








Modelarea piețelor financiare

Seminarul 14: Deep learning și modele fundaționale pentru serii de timp — aplicații



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

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

Prognoze mai bune sau doar modele mai complexe?

- Întrebarea seminarului
  - **rămîn valide prognozele de deep learning și ale modelelor fundaționale în urma unei inferențe corecte?**
- Începeți cu A1–A4 și consultați noțiunile indicate prin linkuri cînd este nevoie; A1 și A2 folosesc VaR (Value-at-Risk) și backtesting din cursurile 7 și 8
- Livrabile
  - răspunsurile la A1–A4; exercițiile suplimentare sînt opționale
  - un tabel și un grafic pentru B1–B4
  - codul corectat din C1
  - o propoziție de interpretare pentru fiecare exercițiu
- Notebook-ul de azi: deschideți notebook-ul seminarului în Colab
  - secțiunile notebook-ului poartă numele exercițiilor (A1, B1, . . .)
- Exercițiile, tipurile și ordinea lor: harta exercițiilor
- Seminar: prof. Daniel Traian Pele, danpele@ase.ro; drd. Antoaneta Amza, antoaneta.amza@csie.ase.ro

Datele folosite

Seria	Simbol, cîmp	Calendar	Perioada de test	Folosit în
S&P 500	`GSPC.INDX`, close	zile de tranzacționare SUA	04.01.2016–18.09.2026 (2 693)	A1, A3, B1, B3
Bitcoin	`BTC-USD.CC`, close	toate zilele	01.01.2019–18.09.2026 (2 818)	B4
BET	`BET`, close	zile de tranzacționare BVB	04.01.2016–18.09.2026 (2 675)	B4
ETF pe BET	`TVBETETF.RO`, adjusted close	zile de tranzacționare BVB	zilele comune cu BET	B4
SPY bare de 5 minute	`SPY.US`, close la 5 minute	ore de tranzacționare SUA	11.10.2022–18.09.2026 (979)	B2

- Prețurile zilnice din datele cursului, cîte o serie pe activ (cîmpul folosit: close / adjusted close); randamentele logaritmice zilnice în % (Notația (1/2)) pe calendarul propriu al fiecărei serii
- Prognozele la o zi calculate pentru Capitolul 14, salvate împreună cu capitolul: randamente (B1, B4), varianța realizată (B2), VaR și ES (B3); fiecare prognoză folosește date pînă la închiderea precedentă

Pregătire: rulați și verificați

```
# notebook section "Setup" defines
# load_csv(name): a forecast table
d = load_csv('ch14_returns_sp500.csv')
print(d.shape, d.index[0].date(),
      d.index[-1].date())
print(d[['y', 'hist', 'lstm',
         'chronos2_q0.50']].head(3).round(4))
rv = load_csv('ch14_rv.csv')
print(rv.shape)
```

- Rezultatul așteptat: (2693, 51), 2016-01-04 până la 2026-09-18; rv: (979, 12)

date	y	hist	lstm	$q_{0,50}$
2016-01-04	−1,5422	0,0265	0,0595	0,0988
2016-01-05	0,2010	0,0263	0,1044	0,0869
2016-01-06	−1,3202	0,0263	0,0587	0,0861

- y: randamentul realizat (%); celelalte coloane: prognozele făcute în seara precedentă
- Dacă dimensiunea sau primul rînd diferă, opriți-vă și verificați numele fișierului

Notația folosită în toate exercițiile (1/2): randamente, R^2 și testul DM

- r_t: randamentul logaritmic zilnic în %, $r_t = 100 \ln(P_t / P_{t-\mathbf{1}})$, cu P_t închiderea din ziua t; $\hat{r}_t$: prognoza lui, făcută în seara precedentă; pierderea portofoliului $L_t = -r_t$; L cu indicele unui model (L_t^A, $L_{k,t}$): pierderea unei prognoze
- $R^{\mathbf{2}}$ **în afara eșantionului**, față de prognoza zero

$$R^{\mathbf{2}} = 1 - \frac{\sum_t (r_t - \hat{r}_t)^{\mathbf{2}}}{\sum_t r_t^{\mathbf{2}}}$$

  - $R^{\mathbf{2}} > 0$: erori pătratice mai mici decît „mereu zero"; poate fi negativ, spre deosebire de $R^{\mathbf{2}}$ în eșantion
- **Testul DM** (Diebold–Mariano, Diebold & Mariano (1995)): $d_t = L_t^A - L_t^B$, pierderea prognozei A minus cea a lui B, pe T zile

$$\mathrm{DM} = \frac{\bar{d}}{\sqrt{\widehat{\mathrm{LRV}}/T}}, \qquad \widehat{\mathrm{LRV}} = \hat{\gamma}_{\mathbf{0}} + 2\sum_{j=\mathbf{1}}^{q} \Big(1 - \frac{j}{q+1}\Big)\hat{\gamma}_j$$

  - $\bar{d}$: media lui d_t; $\hat{\gamma}_j$: autocovarianța lui d_t la lag-ul j; $\widehat{\mathrm{LRV}}$: varianța pe termen lung Newey–West cu q lag-uri
  - aproximativ $N(0,1)$ sub acuratețe egală; negativ: A mai bun; $|\mathrm{DM}| > 1{,}96$ respinge la 5%

Notația folosită în toate exercițiile (2/2): pierderi pentru varianță și riscul de coadă

- RV (realised variance): varianța realizată, suma pătratelor randamentelor la 5 minute dintr-o zi; $\hat{h}_t$: o prognoză a variației
- **Pierderea QLIKE** (quasi-likelihood, Patton (2011)): 0 doar pentru o prognoză perfectă

$$L_t = \frac{\mathrm{RV}_t}{\hat{h}_t} - \log \frac{\mathrm{RV}_t}{\hat{h}_t} - 1$$

- **VaR și ES (Expected Shortfall)** la nivelul α (Capitolul 7), din cuantilele q_u ale lui r_{t+1}, $u \in (0,1)$

$$\mathrm{VaR}_\alpha = -q_\alpha, \qquad \mathrm{ES}_\alpha = -\frac{1}{\alpha}\int_0^\alpha q_u \, du$$

  - numere pozitive (pierderi); ES face media tuturor cuantilelor sub q_α, deci $\mathrm{ES}_\alpha \geq \mathrm{VaR}_\alpha$
  - indicatorul de depășire $I_t = 1\{L_t > \mathrm{VaR}_t\}$; un VaR corect are $P(I_t = 1) = \alpha$
- **Testul Kupiec** (Kupiec (1995)): x depășiri în T zile, rata observată $\hat{\pi} = x/T$

$$\mathrm{LR}_{uc} = -2\big[(T-x)\ln(1-\alpha) + x\ln\alpha - (T-x)\ln(1-\hat{\pi}) - x\ln\hat{\pi}\big] \sim \chi^2_1$$

  - $\mathrm{LR}_{uc} > 3{,}84$ respinge acoperirea corectă la 5%, atît pentru prea multe, cît și pentru prea puține depășiri

Rețele neuronale și coborîrea pe gradient

- **Un strat** al unei rețele neuronale

$$h = \varphi(Wx + b), \qquad \tanh(a) = \frac{e^a - e^{-a}}{e^a + e^{-a}}, \qquad \mathrm{ReLU}(a) = \max(a, 0)$$

