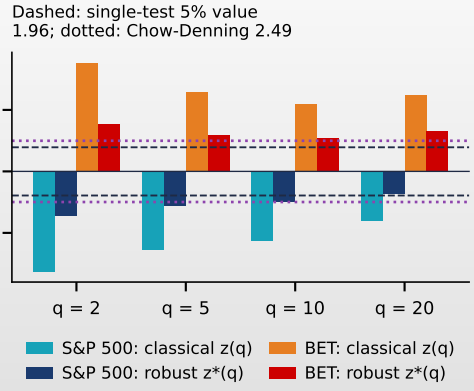
	q	$\widehat{\mathrm{VR}}(q)$	$z(q)$	p (RW1)	$z^*(q)$	p (RW3)
S&P 500	2	0,901	−8,15	$< 0{,}0001$	−3,66	0,0003
	5	0,829	−6,39	$< 0{,}0001$	−2,78	0,0055
	10	0,767	−5,65	$< 0{,}0001$	−2,49	0,013
	20	0,754	−4,06	$< 0{,}0001$	−1,85	0,064
BET	2	1,108	8,83	$< 0{,}0001$	3,84	0,0001
	5	1,173	6,44	$< 0{,}0001$	2,98	0,0029
	10	1,225	5,44	$< 0{,}0001$	2,73	0,0063
	20	1,377	6,19	$< 0{,}0001$	3,31	0,0009

- Valori p bilaterale: $p = 2(1 - \Phi(|z|))$, pentru $z(q)$ sub RW1 și pentru $z^*(q)$ sub RW3
- Chow–Denning: S&P 500 3,66 ($p = 0{,}0010$), BET 3,84 ($p = 0{,}0005$), ambele peste 2,49: testul comun respinge RW3 pentru ambele piețe

MFM_ch2_seminar

B1 [Rezolvat]: Grafic





MFM_ch2_seminar

B1 [Rezolvat]: Interpretare

- Interpretare: inversare la S&P 500 ($\widehat{\text{VR}} < 1$), semnificativă cu statistica robustă pînă la $q = 10$; persistență la BET ($\widehat{\text{VR}} > 1$) la toate cele patru orizonturi
- De ce $|z| > |z^*|$: din cauza grupării volatilității, $\hat{\delta}_1 = 4{,}967$ pentru S&P 500, deci eroarea standard robustă la $q = 2$ este de $\sqrt{\hat{\delta}_1} \approx 2{,}2$ ori cea i.i.d.; statistica robustă are aproximativ jumătate din valoarea absolută ($|z|/|z^*|$ între 1,87 și 2,30)
- Respingerea ipotezei nule nu oferă automat o regulă profitabilă de tranzacționare; trebuie evaluate costurile, riscul și stabilitatea rezultatului (B2)

MFM_ch2_seminar

B1: verifică înainte să continui

- Explică de ce estimarea punctuală poate rămîne aceeași, iar decizia de respingere se poate schimba.
- Pentru fiecare concluzie, numește ipoteza nulă și orizontul testat.
- Nerespingerea nu dovedește independența și nici eficiența pieței.

C1: aplică metoda pe cont propriu

- Ai nevoie de fișa din A1 și de testul robust din B1.
- Verifică singur fiecare afirmație a răspunsului AI înainte de a consulta rezolvarea.
- Pentru fiecare afirmație, identifică secvența de cod care o calculează și ipoteza care o susține.

C1 [Propus]: Critica unui răspuns AI — răspunsul

AI: Artificial Intelligence — inteligență artificială.

- **Cererea** trimisă de un student unui asistent AI: „Testează cu raportul variațelor dacă S&P 500 (închideri zilnice, 2000–2026) urmează un mers aleator. Dă-mi cod Python și interpretează rezultatul”

```
r = np.log(sp).diff().dropna()      # sp: închiderile brute S&P 500; verificați curățarea față de B1
T = len(r)
def vr(q):
    rq = r.rolling(q).sum().dropna()  # randamente suprapuse pe q zile
    return rq.var() / (q * r.var())
q = 5
z = (vr(q) - 1) / np.sqrt(2 * (2*q - 1) * (q - 1) / (3 * q * T))
p = 2 * stats.norm.sf(abs(z))
```

- **Interpretarea AI-ului**: (a) „$\mathrm{VR}(2) = 0{,}900$, $\mathrm{VR}(5) = 0{,}829$, $z(5) = -6{,}41$, $p < 0{,}001$: mersul aleator este respins”; (b) „deoarece $\mathrm{VR}(2) = 1 + 2\rho_1$, $\rho_1 \approx -0{,}050$”; (c) „$\mathrm{VR} < 1$ înseamnă momentum: cumpărați S&P 500 după zilele de creștere”

◂ Înapoi ▸ Lista exercițiilor

C1 [Propus]: Critica unui răspuns AI — enunț

- **Întrebare**: ce afirmații ale răspunsului AI sînt susținute efectiv de codul lui?
- **Date**: codul și interpretarea de pe slide-ul anterior; închiderile brute S&P 500 din perioada B1; aplicați regula de curățare a indicilor (slide-ul Datele folosite) pentru verificările numerice
- **Cerințe**
  1. Verificați patru puncte într-un tabel (linia, afirmația, verdictul (susținută / eroare / diferență de convenție / necalculată de codul afișat), verificarea): convenția estimatorului, ipoteza nulă și eroarea ei standard, identitatea VR, interpretarea economică.
  2. Precizați ce cifre evaluează efectiv liniile afișate; definirea unei funcții nu o execută.
  3. Efectuați două verificări numerice: $\hat{\rho}_1$ calculat direct și $z^*(5)$ din `variance_ratio`.
  4. Scrieți interpretarea corectată în două fraze.
- **De raportat**: tabelul de verificare, cele două verificări, interpretarea corectată; notați fiecare eroare ca în `AI_ERRORS.md`
- **Model**: A1, B1 [Rezolvat]
- Notebook: secțiunea B1

Ce trebuie să poți face după aceste exerciții

- Poți calcula un raport al variațelor, alege o procedură de inferență și formula o concluzie economică prudentă.
- Păstrează fișa A1, comparația B1 și verificarea C1 ca dovezi ale acestor competențe.

Alte aplicații și teme de aprofundare

- ⊡ Aplicații: verifică o altă selecție
- ⊡ Aplicații: efecte de calendar
- ⊡ Pentru aprofundare: memorie lungă
- ⊡ Pentru aprofundare: alte teste și piețe
- ⊡ Pentru aprofundare: prognoză
- ⊡ Pentru aprofundare: un studiu de eveniment
- ⊡ Pentru aprofundare: proiecte de cercetare

B2 [Propus]: Inversarea S&P 500 în afara crizelor (1/2)

- **Întrebare**: rămîne inversarea S&P 500 din B1 dacă eliminăm cele două perioade de criză?
- **Date**
  - S&P 500, randamente logaritmice zilnice 2000–2026, $T = 6\,717$; zilele de criză, fixate dinainte: 1 sept. 2008 – 30 iun. 2009 și 15 febr. – 30 iun. 2020
  - Eliminarea unei perioade lasă segmente contigue. O pereche (r_t, r_{t-1}) cu $t = 1$ iul. 2009 și $t - 1 = 29$ aug. 2008 ar lega două zile aflate la zece luni distanță, deci nu este permisă; la fel, nicio sumă pe 5 zile nu poate traversa un gol

înapoi Lista exercițiilor

B2 [Propus]: Inversarea S&P 500 în afara crizelor (2/2)

Cerințe

1. Pe întreaga selecție, calculați $\hat{\rho}_1$, t i.i.d. $= \hat{\rho}_1\sqrt{T}$, t robust $= \hat{\rho}_1/\sqrt{\hat{\tau}_1}$ și $z^*(5)$.
2. Eliminați zilele de criză și repetați în segmentele rămase, cu media comună și suma comună a abaterilor pătratice. T rămîne numărul randamentelor păstrate; în numitorul VR folosiți numărul sumelor valide din segmente. Raportați randamentele păstrate, perechile valide și sumele valide pe 5 zile.
3. Decideți la 5% cu t robust: SE $\sqrt{\hat{\tau}_1}$ este derivată sub $\rho_1 = 0$, deci $\pm 1{,}96\sqrt{\hat{\tau}_1}$ este o bandă de test în jurul lui zero, nu un interval în jurul lui $\hat{\rho}_1$.
4. Raportați intervale de 95% pentru ρ_1 în ambele selecții cu SE HAC (*Heteroskedasticity and Autocorrelation Consistent*, erori standard robuste la heteroscedasticitate și autocorelație) a estimatorului (Newey–West pe funcția de influență, Capitolul 2), valabilă și cînd $\rho_1 \neq 0$.

B2 [Propus]: Este inversarea S&P 500 un efect al crizelor? — enunț

- **De raportat**
  - un tabel cu două rînduri și o frază
  - cum se modifică estimația?
  - este legitimă eliminarea zilelor de criză?
- **Model**: B1 [Rezolvat]
- Notebook: secțiunea B2

Efecte de calendar

- Folosește B1 ca să separi estimarea de incertitudinea ei; aici introducem regresia și corecțiile pentru testări multiple.
- Compară mediile pe zile ale săptămînii și arată dacă rezultatul se menține după corecția pentru familia de teste.
- Înapoi la alegerea modulelor.
- **TOM**: Turn of the Month — schimbarea lunii (efectul schimbării lunii).
- **OLS**: Ordinary Least Squares — metoda celor mai mici pătrate.

Efectele de calendar ca regresie

- **Schimbarea lunii** (TOM), fereastra din Lakonishok & Smidt (1988): ultima zi de tranzacționare a lunii și primele trei ale lunii următoare

$$r_t = a + b\,\mathrm{TOM}_t + u_t$$

  - efectul lunar a fost documentat întîi de Ariel (1987), pe o fereastră mai lungă (ultima zi și prima jumătate a lunii)
  - $\mathrm{TOM}_t = 1$ într-o zi de schimbare a lunii, 0 în caz contrar (variabilă indicator); u_t: reziduul
- Cu un regresor 0/1, OLS compară două medii

$$\hat{a} = \bar{r}_{\mathbf{alte}}, \qquad \hat{b} = \bar{r}_{\mathbf{TOM}} - \bar{r}_{\mathbf{alte}}$$

  - H_0: $b = 0$, fără efect TOM; $t = \hat{b}/\,\mathrm{SE}(\hat{b})$, SE: eroarea standard
- Celelalte efecte din B4
  - **ianuarie**: aceeași regresie, cu randamente lunare și o variabilă indicator pentru luna ianuarie
  - **ziua săptămînii**: cinci indicatori; un test Wald al egalității mediilor, $W = (\mathrm{R}\hat{\boldsymbol{\beta}})'(\mathrm{R}\hat{\mathrm{V}}\mathrm{R}')^{-1}\mathrm{R}\hat{\boldsymbol{\beta}} \approx \chi^2_4$
  - $\mathrm{R}\hat{\boldsymbol{\beta}}$: cele patru diferențe față de media zilei de luni; $\hat{\mathrm{V}}$: covarianța lor Newey–West (slide-ul următor)
  - χ^2_m: distribuția hi-pătrat cu m grade de libertate, adică distribuția sumei pătratelor a m variabile $N(0,1)$ independente; aici $m = 4$
  - respingem egalitatea mediilor la 5% cînd $W > 9{,}49$, cuantila de 95% a distribuției χ^2_4
  - cu o singură restricție, $W = t^2$, iar pragul devine $1{,}96^2 = 3{,}84$