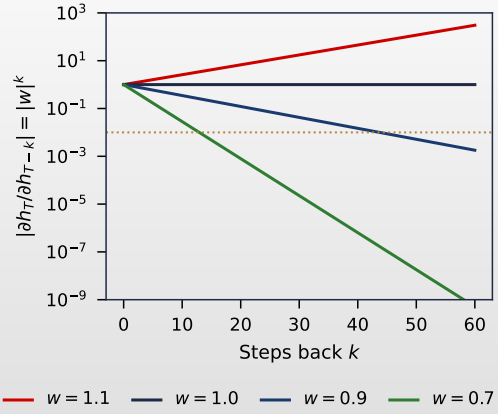
  - x: intrările; W: matricea ponderilor; b: vectorul termenilor liberi; φ: funcția de activare, aplicată element cu element
  - tanh ia valori în $(-1, 1)$, cu derivata $1 - \tanh^2(a) \leq 1$

- **Antrenarea**: coborîre pe gradient pe o funcție de pierdere $\mathcal{L}(\theta)$, de exemplu eroarea pătratică medie (MSE), θ = toate ponderile

$$\theta_{k+1} = \theta_k - \eta \nabla_\theta \mathcal{L}(\theta_k)$$

  - $\eta > 0$: rata de învățare; Adam rescalează fiecare coordonată a gradientului cu mărimea ei recentă
  - **propagarea inversă** (*backpropagation*, Rumelhart, Hinton & Williams (1986)): gradientul prin regula lanțului, strat cu strat, de la ieșire spre intrare
  - **oprirea timpurie** (*early stopping*): oprim antrenarea cînd pierderea pe o perioadă de validare nu mai scade; **seed-ul** (sămînța generatorului aleator) fixează punctul de pornire

Rețele recurente și dispariția gradientului



- **Rețea neuronală recurentă** (RNN): o stare h_t, actualizată în fiecare zi

$$h_t = \tanh(Wh_{t-\mathbf{1}} + Ux_t + b), \qquad \hat{y}_{t+\mathbf{1}} = v^\top h_t + c$$

  - x_t: intrarea; U: ponderile ei; v, c: ponderile de ieșire și termenul liber; aceeași matrice W la fiecare pas

- **Gradientul în timp** (Pascanu, Mikolov & Bengio (2013)): cîte o jacobiană la fiecare pas

$$\frac{\partial h_T}{\partial h_{T-k}} = \prod_{j=\mathbf{0}}^{k-\mathbf{1}} D_{T-j} W, \qquad D_t = \operatorname{diag}(1 - h_t^{\mathbf{2}})$$

- Cazul scalar, $D_t = 1$: produsul este w^k (graficul)
  - $w = 0{,}9$: $0{,}9^{\mathbf{20}} = 0{,}12$, $0{,}9^{\mathbf{50}} = 0{,}005$; o zi de acum 50 de pași aproape nu mai influențează ponderile: gradientul *dispare*
  - $|w| > 1$: gradientul explodează

Instrumente matriceale: raza spectrală, norma, VAR(1)

- **Raza spectrală** a unei matrice pătrate W, cu valorile proprii λ_i

$$\rho(W) = \max_i |\lambda_i|$$

  - pentru o matrice triunghiulară, valorile proprii sînt elementele de pe diagonală
- **Norma spectrală**: cea mai mare alungire a unui vector unitar

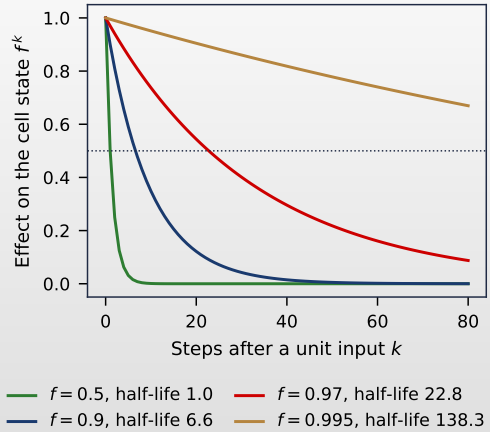
$$\|W\|_{\mathbf{2}} = \max_{\|v\|=\mathbf{1}} \|Wv\| = \sqrt{\lambda_{\max}(W^\top W)}, \qquad \|AB\|_{\mathbf{2}} \leq \|A\|_{\mathbf{2}} \|B\|_{\mathbf{2}}$$

  - $\rho(W) \leq \|W\|_{\mathbf{2}}$, cu egalitate pentru W simetrică
- **Modelul VAR(1)** (*Vector AutoRegression*, vector autoregresiv de ordinul 1) și funcția de răspuns la impuls la orizontul k

$$h_t = Wh_{t-\mathbf{1}} + \varepsilon_t \;\Rightarrow\; h_t = \sum_{k \geq \mathbf{0}} W^k \varepsilon_{t-k}, \qquad \frac{\partial h_t}{\partial \varepsilon_{t-k}} = W^k$$

  - stabil (șocurile se sting) dacă și numai dacă $\rho(W) < 1$

Celula LSTM



- **LSTM** (long short-term memory, Hochreiter & Schmidhuber (1997)): porțile f_t, i_t, $o_t \in (0,1)$, $\sigma(a) = 1/(1+e^{-a})$

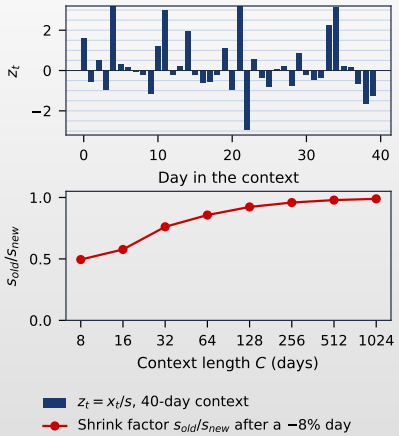
$$f_t = \sigma(W_f[h_{t-\mathbf{1}}, x_t] + b_f), \quad i_t = \sigma(W_i[h_{t-\mathbf{1}}, x_t] + b_i)$$
$$c_t = f_t \odot c_{t-\mathbf{1}} + i_t \odot g_t, \qquad h_t = o_t \odot \tanh(c_t)$$

  - f_t: poarta de uitare; i_t: poarta de intrare; o_t: poarta de ieșire; $g_t = \tanh(W_g[h_{t-\mathbf{1}}, x_t] + b_g)$: conținutul candidat
  - c_t: starea celulei; $\odot$: produs element cu element; $[h_{t-\mathbf{1}}, x_t]$: vectorul concatenat; σ: funcția sigmoid
- Poartă de uitare constantă f: $c_t = fc_{t-\mathbf{1}} + \text{intrare}$, un proces AR(1) (*AutoRegressive*, autoregresiv) cu coeficientul f

$$\text{timpul de înjumătățire} = \frac{\ln 0{,}5}{\ln f}$$

  - f aproape de 1: persistență mai mare pe un orizont finit; pentru $f < 1$ efectul tot dispare geometric (graficul)

Modele fundaționale și scalarea Chronos



- **Model fundațional** pentru serii de timp: preantrenat o singură dată pe multe serii, folosit **zero-shot**: fără estimare pe seria-țintă
  - intrarea: ultimele C valori (**contextul**, aici $C = 512$ zile); ieșirea: o grilă de cuantile q_u ale valorii următoare
  - prognoza punctuală: mediana $q_{0.5}$ sau media $\int_0^1 q_u \, du$
- **Chronos original** (Ansari et al. (2024)): scalarea prin medie și tokenii (intervale numerotate); Chronos-2 folosește standardizare și sinus hiperbolic invers

$$s = \frac{1}{C} \sum_t |x_t|, \qquad z_t = \frac{x_t}{s}, \qquad j = \operatorname{round}\left(\frac{z_t + 15}{\Delta}\right)$$

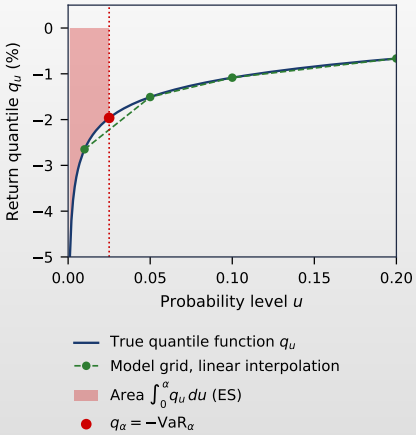
  - $\Delta = 30/4\,092$: lățimea unui interval; 4 093 de intervale pe $[-15, 15]$; intervalele sînt modelate ca niște cuvinte
  - cum schimbă un crah M scala s și celelalte z_t: derivați în A3 (panoul de jos, simulat)

Prognozele pe care le testăm

Prognoza	Descriere	Folosită în
Media istorică, AR(1)	repere clasice: media pe o fereastră care crește, AR(1) pe o fereastră mobilă de 1 000 de zile	B1
LSTM	rețea recurentă; randamente: media a 5 seed-uri, log RV: media a 3 seed-uri	B1, B2
Chronos-2	119 milioane de parametri, cuantile 1%, 5%, . . . , 99%; ponderi publicate pe 30.10.2025 (Ansari et al. (2025))	A1, B1, B2, B3
Chronos-Bolt	48 de milioane de parametri, cuantile 10%–90%; publicat pe 25.11.2024	B1, B2
TimesFM-2.5	231 de milioane de parametri, media și cuantilele 10%–90%; publicat pe 01.10.2025 (Das et al. (2024))	B1, B2, B3
FHS + volatilitatea unui model fundațional	hibrid: scala $\hat{\sigma}_t$ = prognoza modelului pentru $\lvert r_t \rvert$, apoi simularea istorică filtrată (FHS, McNeil & Frey (2000)) pe $r_t / \hat{\sigma}_t$	B3

- Fiecare prognoză este făcută în seara dinainte, din datele de pînă la acea închidere; la seminar se testează doar prognozele salvate; nu se antrenează nicio rețea
- **Contaminare**: un model preantrenat poate să fi văzut seriile noastre; un test fără contaminare a ponderilor folosește versiuni fixate ale ponderilor (*checkpoints*: revizie și amprentă SHA-256) și doar zilele de după publicarea lor, din 3 noiembrie 2025 (circa 220 de zile de tranzacționare, în A2 și C1; 218 zile cu RV în B2)
- Mai mult în Capitolul 14: atenția și patch-urile, designul complet al testelor, rezultatele de după publicare

VaR și ES dintr-o grilă de cuantile



- Un model furnizează cuantile doar pe o grilă, de exemplu $u \in \{0{,}01; 0{,}05; 0{,}10; \ldots\}$; între punctele grilei, **interpolare liniară**

$$q_u \approx q_{u_1} + \frac{u - u_1}{u_2 - u_1}\left(q_{u_2} - q_{u_1}\right), \quad u_1 \le u \le u_2$$

- ES_α din Notația (2/2): aria de sub funcția cuantilă pînă la α, împărțită la α (aria roșie; Student-t cu 4 grade de libertate)
  - sub primul nivel al grilei modelul nu spune nimic: coada trebuie presupusă (A1, A7)
- Reperul distribuției Normale cu media zero, cu $z_{1-\alpha} = \Phi^{-1}(1-\alpha)$

$$\frac{\mathrm{ES}_\alpha}{\mathrm{VaR}_\alpha} = \frac{\varphi(z_{1-\alpha})}{\alpha\, z_{1-\alpha}}$$

  - φ, Φ: densitatea și funcția de repartiție a distribuției Normale standard; cozile mai groase dau un raport mai mare

Depășirea prognozei zero

- **Intervalul bootstrap pe blocuri** pentru R^2: reeșantionăm blocuri de 20 de zile consecutive (circular), recalculăm R^2, citim percentilele 2,5% și 97,5%
- **Proporția semnelor corecte** $\frac{1}{T}\sum_t 1\{\text{sign}\,\hat{r}_t = \text{sign}\,r_t\}$; reperul: proporția zilelor de creștere, nu 50%
- **Modele imbricate**: media istorică și AR(1) conțin prognoza zero drept caz particular; sub H_0, zgomotul de estimare face testul DM prea conservator
- **Clark–West** (Clark & West (2007)): diferența de pierderi ajustată, cu $\hat{r}_{0,t}$ reperul și $\hat{r}_{1,t}$ modelul mai mare

$$f_t = (r_t - \hat{r}_{0,t})^2 - \left[(r_t - \hat{r}_{1,t})^2 - (\hat{r}_{0,t} - \hat{r}_{1,t})^2\right]$$

  - reperul zero ($\hat{r}_{0,t} = 0$): $f_t = 2 r_t \hat{r}_{1,t}$
  - test t unilateral cu erori standard HAC (*Heteroskedasticity and Autocorrelation Consistent*, robuste la heteroscedasticitate și autocorelație) pentru $E[f_t] = 0$; respingem dacă $t > 1{,}645$ la 5%
- **Mai multe modele, o singură familie**: căutarea printre șase modele găsește un „cîștigător" din noroc; testele de pe slide-ul următor țin cont de aceasta

Mai multe modele simultan: SPA, Romano–Wolf, mulțimea de încredere a modelelor

- $d_{k,t} = L_{\mathbf{0},t} - L_{k,t}$: cîștigul modelului k față de reper; $\bar{d}_k$: media lui; $\hat{\omega}_k$: abaterea standard de termen lung, estimată prin bootstrap; $n = T$ este numărul zilelor de prognoză comune
- **Testul SPA** (*Superior Predictive Ability*, Hansen (2005)): $H_\mathbf{0}$: niciun model nu depășește reperul, $\max_k E[d_{k,t}] \leq 0$

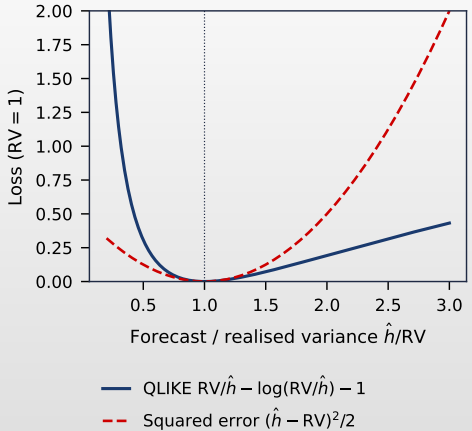
$$T^{SPA} = \max_k \max\left(\frac{\sqrt{n}\,\bar{d}_k}{\hat{\omega}_k},\ 0\right)$$