Erorile standard Newey–West (HAC)

- **HAC** (erori robuste la heteroscedasticitate și autocorelație) Newey & West (1987): estimăm dispersia lui $\hat{\beta}$ permițînd reziduuri cu dispersie variabilă și corelații pe L laguri

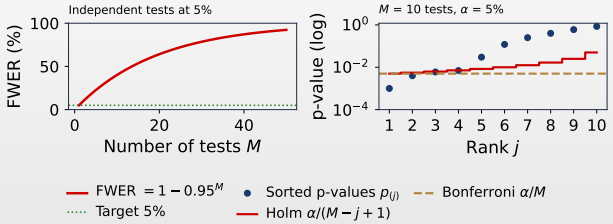
$$\widehat{\mathrm{Var}}(\hat{\beta}) = (\mathrm{X}'\mathrm{X})^{-1}\,\hat{\mathrm{S}}\,(\mathrm{X}'\mathrm{X})^{-1}, \qquad \hat{\mathrm{S}} = \sum_{j=-L}^{L}\Big(1-\frac{|j|}{L+1}\Big)\sum_{t} \mathrm{x}_t\hat{u}_t\hat{u}_{t-j}\mathrm{x}'_{t-j}$$

  - X: matricea $T \times 2$ cu rîndurile $\mathrm{x}'_t = (1, \mathrm{TOM}_t)$; $\hat{u}_t$: reziduurile OLS
  - ponderile $1 - |j|/(L+1)$ (Bartlett): sînt 1 la $j = 0$ și scad liniar pînă la 0 după lagul L
  - $L = 0$: erorile White (robuste la heteroscedasticitate); $L = 5$ în B3
  - L se fixează înainte de a vedea rezultatele: 5 pentru date zilnice (B3), 3 sau 12 pentru date lunare (B4, B8)

- Erorile OLS clasice
  - $\widehat{\mathrm{Var}}(\hat{\beta}) = \hat{\sigma}_u^2(\mathrm{X}'\mathrm{X})^{-1}$: dispersie constantă, fără autocorelație
  - erorile standard HAC și OLS diferă cînd $\mathrm{x}_t\hat{u}_t$ este heteroscedastic sau autocorelat; eroarea HAC este mai mare la autocorelație pozitivă și poate fi mai mică la autocorelație negativă
  - estimația punctuală rămîne aceeași, se schimbă doar incertitudinea

Testarea multiplă: FWER, Bonferroni și Holm

FWER: Family-Wise Error Rate — probabilitatea cel puțin unei respingeri false în familia de teste.



- **FWER** pentru M teste independente la 5%: $1 - 0{,}95^M$ (stînga), $M = 10$: 40,1%
  - **Bonferroni**: respingem ipoteza nulă a testului i dacă $p_i < \alpha/M$; p_i: valoarea p a testului, α: nivelul dorit pentru FWER
  - **Holm** Holm (1979): ordonăm $p_{(1)} \leq \cdots \leq p_{(M)}$ și respingem succesiv cît timp $p_{(j)} < \alpha/(M - j + 1)$; ne oprim la prima ipoteză pe care nu o respingem
  - ambele păstrează FWER $\leq \alpha$; Holm respinge cel puțin la fel de des (dreapta: 4 respingeri față de 2)
  - în finanțe: sute de anomalii publicate au fost testate pe aceleași randamente Harvey, Liu & Zhu (2016); o procedură secvențială care ține cont de dependența dintre teste: Romano & Wolf (2005)

B3 [Rezolvat]: Efectul schimbării lunii cu erori HAC — enunț

- **Întrebare**: sînt randamentele mai mari în zilele de schimbare a lunii decît în celelalte zile?
- **Date**
  - S&P 500 și BET, randamente logaritmice zilnice 2000–2026 în puncte de bază (bp; 1 bp = 0,01%), $r_t = 10\,000 \ln(P_t / P_{t-1})$, fiecare pe calendarul propriu de tranzacționare; ultima lună incompletă (septembrie 2026) este eliminată
  - TOM_t: indicatorul zilelor de schimbare a lunii din Efectele de calendar ca regresie; exemplu: 31 iul. și 3, 4, 5 aug. 2026 sînt zile TOM
- **Cerințe**
  1. Construiți TOM_t din calendarul de tranzacționare (rangul zilei în luna ei, de la început și de la sfîrșit).
  2. Estimați $r_t = a + b\,\text{TOM}_t + u_t$ prin OLS; $\hat{a}$ este media în celelalte zile, $\hat{a} + \hat{b}$ media în zilele TOM.
  3. Calculați erorile standard OLS și HAC Newey–West ale lui $\hat{b}$ cu 5 laguri Newey & West (1987), ambele statistici t și valoarea p HAC.
  4. Transformați $\hat{b}$ într-un randament logaritmic suplimentar aproximativ pe lună (4 zile TOM), înainte de costuri.
- **De raportat**: un tabel (proporția zilelor TOM, $\hat{a}$, $\hat{b}$, ambele SE și t), un grafic cu intervale de 95%, o frază despre motivul diferenței dintre OLS și HAC
- Notebook: secțiunea B3

◂ Înapoi ▸ Lista exercițiilor

B3 [Rezolvat]: Matricea regresorilor

data	r_t (bp)	1	TOM_t
2026-07-27	+1,6	1	0
2026-07-28	+21,0	1	0
2026-07-29	−152,8	1	0
2026-07-30	+164,7	1	0
2026-07-31	+69,8	1	1
2026-08-03	+146,8	1	1
2026-08-04	+177,4	1	1
2026-08-05	−16,8	1	1
2026-08-06	−17,6	1	0
2026-08-07	+61,7	1	0

- Rîndurile S&P 500 în jurul schimbării de lună iulie/august 2026: regresorii sînt o coloană de valori 1 și indicatorul TOM
- OLS: $\hat{b} = \bar{r}_{\mathbf{TOM}} - \bar{r}_{\mathbf{alte}}$; S&P 500 $5{,}1 - 1{,}9 \approx 3{,}2$ bp; BET $20{,}4 - 3{,}1 \approx 17{,}3$ bp
- HAC: dispersia Newey–West din Erorile standard Newey–West (HAC) cu $L = 5$, adică ponderile $1 - |j|/6$
- În cod: `sm.OLS(r, X).fit(cov_type='HAC', cov_kwds={'maxlags': 5})`

MFM_ch2_calendar_anomalies

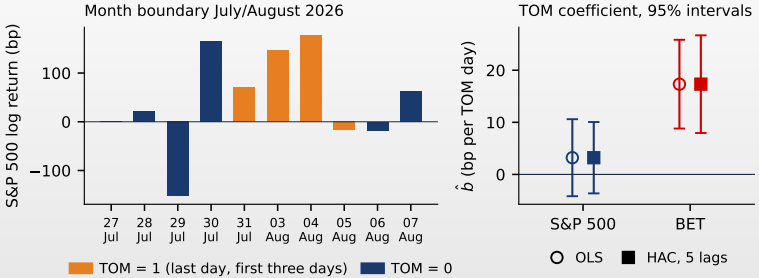
B3 [Rezolvat]: Estimări

	Proporția TOM	$\hat{a}$ (bp)	$\hat{b}$ (bp)	SE_{OLS}	t_{OLS}	SE_{HAC}	t_{HAC} (p)
S&P 500	19,1%	1,9	3,2	3,8	0,85	3,5	0,92 (0,36)
BET	19,2%	3,1	17,3	4,3	3,99	4,8	3,62 (0,0003)

- BET: intervalul HAC de 95% $[8{,}0; 26{,}7]$ bp; SE HAC depășește SE OLS deoarece $x_t \hat{u}_t$ este autocorelat pozitiv: cele patru zile TOM sînt consecutive, deci în interiorul fiecărui bloc produsul preia autocorelația randamentelor ($\hat{\rho}_1 = 0{,}109$); t scade de la 3,99 la 3,62, iar ipoteza nulă rămîne respinsă
- S&P 500: $[-3{,}6; 10{,}1]$ bp conține 0; aici SE HAC este sub SE OLS, deoarece autocorelația este negativă ($\hat{\rho}_1 = -0{,}099$)
- $4\hat{b} \approx 69$ bp (BET): un randament logaritmic lunar suplimentar aproximativ față de media celorlalte zile, înainte de costuri, nu randamentul unei strategii

MFM_ch2_calendar_anomalies

B3 [Rezolvat]: Rezultate și interpretare



- Graficul, panoul din dreapta: intervalul BET nu conține 0 cu niciuna dintre erorile standard; intervalul S&P 500 conține 0 cu ambele
- Interpretare: un singur efect, pe o singură piață, a fost ales dintre multe posibile: corectați pentru efectele testate (B4) înainte de a căuta un mecanism (intrările de salarii și pensii)

MFM_ch2_calendar_anomalies

B4 [Propus]: Cincisprezece teste de calendar și Holm — enunț (1/2)

- **Întrebare**: care dintre cele cincisprezece efecte de calendar rămîn semnificative după o corecție pentru testare multiplă?
- **Date**: cele 15 valori p din tabelul de mai jos: 5 piețe × 3 efecte (ziua săptămînii: testul Wald al egalității mediilor, HAC 5 laguri; ianuarie: randamente lunare, HAC 3 laguri; TOM: ca în B3), 2000–2026

Piața	Efectul	p	Piața	Efectul	p
S&P 500	ziua săptămînii (medii egale)	0,826	Euro Stoxx 50	schimbarea lunii	0,227
S&P 500	ianuarie	0,426	WIG20 (Polonia)	ziua săptămînii (medii egale)	0,227
S&P 500	schimbarea lunii	0,358	WIG20 (Polonia)	ianuarie	0,686
BET (România)	ziua săptămînii (medii egale)	0,305	WIG20 (Polonia)	schimbarea lunii	0,085
BET (România)	ianuarie	0,137	BUX (Ungaria)	ziua săptămînii (medii egale)	0,933
BET (România)	schimbarea lunii	0,0003	BUX (Ungaria)	ianuarie	0,024
Euro Stoxx 50	ziua săptămînii (medii egale)	0,632	BUX (Ungaria)	schimbarea lunii	0,0003
Euro Stoxx 50	ianuarie	0,774			

◂ Înapoi ▸ Lista exercițiilor

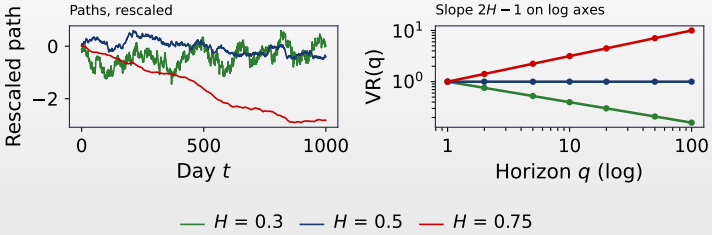
B4 [Propus]: Cincisprezece teste de calendar și Holm — enunț (2/2)

- **Cerințe**
  1. Numărați respingerile fără corecție la 5% și calculați FWER pentru 15 teste independente.
  2. Aplicați Bonferroni: respingeți dacă $p_i < 0{,}05/15$.
  3. Aplicați Holm: ordonați valorile p, comparați $p_{(j)}$ cu $0{,}05/(16-j)$ și opriți-vă la prima nerespingere.
- **De raportat**: tabelul ordonat cu ranguri, praguri și trei coloane de decizie (fără corecție, Bonferroni, Holm) și o frază despre ianuarie la BUX
- **Model**: B3 [Rezolvat]
- Notebook: secțiunea B4

Memorie lungă

- După A1 și B1, deosebește un exponent Hurst implicit de dovezile de memorie lungă.
- Compară scalarea pe orizonturi cu benzile de referință sub ipoteza nulă; ferestrele suprapuse descriu variația, dar nu oferă teste independente.
- Înapoi la alegerea modulelor.

Exponentul Hurst (1/2)

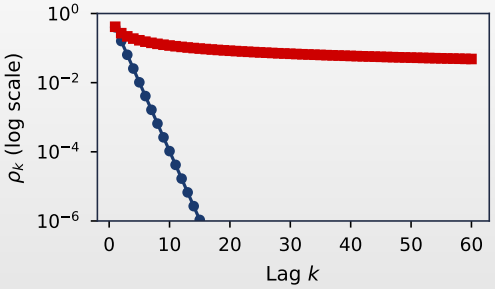


Exponentul Hurst (2/2)

fGn: fractional Gaussian noise — zgomot gaussian fracționar.

- **Exponentul Hurst** $H \in (0,1)$, introdus în studiul hidrologiei Nilului Hurst (1951), descrie cum crește dispersia seriei cumulate odată cu orizontul
  - în Hurst (1951), estimarea folosește amplitudinea rescalată R/S (*rescaled range*): amplitudinea abaterilor cumulate de la medie, împărțită la abaterea standard; acest raport crește ca n^H odată cu lungimea n a selecției
  - creșteri autosimilare (zgomot gaussian fracționar, fGn: suma pe q zile are distribuția unei zile, scalată cu q^H): $\mathrm{Var}(r_t + \cdots + r_{t-q+\mathbf{1}}) = q^{\mathbf{2}H}\mathrm{Var}(r_t)$, deci $\mathrm{VR}(q) = q^{\mathbf{2}H-\mathbf{1}}$
  - $H = 0{,}5$: fără memorie; $H > 0{,}5$: traiectorii netede, cu tendințe; $H < 0{,}5$: antipersistență, traiectorii rugoase (pagina anterioară, stînga)
- Exponentul implicit $H(q)$: valoarea H care rezolvă $\mathrm{VR}(q) = q^{\mathbf{2}H-\mathbf{1}}$ la un singur orizont q; formula explicită pentru $H(q)$ se deduce în A2

Memoria scurtă și memoria lungă



AR(1), $\phi = 0.4$: $\rho_k = \phi^k$

fGn, $H = 0.75$: $\rho_k \approx H(2H-1)k^{2H-2}$

- **Memorie scurtă**: AR(1) (VR ca sumă ponderată)

$$r_t = \phi\, r_{t-1} + u_t, \qquad \rho_k = \phi^k$$

  - scădere geometrică: o dreaptă pe scala logaritmică (albastru)
  - $\mathrm{VR}(q) \to (1+\phi)/(1-\phi)$, mărginit, deci $H(q) \to 0{,}5$ cînd q crește

- **Memorie lungă**, fGn cu $H > 0{,}5$

$$\rho_k = \tfrac{1}{2}\left(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}\right)$$

  - scădere hiperbolică $\rho_k \approx H(2H-1)k^{2H-2}$ (roșu): la $k = 20$, 0,084 față de 10^{-8} pentru AR(1)

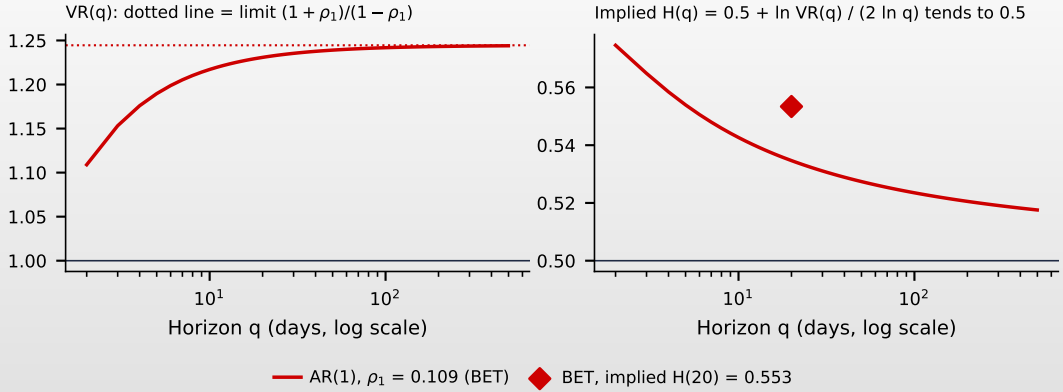
- Un singur orizont q nu permite deosebirea celor două tipuri de memorie; comparați $H(q)$ pentru mai multe valori ale lui q (A2, A3)
  - un test formal: statistica V din Lo (1991), un R/S calculat cu o abatere standard pe termen lung (HAC)

A2 [Rezolvat]: Exponentul Hurst implicit pentru BET — enunț

- **Întrebare**: implică valoarea $\widehat{\mathrm{VR}}(20) > 1$ a BET existența memoriei lungi?
- **Date**: BET, randamente logaritmice zilnice 2000–2026 (prețuri de închidere): $\widehat{\mathrm{VR}}(20) = 1{,}377$, $\hat{\rho}_1 = 0{,}109$
- **Cerințe**
  1. Logaritmați relația zgomotului gaussian fracționar $\mathrm{VR}(q) = q^{2H-1}$ (Exponentul Hurst) și exprimați $H(q)$ implicit.
  2. Calculați $H(20)$ implicit pentru BET.
  3. Calculați VR(20) și $H(20)$ pentru un AR(1) cu memorie scurtă, cu $\rho_k = 0{,}109^k$, folosind suma ponderată din Raportul variațelor ca sumă ponderată.
- **De raportat**: formula, cele două valori $H(20)$ cu trei zecimale și dacă un singur orizont poate identifica memoria lungă
- Notebook: secțiunea A2

◂ Înapoi ▸ Lista exercițiilor

A2 [Rezolvat]: Grafic



MFM_ch2_seminar

A2 [Rezolvat]: Interpretare

- (1) $\ln \mathrm{VR}(q) = (2H-1)\ln q$, deci $H(q) = 0{,}5 + \ln \mathrm{VR}(q)/(2\ln q)$
- (2) $\ln 1{,}377 = 0{,}3199$, $2\ln 20 = 5{,}991$: $H(20) = 0{,}5 + 0{,}3199/5{,}991 = 0{,}553$
- (3) AR(1): $\mathrm{VR}(20) = 1{,}231$, $H(20) = 0{,}535$ și $H(250) = 0{,}520$ (curba din grafic)
  - Curba este teoretică și merge pînă la $q = 500$ doar ca să arate limita $H(q) \to 0{,}5$
  - Orizonturile ei nu sînt cele patru teste empirice, care rămîn la $q = 2, 5, 10, 20$
- Interpretare: memoria scurtă explică singură cea mai mare parte a lui $H(20)$ pentru BET, iar $H(q)$ revine spre 0,5: un singur orizont nu poate identifica memoria lungă

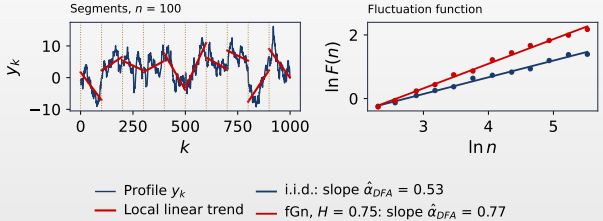
MFM_ch2_seminar

A3 [Propus]: Exponentul Hurst implicit pentru S&P 500 — enunț

- **Întrebare**: este S&P 500 „antipersistent" sau prezintă doar inversări pe termen scurt?
- **Date**: S&P 500, randamente logaritmice zilnice 2000–2026 (prețuri de închidere): $\widehat{\mathrm{VR}}(20) = 0{,}754$; un reper cu memorie scurtă, modelul AR(1) cu $\rho_1 = -0{,}1$, deci $\rho_k = (-0{,}1)^k$
- **Cerințe**
  1. Calculați $H(20)$ implicit pentru S&P 500 cu formula din A2.
  2. Calculați $\mathrm{VR}(20)$ și $H(20)$ pentru reperul AR(1).
  3. Calculați $\mathrm{VR}(250)$, $H(250)$ și limita $\mathrm{VR}(\infty) = (1+\rho_1)/(1-\rho_1)$ pentru reperul AR(1).
  4. Reprezentați $\mathrm{VR}(q)$ și $H(q)$ implicit pentru reper, cu q de la 2 la 500, alături de curba din A2.
- **De raportat**: cele șase valori, graficul și o frază care răspunde la întrebarea: implică $H(20) < 0{,}5$ memorie lungă?
- **Model**: A2 [Rezolvat]
- Notebook: secțiunea A3

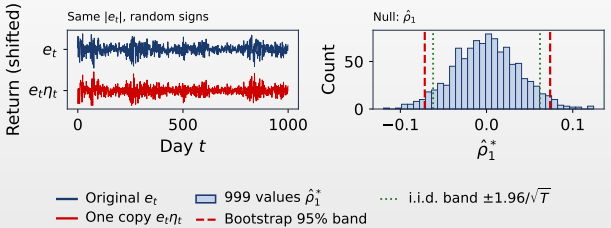
◂ Înapoi ▸ Lista exercițiilor

Analiza fluctuațiilor fără tendință (DFA)



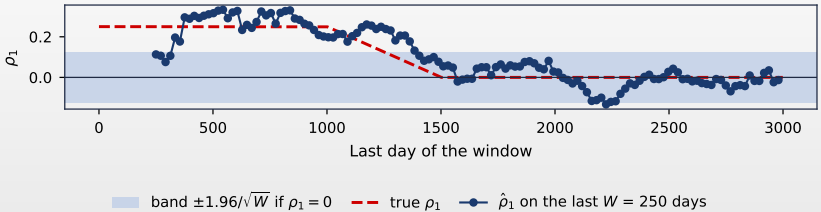
- **DFA** (*Detrended Fluctuation Analysis*, analiza fluctuațiilor fără tendință) Peng et al. (1994): estimează H urmărind cum cresc fluctuațiile profilului, după eliminarea tendinței, odată cu lungimea segmentului
  - profilul $y_k = \sum_{i \le k}(r_i - \bar{r})$; îl împărțim în segmente de lungime n și estimăm o dreaptă în fiecare (stînga)
  - $F(n)$: rădăcina mediei pătratelor reziduurilor față de dreptele locale; $F(n) \propto n^{\alpha_{DFA}}$
  - $\hat{\alpha}_{DFA}$: panta OLS a lui $\ln F(n)$ în funcție de $\ln n$ (dreapta); 0,5 fără memorie, peste 0,5 cu memorie lungă

Bootstrap-ul cu schimbări de semn (wild bootstrap)



- **Bootstrap-ul cu schimbări de semn** (*wild bootstrap*) Kim (2006): construim serii noi $e_t^* = e_t\eta_t$, cu semne $\eta_t = \pm 1$ alese independent, fiecare cu probabilitatea $1/2$
  - $|e_t^*| = |e_t|$: gruparea volatilității se păstrează; semnele aleatoare elimină orice autocorelație, deci H_0 este adevărată în fiecare copie
  - repetăm de B ori (aici 999), calculăm statistica de fiecare dată; cuantilele de 2,5% și 97,5% formează banda sub ipoteza nulă
  - pentru seria simulată din grafic ($T = 1\,000$), banda pentru $\hat{\rho}_1$ este $[-0{,}072; 0{,}074]$, mai largă decît banda i.i.d. $\pm 0{,}062$

Ferestre mobile și piețe adaptive



AMH: Adaptive Markets Hypothesis — ipoteza că eficiența se schimbă odată cu adaptarea participanților.