  - valoarea p: proporția replicilor bootstrap ale maximului recentrat care depășesc T^{SPA}; bootstrap staționar, blocuri de lungime aleatoare, cu media 20 de zile
- **Romano–Wolf** (Romano & Wolf (2005)): procedură descendentă; respingem orice model peste cuantila bootstrap de 95% a maximului, îl eliminăm, repetăm; o valoare p ajustată pentru fiecare model
- **Mulțimea de încredere a modelelor** (MCS, *Model Confidence Set*, Hansen, Lunde & Nason (2011)): mulțimea care conține cel mai bun model cu probabilitatea 90%

$$t_m = \frac{\bar{L}_m - \bar{L}_\cdot}{\widehat{\text{se}}(\bar{L}_m - \bar{L}_\cdot)}, \qquad T_{\max} = \max_m t_m$$

  - $\bar{L}_m$: pierderea medie a modelului m; $\bar{L}_\cdot$: media pe modelele rămase; dacă valoarea p bootstrap a lui $T_{\max}$ este sub 0,10, eliminăm cel mai slab model și repetăm

Prognoze de volatilitate: HAR, GARCH-t, QLIKE



- **HAR** (heterogeneous autoregressive, Corsi (2009)): mediile zilnică, săptămînală și lunară ale RV trecute

$$\mathrm{RV}_{t+\mathbf{1}} = b_{\mathbf{0}} + b_{\mathbf{1}}\mathrm{RV}_t + b_{\mathbf{5}}\overline{\mathrm{RV}}_t^{(\mathbf{5})} + b_{\mathbf{22}}\overline{\mathrm{RV}}_t^{(\mathbf{22})} + u_{t+\mathbf{1}}$$

  - $\overline{\mathrm{RV}}_t^{(j)}$: media ultimelor j zile; log-HAR: aceeași regresie în $\log \mathrm{RV}$, retransformată cu $\exp(\cdot + \hat{s}^{\mathbf{2}}/2)$

- **GARCH(1,1)**-t (*Generalised AutoRegressive Conditional Heteroskedasticity*): varianța doar din randamentele zilnice

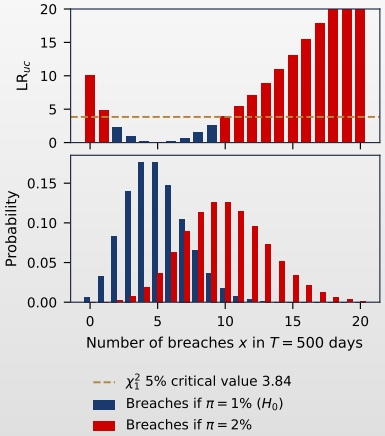
$$\sigma_t^{\mathbf{2}} = \omega + \alpha r_{t-\mathbf{1}}^{\mathbf{2}} + \beta\sigma_{t-\mathbf{1}}^{\mathbf{2}}, \qquad r_t = \sigma_t z_t$$

  - z_t: Student-t, varianță unitară; $\alpha + \beta < 1$: persistența

- **QLIKE** (Notația (2/2)): penalizează subestimarea mai mult decît supraestimarea; valoarea ei așteptată este minimă pentru $\hat{h} = E[\mathrm{RV} \mid \text{informație}]$ (A10)

- Reperul mers aleator: $\hat{h}_{t+\mathbf{1}} = \mathrm{RV}_t$

Mărimea și puterea testului Kupiec

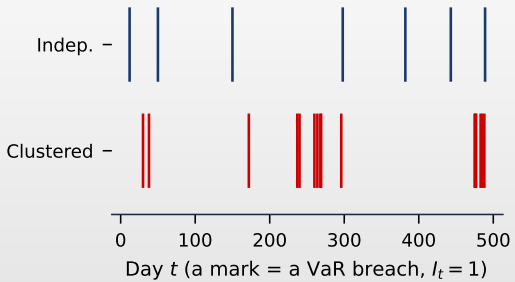


- Depășirile unui VaR cu probabilitatea adevărată de depășire π: distribuția binomială

$$P(x = k) = \binom{T}{k} \pi^k (1 - \pi)^{T-k}, \quad k = 0, \ldots, T$$

  - media $T\pi$; $T = 500$, $\pi = 1\%$: 5 depășiri așteptate
- **Regiunea de respingere**: valorile x cu $\mathrm{LR}_{uc}(x) > 3{,}84$ (barele roșii, sus)
- **Mărimea exactă**: $P_{\pi=\alpha}(x \in \text{regiune})$; **puterea**: $P_\pi(x \in \text{regiune})$ pentru un π greșit
  - x este discret, deci mărimea nu este exact 5%
- Graficul: $T = 500$, $\alpha = 1\%$: respingem dacă $x \leq 1$ sau $x \geq 10$; mărimea 7,1%; față de $\pi = 2\%$ puterea este doar 54%
- O nerespingere pe un eșantion scurt este o dovadă slabă a acoperirii corecte

Depășiri grupate: Christoffersen și DQ



Independent: $x = 7$, $\hat{\pi}_{01} = 0.014$, $\hat{\pi}_{11} = 0.00$

Clustered: $x = 16$, $\hat{\pi}_{01} = 0.027$, $\hat{\pi}_{11} = 0.19$

- Un VaR corect este depășit cu probabilitatea α *independent* de ziua precedentă
- **Testul de independență Christoffersen** (Christoffersen (1998)): n_{ij} = numărul zilelor cu $I_{t-1} = i$, $I_t = j$

$$\hat{\pi}_{01} = \frac{n_{01}}{n_{00} + n_{01}}, \qquad \hat{\pi}_{11} = \frac{n_{11}}{n_{10} + n_{11}}$$

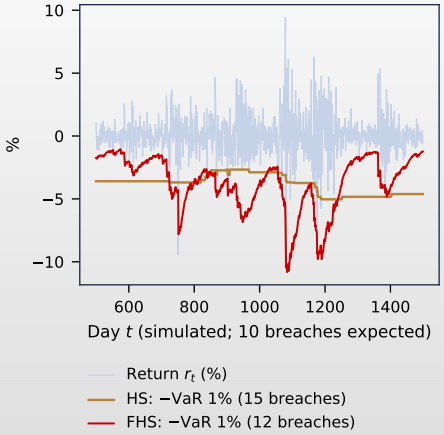
  - test LR (*Likelihood Ratio*, testul raportului de verosimilitate) pentru $\pi_{01} = \pi_{11}$, χ^2_1; depășirile grupate dau $\hat{\pi}_{11} \gg \hat{\pi}_{01}$ (graficul). Acoperirea condiționată testează rata și independența împreună: $\mathrm{LR}_{cc} = \mathrm{LR}_{uc} + \mathrm{LR}_{ind} \sim \chi^2_2$ sub ipoteza nulă

- **Testul cuantilei dinamice** (DQ, Engle & Manganelli (2004)): regresăm $I_t - \alpha$ pe o constantă, pe depășirile anterioare și pe VaR_t

$$\mathrm{DQ} = \frac{\hat{\beta}^\top X^\top X \hat{\beta}}{\alpha(1-\alpha)} \sim \chi^2_k$$

  - X: cei k regresori; $\hat{\beta}$: coeficienții estimați prin metoda celor mai mici pătrate; un VaR corect anulează toți coeficienții

Simularea istorică, FHS și pierderea FZ0



- **Simularea istorică** (HS, *Historical Simulation*): VaR = minus cuantila empirică de nivel α a ultimelor W randamente
- **HS filtrată** (FHS, McNeil & Frey (2000)): standardizăm, luăm cuantila, rescalăm cu volatilitatea de azi

$$\mathrm{VaR}_{\alpha,t} = -\hat{\sigma}_t \, \hat{q}_\alpha \left(\left\{ \frac{r_s}{\hat{\sigma}_s} \right\}_{s=t-W}^{t-1} \right)$$