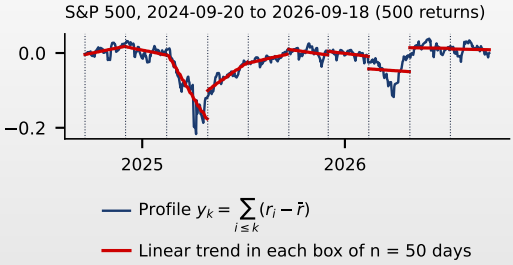
- **Fereastra mobilă** (*rolling window*): reestimăm pe ultimele W zile, deplasînd fereastra cu un pas fix
  - simulare: $\rho_1 = 0{,}25$ timp de 1 000 de zile, apoi scade spre 0; $W = 250$, pasul 21 de zile
  - ferestrele suprapuse folosesc în mare parte aceleași date; estimațiile din ferestre vecine sînt corelate, așa că numărul respingerilor are doar rol descriptiv
- **Ipoteza piețelor adaptive** (AMH) Lo (2004): eficiența se schimbă în timp, prin concurență și învățare
  - un tipar predictibil atrage investitori care îl exploatează pînă cînd dispare, ca în simularea de după ziua 1 000

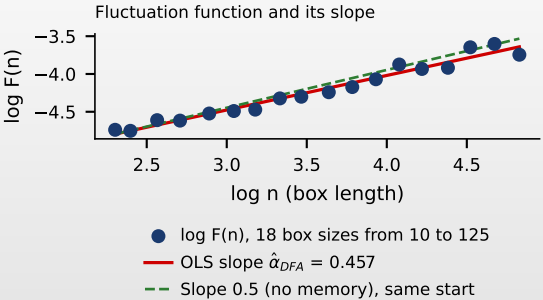
B5 [Rezolvat]: DFA pe ferestre mobile pentru S&P 500 — enunț

- **Întrebare**: iese vreodată exponentul DFA al S&P 500 din banda sub ipoteza nulă care permite gruparea volatilității?
- **Date**
  - S&P 500, randamente logaritmice zilnice 2000–2026; ferestre de 500 de randamente deplasate cu 21 de zile (297 de ferestre); funcția `dfa_hurst(x)`
  - Grila DFA: 18 lungimi de segment echidistante în $\log n$ de la 10 la 125 (un sfert din fereastră), trunchiate la întregi $(10, 11, 13, \ldots, 107, 124)$, fără duplicate
  - Pentru o lungime n, profilul se împarte în $k = \lfloor 500/n \rfloor$ segmente; cele $500 - kn$ puncte rămase se ignoră
  - Sămînța generatorului aleator (*seed*) în fiecare fereastră: 42 + indicele ultimei zile a ferestrei
- **Cerințe**
  1. Pentru ultima fereastră (2024-09-20 – 2026-09-18) construiți profilul, eliminați tendința pe segmente și estimați regresia lui $\log F(n)$ pe $\log n$.
  2. Calculați α_{DFA} în fiecare fereastră.
  3. În fiecare fereastră construiți banda sub ipoteza nulă prin schimbări de semn (Bootstrap-ul cu schimbări de semn (wild bootstrap)) din 199 de serii $e_t \eta_t$; păstrați cuantilele de 2,5% și 97,5% ale lui α_{DFA}.
  4. Calculați proporția ferestrelor în afara propriei benzi și în afara unei benzi fixe i.i.d. Student-t_4 (300 de simulări de lungime 500): distribuția Student t cu 4 grade de libertate are cozi mult mai groase decît distribuția Normală, deci această bandă păstrează cozile groase, dar nu și gruparea volatilității.
- **De raportat**: dreapta estimată pentru o fereastră, graficul pe ferestre mobile, ambele proporții și o frază despre ipotezele benzii
- Notebook: secțiunea B5

◂ Înapoi ▸ Lista exercițiilor

B5 [Rezolvat]: O fereastră, pas cu pas

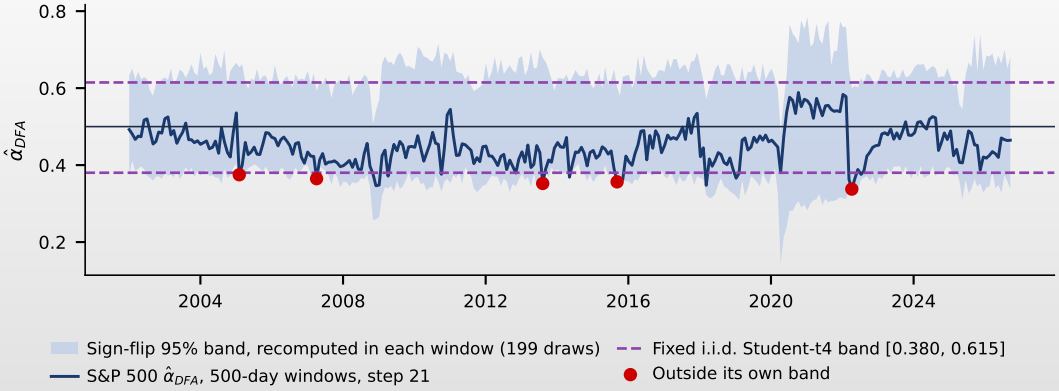




- Stînga: profilul y_k al ultimei ferestre, fără tendință pe fiecare segment, ca în Analiza fluctuațiilor fără tendință (DFA)
- 18 lungimi de segment de la 10 la 125; panta OLS a lui $\log F(n)$ pe $\log n$ este $\hat{\alpha}_{DFA} = 0{,}457$ (dreapta), aproape de 0,5

MFM_ch2_seminar

B5 [Rezolvat]: Toate ferestrele: grafic



MFM_ch2_seminar

B5 [Rezolvat]: Toate ferestrele: interpretare

- În afara propriei benzi: 1,7% din 297 de ferestre; în afara benzii fixe Student-t_4 [0,380; 0,615]: 6,7%
- Media $\hat{\alpha}_{DFA} = 0{,}451$; minimul 0,338 (fereastra încheiată la 2022-04-08), maximul 0,589 (2020-10-08); banda este cea mai largă (0,503) în fereastra încheiată la 2021-01-08
- Interpretare: banda este **punctuală**: acoperire de 95% pentru fiecare fereastră luată separat, nu pentru toate cele 297 ferestre deodată; sub ipoteza nulă, proporția așteptată în medie a ferestrelor din afara ei este de circa 5%, nu o valoare garantată în ferestre corelate
- Este un reper sub ipoteza unor semne simetrice și independente condiționat de $|e_t|$, nu un interval de încredere
- Ferestrele suprapuse sînt corelate, deci proporțiile au rol descriptiv și nu sînt un test al stabilității; proporția de 1,7% nu indică memorie peste nivelul așteptat din întîmplare; nerespingerea nu dovedește absența memoriei

MFM_ch2_seminar

B6 [Propus]: Eficiența adaptivă a Bitcoin — enunț

- **Întrebare**: a devenit Bitcoin mai eficient după primii ani, așa cum ar permite AMH Lo (2004)?
- **Date**: Bitcoin (USD), prețuri de închidere, randamente logaritmice zilnice 7 zile din 7, din 1 ian. 2011 pînă la 18 sept. 2026 (fișierul BTC-USD.CC.csv, inclusiv primii ani); 180 de ferestre de 365 de zile deplasate cu 30 de zile; H_0: RW3 în fiecare fereastră
- **Cerințe**
  1. Calculați $z^*(5)$ robust în fiecare fereastră.
  2. În fiecare fereastră construiți banda sub ipoteza nulă prin schimbări de semn pentru $z^*(5)$, cu 999 de replicări (seed = 42 + indicele ultimei zile a ferestrei), și păstrați cuantilele de 2,5% și 97,5%.
  3. Separat pentru ferestrele încheiate în 2012–2016 și 2017–2026, numărați-le pe cele în afara benzii și în afara pragurilor $\pm 1{,}96$ și comparați cu $0{,}05\times$ numărul ferestrelor.
  4. Comparați ferestrele 2012–2016 cu Urquhart (2016) (slide-ul următor): ce teste ale lui vizează RW3, care pot respinge doar din cauza grupării volatilității și concordă concluziile?
- **De raportat**: graficul, ferestrele semnalate, cu z^* și banda corespunzătoare, ambele numere de ferestre și o frază: ce stabilesc aceste numere și ce nu pot stabili?
- **Model**: B5, B1 [Rezolvat]
- Notebook: secțiunea B6

◂ Înapoi ▸ Lista exercițiilor

B6: ce detectează testele lui Urquhart

BDS: Brock–Dechert–Scheinkman–LeBaron test — test al independenței observațiilor.

- Comparați cu Urquhart (2016): 2010–2016, un set de teste clasice
  - Ljung–Box Ljung & Box (1978): $Q(m) = T(T+2)\sum_{k=1}^{m} \hat{\rho}_k^2/(T-k)$, aproximativ χ^2_m pentru randamente i.i.d., testează împreună primele m autocorelații
  - testul secvențelor și testul rangurilor al lui Bartels (1982), care înlocuiește randamentele cu rangurile lor și este robust la cozi groase
  - VR automat Choi (1999): orizontul q este ales din date
  - BDS Brock, Dechert, Scheinkman & LeBaron (1996): testează independența după cît de des perechi de istorii pe m zile rămîn apropiate; respinge și sub RW3 (gruparea volatilității), deci nu este un test al RW3
  - amplitudinea rescalată R/S din Exponentul Hurst

Alte teste și piețe

- După B1, examinează orizonturi testate împreună, testul secvențelor și piețe de comparație.
- Identifică abaterea detectată de fiecare test și separă familia orizonturilor de familia piețelor.
- Înapoi la alegerea modulelor.

Înainte să compari piețele

- B7 combină teste pe orizonturi și pe piețe; acestea sînt două familii diferite de ipoteze.

A4 [Propus]: Raportul variațelor și distribuția lui sub ipoteza nulă — enunț

- **Întrebări**
  - de ce $\sqrt{T}(\widehat{\mathrm{VR}}(q) - 1)$ are aproximativ distribuția Normală cu dispersia $2(2q-1)(q-1)/(3q)$?
  - cînd nu mai este valabilă această aproximare?
- **Date**: r_t staționar în covarianță: media și dispersia $\gamma_{\mathbf{0}}$ sînt constante, iar autocovariațele $\gamma_k = \mathrm{Cov}(r_t, r_{t-k})$ depind doar de lagul k, nu de t, $\rho_k = \gamma_k/\gamma_{\mathbf{0}}$; sub RW1 $\sqrt{T}(\hat{\rho}_{\mathbf{1}}, \ldots, \hat{\rho}_{q-\mathbf{1}})$ tinde la $N(0, I_{q-\mathbf{1}})$, $I_{q-\mathbf{1}}$ matricea identitate (componente independente, dispersii 1); S&P 500: $T = 6\,717$, $\widehat{\mathrm{VR}}(2) = 0{,}901$
- **Cerințe**
  1. Pentru $q = 3$, scrieți $\mathrm{Var}(r_t + r_{t-\mathbf{1}} + r_{t-\mathbf{2}})$ ca sumă a celor nouă covariațe și numărați de cîte ori apar $\gamma_{\mathbf{0}}$, $\gamma_{\mathbf{1}}$, $\gamma_{\mathbf{2}}$.
  2. Generalizați: arătați că $\mathrm{VR}(q) = 1 + \sum_{k=\mathbf{1}}^{q-\mathbf{1}} a_k \rho_k$ cu $a_k = 2(1 - k/q)$.
  3. Sub RW1, cu $a = (a_{\mathbf{1}}, \ldots, a_{q-\mathbf{1}})'$, calculați $a' I a = \sum_k a_k^{\mathbf{2}}$ și simplificați-l la $2(2q-1)(q-1)/(3q)$; evaluați-l pentru $q = 2, 5, 20$.
  4. Calculați $z(2)$ pentru S&P 500 și explicați de ce I trebuie înlocuit cu $\mathrm{diag}(\delta_k)$, matricea diagonală a dispersiilor robuste δ_k, sub RW3.
- **De raportat**: numărarea, dispersia în două coloane (asimptotică, a lui $\sqrt{T}(\widehat{\mathrm{VR}} - 1)$; în selecție finită, a lui $\widehat{\mathrm{VR}}$: împărțită la T) și $z(2)$
- **Model**: A1 [Rezolvat]
- Notebook: secțiunea A4

◂ Înapoi ▸ Lista exercițiilor

Testul secvențelor

- **Secvența**: un bloc maximal de randamente consecutive cu același semn Wald & Wolfowitz (1940)
  - n_1: numărul zilelor pozitive, n_2: al zilelor negative, $n = n_1 + n_2$
  - R: numărul secvențelor, egal cu numărul schimbărilor de semn $+1$
- Sub H_0 (semne independente), condiționat de n_1 și n_2

$$\mathbb{E}[R] = \frac{2n_1n_2}{n} + 1, \qquad \mathrm{Var}[R] = \frac{2n_1n_2(2n_1n_2 - n)}{n^2(n-1)}, \qquad z = \frac{R - \mathbb{E}[R]}{\sqrt{\mathrm{Var}[R]}} \approx N(0,1)$$

  - prea puține secvențe ($z < 0$): semnele persistă; prea multe ($z > 0$): semnele alternează
  - pentru n mic folosim distribuția exactă a lui R condiționată de n_1 și n_2, nu aproximarea Normală
- Exemplu: ++−−+−: $n_1 = 3$, $n_2 = 3$, $R = 4$
  - $\mathbb{E}[R] = 2 \cdot 9/6 + 1 = 4$: numărul observat de secvențe coincide cu cel așteptat

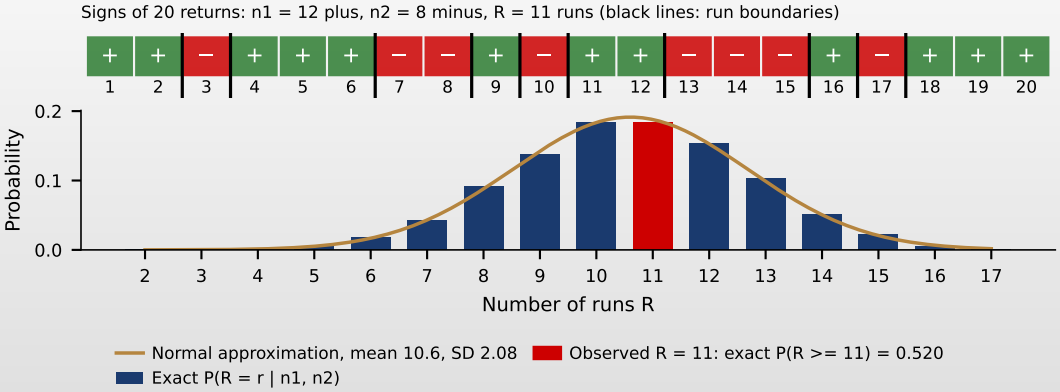
A5 [Rezolvat]: Testul secvențelor, pas cu pas — enunț

- **Întrebare**: sînt 20 de semne compatibile cu independența?
- **Date**: semnele a 20 de randamente zilnice consecutive, fără zerouri: ++-+++--+-++---+-+++; sub H_0 semnele sînt independente, cu o probabilitate constantă a unei zile pozitive Wald & Wolfowitz (1940)
- **Cerințe**
  1. Numerotați semnele de la 1 la 20, marcați fiecare schimbare de semn și numărați n_1, n_2 și secvențele R.
  2. Calculați $\mathbb{E}[R]$ și $\mathrm{Var}[R]$ cu formulele din Testul secvențelor.
  3. Calculați z aproximativ și valoarea p bilaterală corespunzătoare.
  4. Comparați cu distribuția exactă a lui R condiționată de n_1 și n_2.
- **De raportat**: n_1, n_2, R, $\mathbb{E}[R]$, $\mathrm{Var}[R]$, z, ambele valori p și ce ar însemna prea puține sau prea multe secvențe
- Notebook: secțiunea A5

◂ Înapoi ▸ Lista exercițiilor

A5 [Rezolvat]: Grafic

SD: Standard Deviation — abaterea standard a observațiilor.



MFM_ch2_seminar

A5 [Rezolvat]: Interpretare

- (1) Secvențele: $++ \mid - \mid +++ \mid -- \mid + \mid - \mid ++ \mid --- \mid + \mid - \mid +++$, deci $n_1 = 12$, $n_2 = 8$, $R = 11$ (10 schimbări de semn)
- (2) $\mathbb{E}[R] = 2 \cdot 12 \cdot 8/20 + 1 = 10{,}60$; $\mathrm{Var}[R] = 192 \cdot 172/(400 \cdot 19) = 4{,}345$
- (3) z aproximativ $= 0{,}40/2{,}085 = 0{,}19$, p bilateral $= 0{,}85$
- (4) Exact: $P(R = 11) = 0{,}183$, $P(R \geq 11) = 0{,}520$, $P(R \leq 11) = 0{,}663$, deci p exact bilateral $= \min\{1,\ 2\min[P(R \geq 11), P(R \leq 11)]\} = 1$
- Interpretare: nu respingem independența
  - prea puține secvențe ar însemna semne persistente
  - prea multe secvențe ar însemna semne alternante

MFM_ch2_seminar

B7 [Propus]: BET-TR și piețele emergente — enunț

- **Întrebare**: sînt piețele emergente mai puțin eficiente, ca grup?
- **Date**: BET-TR (prima închidere: 23 sept. 2014) și WIG20, BUX, PX, BIST 100, Bovespa, IPC, Nifty 50, Shanghai, prețuri de închidere, 2000–2026, fiecare pe calendarul propriu
- **Cerințe**
  1. Începeți cu BET și BET-TR pe aceeași perioadă (limitele BET-TR), pe observațiile păstrate din fiecare serie: $z(2)$ clasic, $z^*(2)$ robust, Chow–Denning.
  2. Repetați pentru ceilalți opt indici și raportați separat trei decizii: $q = 2$ clasic, $q = 2$ robust, test comun robust pe patru orizonturi.
  3. Înaintea oricărei concluzii despre grup, corectați valorile p ale testelor comune pentru BET-TR și cei opt indici enumerați (nouă teste, Holm).
- **De raportat**: un tabel (un rînd pe piață), comparația BET pe aceeași perioadă și o frază: cît se schimbă diferența BET versus BET-TR cînd perioadele coincid? Observațiile păstrate diferite nu permit atribuirea întregii diferențe rămase dividendelor.
- **Model**: B1 [Rezolvat]
- Notebook: secțiunea B7

◂ Înapoi ▸ Lista exercițiilor

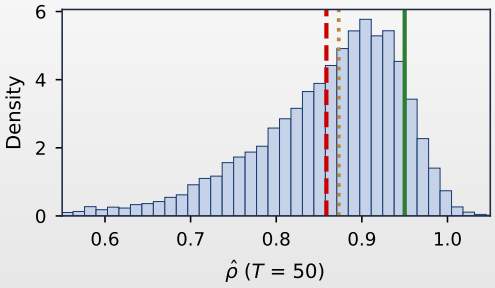
Prognoză

- După partea din curs despre regresii predictive, evaluează prognoze folosind numai informația disponibilă la data prognozei.
- Compară erorile pătratice în afara eșantionului cu reperul mediei istorice; ajustarea în eșantion nu dovedește o îmbunătățire a prognozei.
- Înapoi la alegerea modulelor.

Înainte să estimezi modelul de prognoză

- Folosește B3 ca exemplu rezolvat de regresie cu inferență HAC.
- În A6, care urmează, vei studia deplasarea asociată unui predictor persistent.

Regresia predictivă și deplasarea Kendall



- **Regresia predictivă**: randamentul perioadei următoare în funcție de un predictor x_t cunoscut azi

$$r_{t+1} = \alpha + \beta x_t + u_{t+1}, \quad x_{t+1} = c + \rho x_t + v_{t+1}$$

  - x_t: de exemplu logaritmul raportului dividend–preț $dp_t = \ln D_t - \ln P_t$, foarte persistent (ρ aproape de 1)
  - u, v: șocuri; o creștere a prețului îl reduce pe x, deci $\operatorname{corr}(u, v) < 0$

- **Deplasarea Kendall** Kendall (1954): în selecții mici, OLS tinde să subestimeze parametrul ρ al unui proces persistent

$$\mathbb{E}[\hat{\rho} - \rho] \approx -\frac{1 + 3\rho}{T}$$

  - graficul: $\rho = 0{,}95$, $T = 50$; media lui $\hat{\rho}$ este 0,859, sub valoarea adevărată
  - cînd u și v sînt corelate, deplasarea se transmite și estimatorului $\hat{\beta}$ Stambaugh (1999); în A6 deducem această legătură

A6 [Propus]: Deplasarea Stambaugh — enunț (1/2)

- **Întrebare**: de ce un indicator de evaluare persistent pare, prea des, un predictor semnificativ?
- **Date**
  - Sistemul, notația și deplasarea Kendall $\mathbb{E}[\hat{\rho} - \rho]$: Regresia predictivă și deplasarea Kendall, cu x_t logaritmul raportului dividend–preț
  - Scriem $u = \psi v + e$, $\psi = \sigma_{uv}/\sigma_v^2$, cu $\sigma_{uv} = \mathrm{Cov}(u, v)$, σ_u, σ_v abaterile standard ale șocurilor și e partea lui u necorelată cu v Stambaugh (1999)
  - Valori anuale ilustrative (ca fracții, nu în procente): $\rho = 0{,}98$, $T = 60$ de ani, $\mathrm{corr}(u, v) = -0{,}9$, $\sigma_u = 0{,}17$, $\sigma_v = 0{,}14$; în procesul simulat $\beta = 0$

înapoi Lista exercițiilor

A6 [Propus]: Deplasarea Stambaugh — enunț (2/2)

- **Context**: aceleași T perechi și același regresor centrat în ambele regresii OLS; șocurile i.i.d. împreună Normale fac ca e să fie independent de întreaga traiectorie a regresorului.
- **Cerințe**
  1. Folosind variabila centrată $\tilde{x}_t = x_t - \bar{x}$, arătați că $\hat{\beta} - \beta = \psi(\hat{\rho} - \rho) + \sum \tilde{x}_t e_{t+\mathbf{1}} / \sum \tilde{x}_t^{\mathbf{2}}$ și deduceți $\mathbb{E}[\hat{\beta} - \beta]$.
  2. Calculați ψ, $\mathbb{E}[\hat{\rho} - \rho]$ și deplasarea lui $\hat{\beta}$.
  3. Calculați eroarea standard asimptotică $\mathrm{SE}(\hat{\beta}) \approx \sigma_u / (\sigma_x \sqrt{T})$, unde $\sigma_x = \sigma_v / \sqrt{1 - \rho^{\mathbf{2}}}$ este abaterea standard staționară, și deplasarea exprimată în unități ale acestei erori standard.
  4. Fixați $\alpha = c = 0$, seed 42, și generați $x_{\mathbf{0}} \sim N(0, \sigma_x^{\mathbf{2}})$ independent de șocuri. Simulați 10 000 de selecții cu șocuri Normale; raportați media și abaterea standard a lui $\hat{\beta}$, eroarea standard OLS medie și proporția selecțiilor cu $|t| > 1{,}96$.
- **De raportat**: deducerea formulei, ψ, deplasarea, SE asimptotică, abaterea standard simulată și explicația efectului unei deplasări pozitive asupra ratei respingerilor false
- **Model**: exemplul Stambaugh lucrat din Capitolul 2
- Notebook: secțiunea A6

Prognoze în afara eșantionului

OS: Out-of-Sample — în afara eșantionului.