  - $\hat{\sigma}_t$: volatilitatea GARCH sau a unui model fundațional; FHS reacționează la volatilitate, HS nu (graficul, simulat)
- **Pierderea FZ0** (Fissler–Ziegel, varianta omogenă de grad zero, Patton, Ziegel & Chen (2019)) pentru perechea (VaR v, ES e), pierderea L, nivelul α

$$\mathrm{FZ0} = \frac{1\{L > v\}(L - v)}{\alpha e} + \frac{v}{e} + \log e - 1$$

  - mai mică este mai bună; modelele se compară prin DM; penalizează atît cozile ratate, cît și VaR și ES inutil de mari

Instrumente pentru pachetele de extindere

- **PIT** (*Probability Integral Transform*, transformarea integrală de probabilitate) (Diebold, Gunther & Tay (1998)): $u_t = \hat{F}_t(r_t)$, funcția de repartiție prognozată, evaluată în valoarea realizată; i.i.d. uniformă dacă prognoza este corectă
  - Berkowitz (Berkowitz (2001)): $z_t = \Phi^{-1}(u_t) \sim N(0,1)$; test LR pentru $\mu = 0$, $\sigma = 1$, χ^2_2
- **Scorul cuantilic** QS (*Quantile Score*, pierderea *pinball*) al unei prognoze de cuantilă q_u, mediat pe nivelurile grilei u_k

$$\mathrm{QS}_t = \frac{2}{K} \sum_{k=1}^{K} \left(1\{r_t < q_{u_k}\} - u_k\right)\left(q_{u_k} - r_t\right) \geq 0$$

- **Depășiri peste prag** (A7): excesele $L - u$ peste un prag ridicat u urmează o distribuție Pareto generalizată (GPD, *Generalised Pareto Distribution*)

$$P(L - u \leq y \mid L > u) = 1 - \left(1 + \frac{\xi y}{\beta}\right)^{-1/\xi}$$

  - ξ: parametrul de formă (mai mare: coadă mai groasă); $\beta > 0$: parametrul de scală
- **Recalibrarea conformală** (Gibbs & Candès (2021)): deplasăm o cuantilă cu cuantila empirică a erorilor recente; varianta adaptivă actualizează nivelul $\alpha_{t+1} = \alpha_t + \gamma(\alpha - I_t)$

Harta exercițiilor

Exercițiul	Întrebarea	Tipul	Model
A1	cum citim VaR și ES dintr-o grilă de cuantile?	Rezolvat	–
A2	ce putere are un backtest pe 220 de zile?	Rezolvat	–
A3	cum modifică un crah scala Chronos?	Rezolvat	–
A4	de ce dispare gradientul unei RNN?	Rezolvat	–
B1	poate un model să depășească prognoza zero pentru randamentele S&P 500?	Rezolvat	–
B2	este un model fundațional mai bun decît log-HAR pentru volatilitate?	Rezolvat	–
B3	trec prognozele VaR ale modelelor fundaționale backtest-urile?	Rezolvat	–
C1	ce este greșit într-un răspuns AI?	Propus	A2, B3
B4	Bitcoin și BET: căutarea unui semnal (acasă)	Propus	B1
Pachete de extindere	A5–A10, B5–B11, C2, C3 (opționale, după seminar)	ambele	la fiecare cerință

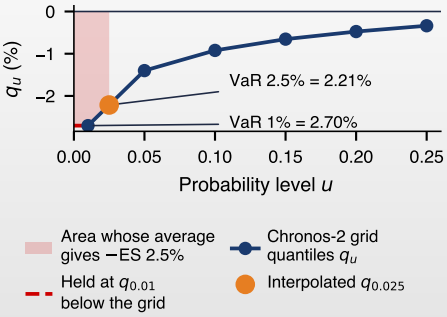
- **[Rezolvat]**: rezolvarea completă în slide-uri și în notebook, un model de urmat; **[Propus]**: derivați sau reproduceți rezultatul
- A: derivări pe hîrtie; B: inferență pe prognozele salvate, în notebook, fiecare cu o întrebare de interpretare; C: întrebări deschise și critica răspunsului unui asistent de inteligență artificială (AI)

A1 [Rezolvat]: VaR 1% și ES 2,5% dintr-o grilă de cuantile — Cerință

- **Întrebarea**: cum transformăm grila de cuantile a unui model fundațional în VaR și ES?
- **Date**: ultima prognoză zero-shot Chronos-2 pentru S&P 500 (*Standard & Poor's*), cu cuantilele pentru 18.09.2026, din contextul de 512 randamente pînă la 17.09.2026: $q_{0,01} = -2{,}704\%$, $q_{0,05} = -1{,}400\%$; nivelurile grilei 1%, 5%, 10%, . . . , 99%
- **Cerințe**
  1. Calculați VaR 1% cu convenția $\mathrm{VaR}_\alpha = -q_\alpha$.
  2. Interpolați liniar $q_{0,025}$ între $q_{0,01}$ și $q_{0,05}$ și calculați VaR 2,5%.
  3. Calculați ES 2,5%, păstrînd $q_u = q_{0,01}$ pentru $u < 0{,}01$ și folosind q_u liniar pe $[0{,}01; 0{,}025]$.
- **Raportați**: VaR 1%, VaR 2,5%, ES 2,5% și raportul ES 2,5% / VaR 2,5%
- **Interpretare**: comparați raportul cu valoarea lui pentru o distribuție Normală de medie zero, la același nivel: ce informație lipsește din grilă?

◂ Înapoi ▸ Lista exercițiilor

A1 [Rezolvat]: Rezolvare



1. VaR 1% $= -q_{0,01} = 2{,}704\%$.
2. $q_{0,025} = q_{0,01} + \frac{0,015}{0,04}(q_{0,05} - q_{0,01}) = -2{,}704 + 0{,}375\,(-1{,}400 + 2{,}704) = -2{,}215\%$, deci VaR 2,5% $= 2{,}215\%$.
3. Împărțim integrala (zona roșie): $\int_0^{0,01} q_u\,du = 0{,}01\,q_{0,01}$ și $\int_{0,01}^{0,025} q_u\,du = 0{,}015\,(q_{0,01} + q_{0,025})/2$; ES 2,5% $= -\frac{1}{0,025}[\ldots] = 2{,}557\%$.