- Prognoză recursivă: la fiecare moment t, estimăm modelul pe datele disponibile pînă la t și calculăm prognoza $\hat{r}_{t+1}$; o comparăm cu media istorică $\bar{r}_t$
- R^2 **în afara eșantionului** Welch & Goyal (2008)

$$R^2_{OS} = 1 - \frac{\sum_t (r_{t+1} - \hat{r}_{t+1})^2}{\sum_t (r_{t+1} - \bar{r}_t)^2}$$

  - $R^2_{OS} > 0$: modelul prognozează mai bine decît media istorică; indicatorul poate lua și valori negative

- **Clark–West** Clark & West (2007): pentru modele imbricate, în care modelul mic este un caz particular al celui mare (media este modelul cu $\beta = 0$)

$$f_{t+1} = (r_{t+1} - \bar{r}_t)^2 - \left[(r_{t+1} - \hat{r}_{t+1})^2 - (\bar{r}_t - \hat{r}_{t+1})^2\right]$$

  - aplicăm un test unilateral bazat pe statistica t a mediei lui f_{t+1}, cu erori HAC; respingem ipoteza de acuratețe egală cînd $t > 1{,}645$, cuantila de 95% a distribuției $N(0,1)$
  - testul din Diebold & Mariano (1995): statistica t HAC a mediei lui $d_{t+1} = (r_{t+1} - \bar{r}_t)^2 - (r_{t+1} - \hat{r}_{t+1})^2$, diferența simplă a erorilor pătratice
  - ultimul termen din f_{t+1} elimină zgomotul estimării lui β, pe care testul Diebold–Mariano îl penalizează

- Restricțiile **Campbell–Thompson** Campbell & Thompson (2008)
  - se impune $\hat{\beta} = 0$ cînd are semnul greșit, iar o prognoză negativă se înlocuiește cu 0

B8 [Propus]: Regresie predictivă și un test în afara eșantionului — enunț (1/2)

- **Întrebare**: poate raportul dividend–preț să prezică randamentele în exces ale S&P 500, în eșantion și în afara eșantionului?

ETF: Exchange-Traded Fund — fond tranzacționat la bursă.

- **Date**
  - Tabel lunar 1934–2026, $T = 1\,111$
  - P: prețul mediu lunar S&P Composite al lui Shiller (pînă la 2026-06), prelungit cu mediile lunare ale închiderilor zilnice S&P 500, scalate cu media raportului Shiller/extensie din ultimele 12 luni comune
  - D: dividendele pe ultimele 12 luni (Shiller, interpolate din totaluri trimestriale), prelungite cu dividendele fondului SPY (*SPDR S&P 500 ETF Trust*), un ETF pe S&P 500
  - dividendele SPY sînt deduse din diferența dintre prețul de închidere și cel ajustat pentru dividende; seriile sînt racordate folosind perioada comună
  - Randamentul în exces: $r_{t+\mathbf{1}} = \ln((P_{t+\mathbf{1}} + D_{t+\mathbf{1}}/12)/P_t) - \ln(1 + \text{TB3MS}_t/1\,200)$, cu TB3MS (*FRED series: 3-month Treasury bill, secondary market rate*, seria FRED a randamentului titlurilor de stat americane pe 3 luni) seria FRED a ratei titlurilor de stat americane la 3 luni (piața secundară), cunoscută la t
  - Predictorul: $dp_t = \ln D_t - \ln P_t$, folosit cu un decalaj de o lună

înapoi Lista exercițiilor

B8 [Propus]: Regresie predictivă și un test în afara eșantionului — enunț (2/2)

- **Cerințe**
  1. Estimați $r_{t+1} = \alpha + \beta\, dp_t + u_{t+1}$ prin OLS cu erori Newey–West (12 laguri); estimați AR(1) pentru dp și $\mathrm{corr}(\hat{u}, \hat{v})$; corectați $\hat{\beta}$ pentru deplasarea Stambaugh (A6).
  2. Din ianuarie 1970, în fiecare lună reestimați pe datele pînă la luna anterioară și prognozați; reperul este media istorică a acelorași date; calculați R^2_{OS} Welch & Goyal (2008) (prognoză pseudo în afara eșantionului, *pseudo-out-of-sample*: recursivă, dar pe date revizuite sau interpolate ulterior, aici dividendele).
  3. Calculați t Clark–West (HAC, 12 laguri, unilateral) Clark & West (2007) și, pentru comparație, t Diebold–Mariano Diebold & Mariano (1995); impuneți restricțiile Campbell–Thompson (semnul pantei, prognoză nenegativă) Campbell & Thompson (2008).

- **De raportat**: $\hat{\beta}$, t, $\hat{\rho}$, deplasarea, R^2_{OS}, ambele statistici t și explicația pentru care testul Diebold–Mariano nu este potrivit în cazul modelelor imbricate
- **Model**: B3; exemplul Stambaugh lucrat din Capitolul 2
- Notebook: secțiunea B8

Un studiu de eveniment

- După partea din curs despre randamentele anormale, separă reacția măsurată la eveniment de semnificația statistică.
- Estimează randamentul normal înaintea evenimentului, cumulează randamentele anormale în fereastra evenimentului și ține cont de incertitudinea predicției.
- Înapoi la alegerea modulelor.

Studiul de eveniment (1/2)

AR (studiu de eveniment): Abnormal Return — randament anormal: randamentul observat minus randamentul dat de modelul de referință.

- **Modelul de piață** pe fereastra de estimare (T_0 zile înaintea evenimentului) MacKinlay (1997)

$$r_{i,t} = a_i + b_i\, r_{m,t} + \varepsilon_{i,t}, \qquad AR_{i,t} = r_{i,t} - \hat{a}_i - \hat{b}_i\, r_{m,t}$$

  - AR este aici randamentul anormal, nu modelul autoregresiv AR(p)
  - $r_{m,t}$: randamentul pieței (BET); $\hat{\sigma}_\varepsilon$: abaterea standard reziduală
  - ziua 0: prima zi de tranzacționare în care piața poate reacționa la anunț
  - $[\tau_1, \tau_2]$: fereastra evenimentului, în zile relative la ziua 0, de exemplu $[0, +5]$

CAR: Cumulative Abnormal Return — randamentul anormal cumulat.

- Randamentul anormal cumulat și dispersia lui

$$CAR[\tau_1, \tau_2] = \sum_{t=\tau_1}^{\tau_2} AR_{i,t}, \qquad \mathrm{Var}(AR_t) = \hat{\sigma}_\varepsilon^2 \left(1 + \mathrm{x}_t' (\mathrm{X}_0' \mathrm{X}_0)^{-1} \mathrm{x}_t\right)$$

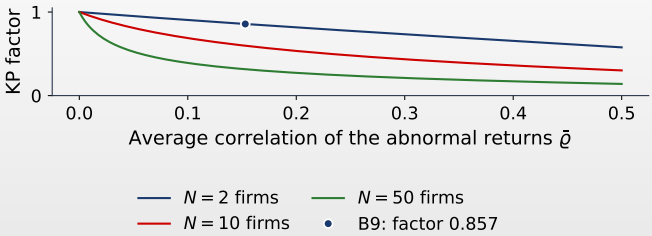
  - $\mathrm{x}_t = (1, r_{m,t})'$; X_0: rîndurile x_t' ale ferestrei de estimare
  - factorul din paranteză, asociat erorii de predicție, depășește 1 cînd $r_{m,t}$ se îndepărtează de valorile obișnuite, de exemplu la o prăbușire a pieței în ziua 0

Studiul de eveniment (2/2)

BMP: Boehmer–Musumeci–Poulsen — test pe randamente anormale standardizate.

- Mai multe firme la aceeași dată: reziduurile lor sînt corelate (corelația medie $\bar{\varrho}$)
  - intuiția, testul transversal simplu pentru N firme: $t = \overline{CAR}/(s_{CAR}/\sqrt{N})$, cu $\overline{CAR}$ și s_{CAR} media și abaterea standard a celor N valori; presupune firme independente
  - Boehmer, Musumeci & Poulsen (1991) (BMP): același test transversal, aplicat randamentelor anormale *standardizate*, fiecare împărțit la abaterea standard a erorii lui de predicție
  - Kolari–Pynnönen Kolari & Pynnönen (2010): înmulțim statistica BMP cu $\sqrt{(1-\bar{\varrho})/(1+(N-1)\bar{\varrho})}$, N: numărul firmelor

Firme corelate într-un studiu de eveniment



- Axa orizontală: corelația medie $\bar{\varrho}$ a randamentelor anormale; axa verticală: factorul Kolari–Pynnönen care înmulțește statistica BMP
- Pentru $\bar{\varrho} = 0$, factorul este 1: firmele independente nu cer corecție
- Factorul scade mai repede cînd sînt mai multe firme: $N = 50$ și $\bar{\varrho} = 0{,}1$ dau 0,39, deci statistica necorectată este de circa 2,6 ori prea mare
- Punctul: B9, două bănci cu $\bar{\varrho} = 0{,}153$, factorul 0,857

B9 [Propus]: Taxa bancară, TLV și BRD împreună — enunț (1/2)

- **Întrebare**: cît din scăderea provocată de taxa bancară rămîne semnificativ dacă inferența respectă regulile unui studiu de eveniment?