⊡ Interpretare: ES/VaR = 1,15, sub valoarea pentru distribuția Normală, $\varphi(1{,}96)/(0{,}025 \cdot 1{,}96) = 1{,}19$

- φ: densitatea distribuției Normale standard
- Grila trunchiază coada de sub 1%, deci acest ES este o aproximare cu coada trunchiată

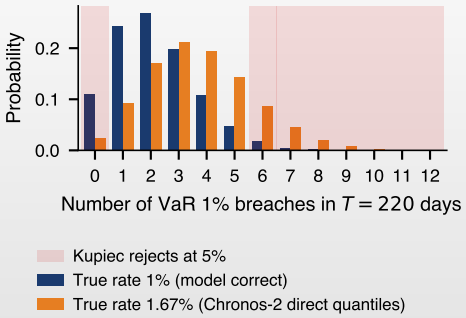
MFM_ch14_seminar

A2 [Rezolvat]: Puterea exactă a testului Kupiec — Cerință

- **Întrebarea**: pot 220 de zile de după publicarea ponderilor să deosebească un VaR 1% corect de unul depășit în 1,67% din zile?
- **Date**: $T = 220$ de zile (3.11.2025–18.09.2026); depășirile $x \sim \mathrm{Bin}(T, \pi)$ (mărimea și puterea testului Kupiec); testul Kupiec bilateral din Notația (2/2), la 5%; 1,67% este rata de depășire a Chronos-2 direct pe S&P 500 în tot eșantionul (B3)
- **Cerințe**
  1. Calculați $\mathrm{LR}_{uc}(x)$ pentru $x = 0, 1, 2, \ldots$ cu $\alpha = 1\%$ și găsiți regiunea de respingere.
  2. Calculați mărimea exactă a testului, probabilitatea de a respinge cînd $\pi = 1\%$.
  3. Calculați puterea, probabilitatea de a respinge cînd $\pi = 1{,}67\%$.
- **Raportați**: regiunea de respingere, mărimea și puterea
- **Interpretare**: ce îi spune unui manager de risc o nerespingere pe 220 de zile?

◂ Înapoi ▸ Lista exercițiilor

A2 [Rezolvat]: Rezolvare



1. $x = 0$: $\mathrm{LR} = -2 \cdot 220 \ln 0{,}99 = 4{,}42 > 3{,}84$, respingem (prea puține depășiri); $x = 5$: LR = 2,65; $x = 6$: LR = 4,51; regiunea de respingere $x = 0$ sau $x \geq 6$.
2. Mărimea $= P_{\mathbf{1\%}}(x = 0) + P_{\mathbf{1\%}}(x \geq 6) = 0{,}134$, nu 0,05: regula folosește pragul asimptotic $\chi^2_{\mathbf{1}}$, o regiune de respingere calibrată exact pe distribuția binomială ar ține mărimea sub 5%.
3. Puterea exactă a acestei reguli $= 0{,}9833^{\mathbf{220}} + P_{\mathbf{1{,}67\%}}(x \geq 6)$ $= 0{,}025 + 0{,}165 = 0{,}189$; depășiri așteptate 2,2 sub $H_{\mathbf{0}}$, 3,7 sub alternativă.

⊡ Interpretare: cele două distribuții se suprapun aproape complet

- În patru cazuri din cinci, un model cu 67% mai multe depășiri decît trebuie trece testul
- Deci „nerespins" nu dovedește acoperirea corectă

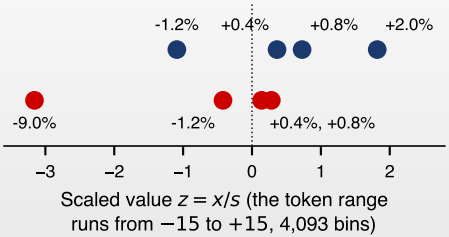
MFM_ch14_seminar

A3 [Rezolvat]: Scalarea prin medie în Chronos: efectul unei zile noi asupra scalei — Cerință

- **Întrebarea**: cît de mult schimbă o zi nouă scala la care Chronos împarte contextul?
- **Date**: scala Chronos s, valorile scalate z și intervalele j din modele fundaționale și scalarea Chronos; un context de patru zile $x = (0{,}8; -1{,}2; 0{,}4; 2{,}0)$ (%)
- **Cerințe**
  1. Fereastra avansează o zi: intră o valoare nouă M și iese x_{out}; derivați s_{nou} în funcție de s_{vechi}, M, x_{out} și C.
  2. Derivați factorul cu care se schimbă valoarea scalată z a fiecărei zile neschimbate.
  3. Pentru contextul de patru zile calculați s, valorile scalate z și indicii j.
- **Raportați**: formula de la pașii 1–2 și un tabel cu x, z, j
- **Interpretare**: cum depinde de C efectul unui crah asupra scalei?

Înapoi Lista exercițiilor

A3 [Rezolvat]: Rezolvare



1. Suma valorilor absolute pierde $|x_{\text{out}}|$ și cîștigă $|M|$; împărțind la C: $s_{\text{nou}} = s_{\text{vechi}} + (|M| - |x_{\text{out}}|)/C$.
2. Zi neschimbată x: $z_{\text{nou}} = x/s_{\text{nou}} = z_{\text{vechi}} \cdot s_{\text{vechi}}/s_{\text{nou}}$.
3. $s = (0{,}8 + 1{,}2 + 0{,}4 + 2{,}0)/4 = 1{,}10$; $z = (+0{,}727;\ -1{,}091;\ +0{,}364;\ +1{,}818)$; $\Delta = 0{,}00733$; $j = (2145, 1897, 2096, 2294)$, lîngă centrul $j = 2046$ al lui $z = 0$.

⊡ Interpretare: un crah mută s cu circa $|M|/C$ (roșu: ultima zi înlocuită cu -9%, toate celelalte zile se micșorează); un context de 512 zile aproape nu reacționează

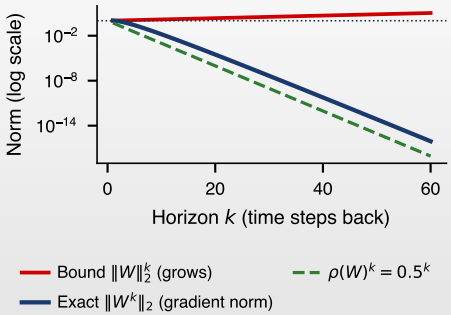
MFM_ch14_seminar

A4 [Rezolvat]: Dispariția gradientului ca răspuns la impuls al unui VAR — Cerință

- **Întrebările**
  - de ce dispare gradientul unei rețele recurente simple?
  - ce mărime determină viteza de dispariție?
- **Date**: RNN din rețele recurente și dispariția gradientului, cu două unități ascunse, $W = \begin{pmatrix} 0{,}5 & 0{,}8 \\ 0 & 0{,}5 \end{pmatrix}$, liniarizată în $h^* = 0$ ($b = 0$, $x = 0$); fără fișier de date
- **Cerințe**
  1. Derivați jacobiana $\partial h_T / \partial h_{T-k}$ prin regula lanțului și demonstrați că norma ei spectrală este cel mult $\|W\|_2^k$.
  2. Arătați că rețeaua liniarizată este un VAR(1) în h_t și că jacobiana este egală cu răspunsul lui la impuls la orizontul k.
  3. Calculați raza spectrală $\rho(W)$, norma $\|W\|_2$, majorarea $\|W\|_2^k$ și valoarea exactă $\|W^k\|_2$ pentru $k = 5, 20, 50$.
- **Raportați**: cele două derivări și un tabel cu cifrele de la pasul 3
- **Interpretare**: majorarea prin normă sau raza spectrală este criteriul potrivit pentru dispariția gradientului?