- **Date**
  - Banca Transilvania (TLV) și BRD (prețuri ajustate pentru dividende și split-uri) și BET (prețuri de închidere), prețuri aliniate pe zilele comune de tranzacționare, apoi randamente logaritmice
  - Cronologia: taxa a fost anunțată în seara de 18 dec. 2018; ziua 0 = 19 dec.; fereastra de estimare: zilele $[-250, -11]$ ($T_0 = 240$); fereastra evenimentului $[0, +5]$
  - X_0 (240×2) și X_e (6×2): rîndurile $[1, r_{m,t}]$ ale ferestrei de estimare și ale ferestrei evenimentului; ι: un vector 6×1 cu toate elementele 1; I_6: matricea identitate 6×6; N, $\bar{\varrho}$: ca în Studiul de eveniment

înapoi Lista exercițiilor

B9 [Propus]: Taxa bancară, TLV și BRD împreună — enunț (2/2)

- **Cerințe**
  1. Pentru TLV estimați modelul de piață pe $[-250, -11]$, calculați $AR_{\mathbf{0}}$ (randamentul anormal din ziua 0), t naiv $= AR_{\mathbf{0}}/\hat{\sigma}_{\varepsilon}$ și t cu dispersia erorii de predicție; repetați pentru BRD.
  2. Pentru portofoliul egal ponderat al celor două bănci estimați modelul de piață pe randamentele lui și calculați $AR_{\mathbf{0}}$ și t; comparați cu t care presupune bănci independente.
  3. Calculați $CAR[0, +5]$ și $CAR[+1, +5]$ cu $\mathrm{Var}(CAR) = \iota' \hat{\sigma}_{\varepsilon}^{\mathbf{2}} [I_{\mathbf{6}} + X_e (X_{\mathbf{0}}' X_{\mathbf{0}})^{-\mathbf{1}} X_e'] \iota$ (adaptat ferestrei folosite). Pentru L zile folosim cele L rînduri din X_e, L valori 1 și I_L: $L = 6$ pentru $[0, +5]$, $L = 5$ pentru $[+1, +5]$ (eliminăm ziua 0).
  4. Calculați corelația reziduurilor $\bar{\varrho}$ și factorul Kolari–Pynnönen $\sqrt{(1 - \bar{\varrho})/(1 + (N - 1)\bar{\varrho})}$ Kolari & Pynnönen (2010).
- **De raportat**: un tabel (TLV, BRD, portofoliu), graficul AR/CAR și dacă inversarea contrazice eficiența semi-tare; țineți cont de faptul că BET conține ambele bănci
- **Model**: Noțiunile despre studiul de eveniment; exemplul taxei bancare din Capitolul 2; B3 pentru HAC. Opțional: exercițiul B7 din Seminarul 0
- Notebook: secțiunea B9

Proiecte de cercetare

- Parcurge întîi modulul de prognoză; folosește un grup de control sau validarea încrucișată (*cross-validation*) ca să verifici critic concluzia.
- Separă alegerea modelului de evaluarea lui și precizează ce explicație alternativă poate elimina comparația.
- Înapoi la alegerea modulelor.

Ce trebuie să parcurgi înaintea proiectelor

- Pentru C2: parcurge B5 despre ferestre mobile și interpretarea prin piețe adaptive.
- Pentru C3: parcurge B8 despre evaluarea prognozelor în afara eșantionului.
- Paginile următoare introduc bootstrapul pe blocuri pentru C2 și regresia ridge cu validare încrucișată pentru C3.

Bootstrap pe blocuri și grup de control (1/2)

- **Bootstrap pe blocuri suprapuse** Künsch (1989): extragem blocuri de cîte ℓ zile consecutive, nu zile izolate
  - extragem uniform începuturile blocurilor din $1, \ldots, T - \ell + 1$ și concatenăm blocurile pînă obținem o serie de lungime T
  - dependența din interiorul fiecărui bloc (autocorelație, gruparea volatilității) se păstrează
  - intervalul percentil: cuantilele de 2,5% și 97,5% ale celor B estimații bootstrap
  - ℓ trebuie să fie suficient de lung pentru dependența relevantă (autocorelație, gruparea volatilității); C2 folosește $\ell = 20$ de zile, o alegere, nu o garanție, deci verificați sensibilitatea la ℓ

Bootstrap pe blocuri și grup de control (2/2)

FTSE: Financial Times Stock Exchange (FTSE Russell, index provider) — furnizorul de indici FTSE Russell.

- **Contrastul înainte/după** pentru un eveniment (o reclasificare)

$$\Delta_i = \hat{\rho}_{\mathbf{1},i}^{\mathbf{după}} - \hat{\rho}_{\mathbf{1},i}^{\mathbf{înainte}}$$

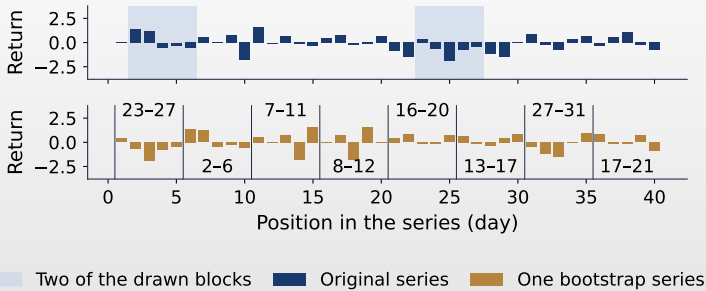
  - reclasificarea: FTSE Russell împarte pieţele în dezvoltate, emergente avansate, emergente secundare şi de frontieră (mărime, lichiditate, acces pentru investitorii străini, decontare)
  - România a trecut de la piaţă de frontieră la piaţă emergentă secundară pe 21 sept. 2020 (C2)
  - i: o piaţă; alte evenimente din aceiaşi ani, precum pandemia sau listările noi, pot modifica şi ele Δ_i şi constituie factori de confuzie

- **Grupul de control**: pieţe fără eveniment

$$\Delta_{\mathbf{tratat}} - \overline{\Delta}_{\mathbf{control}}$$

  - interpretarea este validă sub ipoteza *tendinţelor paralele*: în absenţa evenimentului, piaţa tratată ar fi evoluat la fel ca pieţele de control

Bootstrap-ul pe blocuri suprapuse: ilustrare grafică



- Sus: 40 de randamente zilnice simulate; jos: o serie bootstrap de aceeași lungime, formată din 8 blocuri de cîte $\ell = 5$ zile consecutive
- Etichetele arată zilele din care provine fiecare bloc: ordinea din interiorul blocului se păstrează, deci și dependența pe termen scurt
- Blocurile se extrag cu înlocuire: zilele 8–11 apar de două ori (blocurile 7–11 și 8–12), iar zilele 32–40 deloc

C2 [Propus]: Eficiența BVB după reclasificarea FTSE din 2020 — enunț (1/2)

- **Întrebare**: s-a apropiat autocorelația de ordinul 1 a BET de zero după reclasificare, față de WIG20 și BUX?
- **Date**
  - FTSE Russell (2019): decizie anunțată la 26 sept. 2019, în vigoare din 21 sept. 2020 (din piață de frontieră în piață emergentă secundară); context: Pele & Mazurencu-Marinescu (2012)
  - FTSE Russell împarte piețele în dezvoltate, emergente avansate, emergente secundare și de frontieră, după mărime, lichiditate, accesul investitorilor străini și decontare
  - Reclasificarea aduce fonduri pasive care urmăresc indicii piețelor emergente
  - BET, WIG20, BUX, prețuri de închidere; înainte: 21 sept. 2015 – 20 sept. 2020; după: 21 sept. 2020 – 21 sept. 2025; datele de după 21 sept. 2025 se folosesc doar într-o extindere raportată separat

Înapoi Lista exercițiilor

C2 [Propus]: Eficiența BVB după reclasificarea FTSE din 2020 — enunț (2/2)

COVID-19: COronaVIrus Disease 2019 — boala provocată de coronavirus, 2019.

- **Cerințe**
  1. Pentru fiecare piață calculați $\hat{\rho}_{\mathbf{1}}$ cu SE HAC (Newey–West pe funcția de influență, Capitolul 2) și $z^*(5)$ în ambele ferestre.
  2. Construiți un interval de 95% pentru schimbarea lui $\hat{\rho}_{\mathbf{1}}$ printr-un bootstrap pe blocuri suprapuse, independent în fiecare perioadă (blocuri de 20 de zile, 999 de replicări, seed 42) Künsch (1989).
  3. Reprezentați $\hat{\rho}_{\mathbf{1}}$ pe ferestre mobile de 250 de zile pentru cele trei piețe, cu banda de test robustă de 5% $0 \pm 1{,}96\sqrt{\hat{\tau}_{\mathbf{1}}}$ (SE sub $\rho_{\mathbf{1}} = 0$; o bandă de test în jurul lui zero, nu un interval în jurul lui $\hat{\rho}_{\mathbf{1}}$), marcînd anunțul, data intrării în vigoare și sfîrșitul ferestrei formale.
  4. Enumerați factorii de confuzie (COVID-19, listarea Hidroelectrica din 2023, lichiditatea) și precizați ce ar presupune un contrast ajustat cu grupul de control.
- **De raportat**: un folder Quantlet și un slide: concluzia, principala limită, legătura cu AMH Lo (2004); evaluați apropierea de zero, nu semnul
- **Model**: B1, B5 [Rezolvat]
- Notebook: secțiunea C2

Mulți predictori: ridge și validare încrucișată

- Cînd numărul J al predictorilor se apropie de numărul observațiilor, OLS ajunge să ajusteze și zgomotul: obținem un R^2 mare în eșantion (*in-sample*), dar prognoze slabe
- **Regresia ridge**: penalizăm coeficienții mari

$$\hat{\mathbf{h}}_\lambda = \arg\min_{\mathbf{h}} \sum_t (y_t - \mathbf{x}_t'\mathbf{h})^2 + \lambda\|\mathbf{h}\|^2 = (\mathbf{X}'\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}'\mathbf{y}$$

  - h: vectorul celor J coeficienți; I: matricea identitate $J \times J$
  - $\lambda \geq 0$: penalizarea; $\lambda = 0$ este OLS, $\lambda \to \infty$ aduce fiecare coeficient la 0
  - predictorii se standardizează întîi (media 0, abaterea standard 1), pentru ca penalizarea să îi trateze la fel
- **Validarea încrucișată**: alegem λ după eroarea de predicție pe datele lăsate în afara estimării
  - excludem pe rînd cîte un an: estimăm fără anul y, prognozăm pentru anul y, apoi repetăm și calculăm media erorilor
- **În eșantion și în afara eșantionului** Martin & Nagel (2022)
  - cînd investitorii învață, testele în eșantion pot respinge eficiența chiar dacă nimic nu este predictibil în timp real; doar rezultatul obținut în afara eșantionului poate fi pus în practică

C3 [Propus]: Big data și eficiența, în eșantion și în afara lui — enunț (1/2)

- **Întrebare**: cu 238 de predictori, predictibilitatea în eșantion a randamentelor sectoriale se regăsește și în afara eșantionului?
- **OOS**: Out-of-Sample — în afara eșantionului.
- **IS**: In-Sample — în eșantion.
- **Date**
  - Studiul de caz din Capitolul 2, Martin & Nagel (2022): învățarea rațională duce la respingeri ale eficienței în testele în eșantion, deși randamentele în afara eșantionului nu pot fi prezise; replicați planul empiric din secț. 7 a lucrării
  - 49 de portofolii sectoriale, Kenneth French Data Library, randamente lunare ponderate cu capitalizarea, iulie 1926 – ultima lună; valorile lipsă $-99{,}99$ și -999 sînt eliminate lună de lună; randamentele sînt împărțite la 100
  - Dimensiuni: în luna t, X_t are dimensiunea $N_t \times 238$ (randamente și pătratele lor la lagurile 2–120), N_t: numărul sectoarelor cu istoric complet de 120 de luni; h_t: vectorul 238×1 al coeficienților
  - $r_{\mathbf{OOS},t+\mathbf{1}} = \frac{\mathbf{1}}{N_{t+\mathbf{1}}} h_t' X_{t+\mathbf{1}}' r_{t+\mathbf{1}}$ este randamentul unui portofoliu long–short cu ponderile $X_{t+\mathbf{1}} h_t / N_{t+\mathbf{1}}$, adică prognozele centrate ale sectoarelor, cu h_t estimat doar pe date pînă la t
  - $r_{\mathbf{IS}}$: același randament, cu h estimat pe o fereastră care conține luna evaluată; diferența dintre cele două arată cît din predictibilitate nu era disponibilă în timp real

Înapoi Lista exercițiilor

C3 [Propus]: Big data și eficiența, în eșantion și în afara lui — enunț (2/2)

- **Cerințe**
  1. Centrați transversal randamentele și fiecare predictor în fiecare lună, scalați fiecare predictor la abaterea standard transversală 1, atribuiți fiecărei luni aceeași pondere.
  2. Estimați prin OLS pe datele cumulate (pooled) coeficienții h_t pe cele 240 de luni încheiate în t (doar date pînă la t); calculați $r_{\mathbf{OOS},t+\mathbf{1}} = \frac{\mathbf{1}}{N_{t+\mathbf{1}}} h_t' X_{t+\mathbf{1}}' r_{t+\mathbf{1}}$ și $r_{\mathbf{IS}}$ pe aceeași fereastră; precizați prima lună cu prognoză.
  3. Calculați media lui $r_{\mathbf{OOS}}$ pe ferestre de 10 ani, cu benzi descriptive ± 2 SE (ca în Fig. 8 din lucrare).
  4. Estimați o regresie ridge din ianuarie 1971 cu predictori standardizați; alegeți penalizarea pe o grilă logaritmică prin validare încrucișată cu cîte un an exclus (procedura poate folosi ani ulteriori: este o validare, nu o prognoză); reprezentați coeficienții pe laguri.
- **De raportat**: graficul de la pasul 3, mediile din prima și din a doua parte a perioadei pentru $r_{\mathbf{IS}}$ și $r_{\mathbf{OOS}}$ cu SE, graficul coeficienților și de ce 49 de portofolii pot diferi de mii de acțiuni
- **Model**: B3, B8
- Notebook: secțiunea C3 (Quantlet-ul de curs `MFM_ch2_big_data`)

Harta exercițiilor

Exercițiul	Întrebarea	Tipul	Model
A1	ce calculează $\widehat{\mathrm{VR}}(2)$, rînd cu rînd?	Rezolvat	–
A2, A3	dovedește un $H(20)$ implicit diferit de 0,5 existența memoriei lungi?	Rezolvat, Propus	A2
B1	este respinsă RW3 pentru S&P 500 și BET cînd folosim erori robuste?	Rezolvat	–
B2	este inversarea S&P 500 un efect al crizelor?	Propus	B1
B3	este real efectul schimbării lunii?	Rezolvat	–
B4	care dintre cele 15 efecte de calendar rămîn semnificative după corecția Holm?	Propus	B3
C1	ce erori conține un răspuns AI despre raportul variațelor?	Propus	A1, B1
După curs	B5 [Rezolvat], B6 [Propus]: DFA și rapoarte ale variațelor pe ferestre mobile, cu benzi bootstrap sub ipoteza nulă	temă	B1, B5
Teme de aprofundare	A4–A6, B7–B9, C2, C3 (opționale, după seminar)	Rezolvat, Propus	la fiecare enunț

- **[Rezolvat]**: găsiți rezolvarea completă în slide-uri și în notebook; folosiți-o ca model
- **[Propus]**: rezolvați exercițiul, apoi discutăm soluția la seminar

- Partea A: derivări pe hîrtie
- Partea B: calcule și inferență pe date reale; fiecare exercițiu se încheie cu o întrebare de interpretare
- Partea C: întrebări deschise și critica unui răspuns AI

Temele de aprofundare: cerințe prealabile

Tema	Exercițiile	Cerințe prealabile
1. Teorie	A4 (distribuția sub H_0), A5 [Rezolvat] (secvențe), A6 (deplasarea Stambaugh)	A1, B1
2. Mai multe piețe	B7	B1
3. Predictibilitate	B8	A6, B3
4. Studiu de eveniment	B9	Noțiunile despre studiul de eveniment; exemplul taxei bancare din Capitolul 2; B3 pentru HAC. Opțional: exercițiul B7 din Seminarul 0
5. Puncte de plecare pentru proiect (Partea C)	C2, C3	B1, B5; studiul de caz din Capitolul 2

- Fiecare exercițiu are secțiunea lui în notebook

Acronime folosite în acest material (1/2)

- **AI**: Artificial Intelligence — *inteligență artificială*
- **AMH**: Adaptive Markets Hypothesis — *ipoteza piețelor adaptive*
- **AR(p)**: AutoRegressive model: model autoregresiv de ordin p.
- **AR**: Abnormal Return: randament anormal, în studiile de eveniment.
- **BDS**: Brock–Dechert–Scheinkman–LeBaron (test) — *test de independență*
- **BET**: Bucharest Exchange Trading (index) — *indicele de referință al Bursei de Valori București*
- **BET-TR**: BET Total Return (index) — *indicele BET cu dividendele reinvestite*
- **BIST**: Borsa İstanbul (BIST 100 index) — *Bursa din Istanbul (indicele BIST 100)*
- **BUX**: Budapesti Értéktőzsde index — *indicele Bursei din Budapesta*
- **BVB**: Bursa de Valori București
- **CD**: Chow–Denning (multiple variance-ratio test) — *testul multiplu Chow–Denning pentru raportul variațelor*
- **COVID-19**: COronaVIrus Disease 2019 — *boala provocată de coronavirus, 2019*
- **DFA**: Detrended Fluctuation Analysis — *analiza fluctuațiilor fără tendință*
- **ETF**: Exchange-Traded Fund — *fond tranzacționat la bursă*
- **FRED**: Federal Reserve Economic Data — *baza de date economice a Rezervei Federale (St. Louis)*
- **FTSE**: Financial Times Stock Exchange (FTSE Russell, index provider) — *furnizorul de indici FTSE Russell*

Acronime folosite în acest material (2/2)

- **FWER**: Family-Wise Error Rate — *probabilitatea de cel puțin o eroare de tip I în familia de teste*
- **GARCH**: Generalized AutoRegressive Conditional Heteroskedasticity — *heteroscedasticitate condiționată autoregresivă generalizată*
- **HAC**: Heteroskedasticity and Autocorrelation Consistent (standard errors) — *erori standard robuste la heteroscedasticitate și autocorelație*
- **IPC**: Índice de Precios y Cotizaciones — *indicele Bursei din Mexic*
- **OLS**: Ordinary Least Squares — *metoda celor mai mici pătrate*
- **PX**: Prague Stock Exchange index (index Burzy cenných papírů Praha) — *indicele Bursei din Praga*
- **RW1**: Random Walk 1 (i.i.d. increments) — *mersul aleator 1 (creșteri i.i.d.)*
- **RW2**: Random Walk 2 (independent increments) — *mersul aleator 2 (creșteri independente)*
- **RW3**: Random Walk 3 (uncorrelated increments) — *mersul aleator 3 (creșteri necorelate)*
- **SE**: Standard Error — *eroare standard*
- **TOM**: Turn of the Month — *schimbarea lunii (efectul schimbării lunii)*
- **VR**: Variance Ratio — *raportul variațelor*
- **WIG20**: Warszawski Indeks Giełdowy 20 — *indicele celor mai mari 20 de companii de la Bursa din Varșovia*

Bibliografie (1/5)

- Ariel, R.A. (1987). A monthly effect in stock returns. *Journal of Financial Economics*, 18(1), 161–174.
- Bartels, R. (1982). The rank version of von Neumann's ratio test for randomness. *Journal of the American Statistical Association*, 77(377), 40–46.
- Boehmer, E., Musumeci, J., Poulsen, A.B. (1991). Event-study methodology under conditions of event-induced variance. *Journal of Financial Economics*, 30(2), 253–272.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
- Brock, W.A., Dechert, W.D., Scheinkman, J.A., LeBaron, B. (1996). A test for independence based on the correlation dimension. *Econometric Reviews*, 15(3), 197–235.
- Campbell, J.Y., Lo, A.W., MacKinlay, A.C. (1997). The Econometrics of Financial Markets. Princeton University Press.
- Campbell, J.Y., Thompson, S.B. (2008). Predicting excess stock returns out of sample: can anything beat the historical average? *Review of Financial Studies*, 21(4), 1509–1531.
- Choi, I. (1999). Testing the random walk hypothesis for real exchange rates. *Journal of Applied Econometrics*, 14(3), 293–308.
- Chow, K.V., Denning, K.C. (1993). A simple multiple variance ratio test. *Journal of Econometrics*, 58(3), 385–401.

Bibliografie (2/5)

- Clark, T.E., West, K.D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1), 291–311.
- Diebold, F.X., Mariano, R.S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253–263.
- Fama, E.F. (1970). Efficient capital markets: a review of theory and empirical work. *Journal of Finance*, 25(2), 383–417.
- Fama, E.F. (1991). Efficient capital markets: II. *Journal of Finance*, 46(5), 1575–1617.
- Franke, J., Härdle, W.K., Hafner, C.M. (2019). Statistics of Financial Markets: An Introduction, ed. a 5-a. Springer.
- French, K.R. (2026). Kenneth R. French Data Library. Tuck School of Business, Dartmouth College.
- FTSE Russell (2019). FTSE Russell announces results of country classification review (comunicat de presă, 26 septembrie 2019).
- Harvey, C.R., Liu, Y., Zhu, H. (2016). . . . and the cross-section of expected returns. *Review of Financial Studies*, 29(1), 5–68.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70.

Bibliografie (3/5)

- Hurst, H.E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116(1), 770–799.
- Kendall, M.G. (1954). Note on bias in the estimation of autocorrelation. *Biometrika*, 41(3–4), 403–404.
- Kim, J.H. (2006). Wild bootstrapping variance ratio tests. *Economics Letters*, 92(1), 38–43.
- Kolari, J.W., Pynnönen, S. (2010). Event study testing with cross-sectional correlation of abnormal returns. *Review of Financial Studies*, 23(11), 3996–4025.
- Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *Annals of Statistics*, 17(3), 1217–1241.
- Lakonishok, J., Smidt, S. (1988). Are seasonal anomalies real? A ninety-year perspective. *Review of Financial Studies*, 1(4), 403–425.
- Ljung, G.M., Box, G.E.P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303.
- Lo, A.W. (1991). Long-term memory in stock market prices. *Econometrica*, 59(5), 1279–1313.
- Lo, A.W. (2004). The adaptive markets hypothesis. *Journal of Portfolio Management*, 30(5), 15–29.

Bibliografie (4/5)

- Lo, A.W., MacKinlay, A.C. (1988). Stock market prices do not follow random walks: evidence from a simple specification test. *Review of Financial Studies*, 1(1), 41–66.
- MacKinlay, A.C. (1997). Event studies in economics and finance. *Journal of Economic Literature*, 35(1), 13–39.
- Martin, I.W.R., Nagel, S. (2022). Market efficiency in the age of big data. *Journal of Financial Economics*, 145(1), 154–177.
- Newey, W.K., West, K.D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703–708.
- Pele, D.T., Mazurencu-Marinescu, M. (2012). Modelling stock market crashes: the case of Bucharest Stock Exchange. *Procedia – Social and Behavioral Sciences*, 58, 533–542.
- Peng, C.-K., Buldyrev, S.V., Havlin, S., Simons, M., Stanley, H.E., Goldberger, A.L. (1994). Mosaic organization of DNA nucleotides. *Physical Review E*, 49(2), 1685–1689.
- Romano, J.P., Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica*, 73(4), 1237–1282.
- Stambaugh, R.F. (1999). Predictive regressions. *Journal of Financial Economics*, 54(3), 375–421.
- Urquhart, A. (2016). The inefficiency of Bitcoin. *Economics Letters*, 148, 80–82.

Bibliografie (5/5)

- Wald, A., Wolfowitz, J. (1940). On a test whether two samples are from the same population. *Annals of Mathematical Statistics*, 11(2), 147–162.
- Welch, I., Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4), 1455–1508.