Înapoi Lista exercițiilor

A4 [Rezolvat]: Rezolvare



1. Un pas: $\partial h_t/\partial h_{t-1} = D_t W$, $D_t = \mathrm{diag}(1 - h_t^2)$; pe k pași $\prod_j D_{T-j} W$; cum $\|D_t\|_2 \leq 1$ și normele sînt submultiplicative, produsul este cel mult $\|W\|_2^k$.
2. În $h^* = 0$, $D_t = I$: $h_t \approx W h_{t-1} + U x_t$, un VAR(1); din $h_T = \sum_k W^k U x_{T-k}$, $\partial h_T/\partial h_{T-k} = W^k$ este răspunsul la impuls.
3. W este triunghiulară: $\rho(W) = 0{,}50$; $\|W\|_2 = 1{,}040$, deci majorarea $1{,}040^{20} = 2{,}20$ crește; exact $W^k = 0{,}5^k \begin{pmatrix} 1 & 1{,}6k \\ 0 & 1 \end{pmatrix}$: $\|W^5\|_2 = 0{,}254$, $\|W^{20}\|_2 = 3{,}1 \times 10^{-5}$, $\|W^{50}\|_2 = 7{,}1 \times 10^{-14}$.

⊡ Interpretare: majorarea prin normă este suficientă, nu necesară

- Gradientul dispare ca $k\,0{,}5^k$ pentru că $\rho(W) < 1$, criteriul de stabilitate al unui VAR
- Gelfand: $\|W^k\|^{1/k} \to \rho(W)$

MFM_ch14_seminar

B1 [Rezolvat]: Modelele față de prognoza zero pentru S&P 500 — Cerință

- **Întrebarea**: prognozează vreunul dintre cele șase modele randamentele zilnice ale S&P 500 mai bine decît prognoza zero, dacă testăm corect?
- **Date**: `ch14_returns_sp500.csv`: randamentul realizat `y` și prognozele `hist` (media istorică), `ar1` (fereastră mobilă de 1 000 de zile), `lstm` (media a 5 seed-uri), `chronos2_point`, `bolt_point`, `timesfm_point` (mediane); 2 693 de zile din 04.01.2016
- **Cerințe**
  1. Calculați R^2 în afara eșantionului pentru fiecare model și un interval bootstrap de 95%, circular, pe blocuri (blocuri de 20 de zile, 1 000 de extrageri).
  2. Testați fiecare model față de zero cu testul DM pe erorile pătratice (bilateral, 5%) și calculați proporția semnelor corecte.
  3. Aplicați testul Clark–West (Clark & West (2007)) mediei istorice și AR(1), care conțin prognoza zero drept caz particular.
  4. Aplicați SPA (Hansen (2005)) și Romano–Wolf (Romano & Wolf (2005)) celor șase modele (bootstrap staționar, bloc mediu de 20 de zile).
- **Raportați**: un grafic cu R^2 și intervalele, un tabel (R^2, interval, DM [p], semne corecte) și valorile p Clark–West, SPA și Romano–Wolf
- **Interpretare**
  - este semnificativ aici un R^2 de circa 1%?
  - de ce 55% semne corecte nu dovedesc abilitate?
- **Întrebare pentru sală**: ce credeți: depășește LSTM prognoza zero?

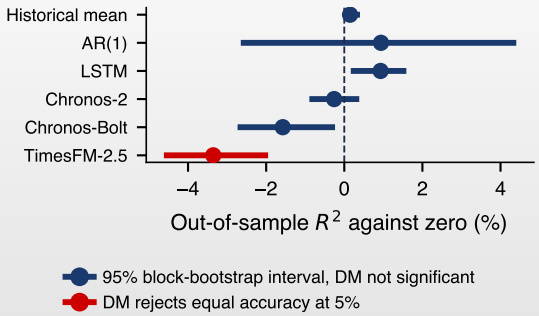
Înapoi Lista exercițiilor

B1 [Rezolvat]: Rezolvare (1/3) — Pas cu pas

1. Aliniați prognozele și randamentele pe cele 2 693 de zile de test; pierderile $e^2_{0,t} = r_t^2$ (zero) și $e^2_{1,t} = (r_t - \hat{r}_t)^2$ (model); $R^2 = 1 - \sum e_1^2 / \sum e_0^2$. Bootstrap: extrageți blocuri de 20 de zile consecutive, circular, recalculați R^2 de 1 000 de ori, luați percentilele de 2,5% și 97,5%; LSTM: R^2 +0,93%, interval [+0,17; +1,59].
2. DM: $d_t = e^2_{1,t} - e^2_{0,t}$, variantă Newey–West cu $\lfloor 4(T/100)^{2/9} \rfloor$ lag-uri; LSTM: DM $= -1{,}70$, $p = 0{,}090 > 0{,}05$: acuratețea egală nu este respinsă, deși intervalul bootstrap exclude zero. Semne corecte LSTM: 53,5%, sub cele 54,6% de zile de creștere.
3. Clark–West cu reperul zero (depășirea prognozei zero): un test t HAC unilateral pentru $E[r_t \hat{r}_t] = 0$ (Clark & West (2006)); valorile p unilaterale, DM față de Clark–West (CW): media istorică 0,048 față de 0,007; AR(1) 0,320 față de 0,055.
4. SPA: bootstrap pentru maximul cîștigurilor studentizate recentrate; $p = 0{,}130$ (H_0: niciun model nu depășește prognoza zero); Romano–Wolf elimină pas cu pas modelele respinse: media istorică 0,164, LSTM 0,161.

⊡ **Răspuns**: nu, după corecția pentru șase modele: SPA $p = 0{,}130$, Romano–Wolf $p = 0{,}161$

B1 [Rezolvat]: Rezolvare (2/3) — Grafic



- Puncte: R^2; bare: intervale bootstrap de 95% pe blocuri; roșu: DM respinge acuratețea egală cu zero la 5%
- TimesFM-2.5 este semnificativ mai slab decît zero; AR(1) are un interval foarte larg, determinat de 2020
- Interpretare: CW detectează micul drift al mediei istorice
  - După căutarea printre șase modele (SPA, Romano–Wolf), niciun model nu depășește prognoza zero

MFM_ch14_seminar

B1 [Rezolvat]: Rezolvare (3/3) — Tabel

Model	R^2 (%)	Interval 95%	DM [p]	Semn corect (%)
Media istorică	+0,15	[−0,02; +0,40]	−1,67 [0,095]	54,6
AR(1)	+0,94	[−2,65; +4,40]	−0,47 [0,639]	54,1
LSTM	+0,93	[+0,17; +1,59]	−1,70 [0,090]	53,5
Chronos-2	−0,26	[−0,89; +0,38]	+0,81 [0,418]	54,7
Chronos-Bolt	−1,57	[−2,73; −0,24]	+1,76 [0,078]	54,6
TimesFM-2.5	−3,35	[−4,62; −1,95]	+3,23 [0,001]	53,6

- Indicele a crescut în 54,6% din zile: regula „mereu în creștere” ghicește deja semnul în această proporție
- LSTM: intervalul bootstrap exclude zero, dar p DM = 0,090; ambele privesc un cîștig predictiv nul, cu aproximări diferite în eșantion finit
  - intervalul: percentile bootstrap pentru un raport, blocuri de 20 de zile; DM: inferență HAC pentru diferența medie de pierdere, cu 8 lag-uri
  - niciunul nu rezistă corecției pentru șase modele

MFM_ch14_sem_partB

B2 [Rezolvat]: Volatilitatea realizată: modele fundaționale față de log-HAR — Cerință

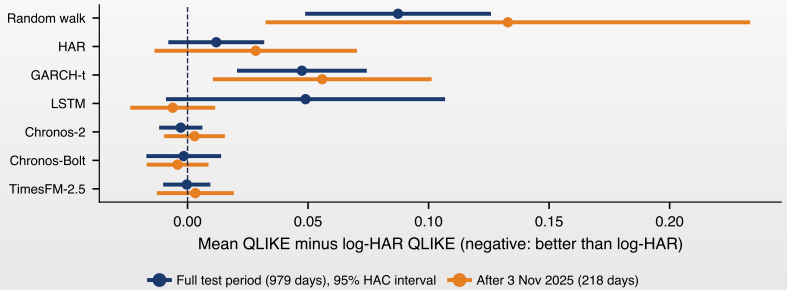
- **Întrebarea**: sînt modelele fundaționale zero-shot mai bune decît log-HAR la prognoza variației realizate a SPY (*SPDR S&P 500 ETF Trust*, fondul tranzacționat la bursă pe indicele S&P 500)?
- **Date**: `ch14_rv.csv`: `RV` (suma a 78 de pătrate ale randamentelor la 5 minute pe grila 09:30–16:00, ora New York-ului; primul randament: deschiderea și închiderea barei de 09:30; doar zilele cu grila completă) și prognozele `RW`, `HAR`, `logHAR` (Corsi (2009); fereastră mobilă de 500 de zile), `GARCH-t`, `LSTM`, `Chronos-2`, `Chronos-Bolt`, `TimesFM-2.5` (media din grilă); 979 de zile
- **Cerințe**
  1. Calculați pierderea QLIKE zilnică a fiecărui model și media ei pe cele 979 de zile.
  2. Testați fiecare model față de log-HAR cu DM pe diferențele QLIKE (bilateral, 5%) și reprezentați diferențele medii cu intervale HAC de 95%.
  3. Calculați mulțimea de încredere a modelelor (MCS) de 90% (Hansen, Lunde & Nason (2011); bootstrap pe blocuri de 10 zile).
  4. Repetați pașii 1–2 pe cele 218 zile de după 3 noiembrie 2025, cînd toate cele trei seturi de ponderi erau publice.
- **Raportați**: graficul și un tabel: QLIKE, DM [p], p MCS, QLIKE și DM după publicare
- **Interpretare**
  - ce informație oferă apartenența la MCS?
  - ce informație nu oferă?

◂ Înapoi ▸ Lista exercițiilor

B2 [Rezolvat]: Rezolvare (1/3) — Pas cu pas

1. Pierderea zilnică $L_{t,m} = \mathrm{RV}_t/\hat{h}_{t,m} - \log(\mathrm{RV}_t/\hat{h}_{t,m}) - 1$; de exemplu la 11.10.2022, RV = 1,578 și prognoza Chronos-2 1,484 (ambele variațe zilnice, în puncte procentuale la pătrat): $L = 1{,}064 - \log 1{,}064 - 1 = 0{,}0020$; medii: Chronos-2 0,223, log-HAR 0,226.
2. $d_t = L_{t,m} - L_{t,\log \mathrm{HAR}}$; Chronos-2: $\bar{d} = -0{,}0028$, DM $= -0{,}61$, deci eroarea standard HAC este $\bar{d}/\mathrm{DM}$, iar intervalul de 95% $[-0{,}0118;\ +0{,}0062]$ conține 0 ($p = 0{,}539$).
3. MCS: bucla de eliminare din mai multe modele simultan, cu pierderile QLIKE; p MCS al unui model este maximul tuturor valorilor p de eliminare pînă la eliminarea sa, astfel încît mulțimile păstrate să fie imbricate.
4. După 3.11.2025: Chronos-2 $\bar{d} = +0{,}0029$, interval $[-0{,}0097;\ +0{,}0155]$: mai larg, pentru că 218 zile conțin mai puțină informație.

B2 [Rezolvat]: Rezolvare (2/3) — Grafic



- Puncte: diferența medie de QLIKE față de log-HAR; bare: intervale HAC de 95%; albastru: toată perioada, portocaliu: după 3.11.2025
- Chronos-2 are cea mai mică QLIKE, dar toate intervalele modelelor fundaționale conțin 0
  - Mersul aleator și GARCH-t sînt mai slabe și sînt eliminate din MCS
- Interpretare: apartenența la MCS înseamnă „nu se distinge la 90%", nu „la fel de bun"

MFM_ch14_seminar

B2 [Rezolvat]: Rezolvare (3/3) — Tabel

	Toată perioada (979 de zile)			După 3.11.2025 (218 zile)	
Model	QLIKE	DM [p]	MCS p	QLIKE	DM [p]
Mers aleator	0,313	+4,44 [<0,001]	0,005	0,356	+2,59 [0,010]
HAR	0,238	+1,17 [0,241]	0,33	0,251	+1,32 [0,187]
log-HAR	0,226	–	0,95	0,223	–
GARCH-t	0,273	+3,45 [<0,001]	0,03	0,279	+2,41 [0,016]
LSTM	0,275	+1,66 [0,097]	0,29	0,217	−0,68 [0,494]
Chronos-2	0,223	−0,61 [0,539]	1,00	0,226	+0,45 [0,650]
Chronos-Bolt	0,224	−0,20 [0,842]	0,95	0,219	−0,63 [0,527]
TimesFM-2.5	0,225	−0,07 [0,948]	0,95	0,226	+0,40 [0,692]

- DM față de log-HAR; pozitiv: modelul este mai slab decît log-HAR
- După publicare, LSTM are cea mai mică QLIKE (0,217), dar numai mersul aleator și GARCH-t diferă semnificativ de log-HAR
- HAR în niveluri supraestimează; GARCH-t reacționează la randamentele zilnice, nu la informația intrazilnică

MFM_ch14_sem_partB

B3 [Rezolvat]: Backtesting pentru prognozele VaR și ES pe S&P 500 — Cerință

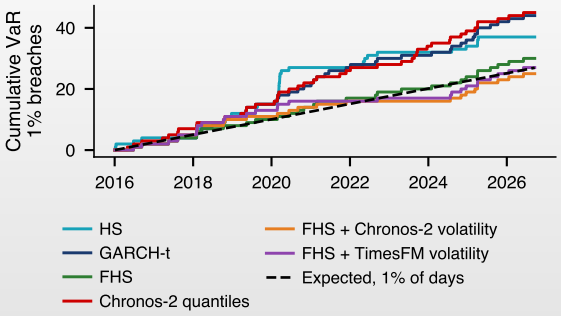
- **Întrebarea**: dau cuantilele directe Chronos-2 sau FHS cu volatilitatea unui model fundațional un VaR 1% și un ES 2,5% utilizabile?
- **Date**: `ch14_risk_sp500.csv`: randamentul `y` și, pentru șase modele, coloanele `<model>|VaR1`, `|VaR2.5`, `|ES2.5`: HS, GARCH-t, FHS (fereastră de 1 000 de zile, reestimare la fiecare 20 de zile), `C2-raw` (cuantilele directe Chronos-2), `C2-FHS` și `TFM-FHS` (FHS cu volatilitatea Chronos-2, respectiv TimesFM); 2 693 de zile
- **Cerințe**
  1. Numărați depășirile VaR 1% ale fiecărui model și reprezentați numărul lor cumulat față de dreapta așteptată de 1%.
  2. Aplicați testul Kupiec și testul de independență Christoffersen (Christoffersen (1998), χ^2_1) la 5%.
  3. Calculați pierderea FZ0 medie (Patton, Ziegel & Chen (2019)) a perechii (VaR 2,5%, ES 2,5%) și testele DM față de FHS.
  4. Pentru FHS și hibridul Chronos-2 calculați testul DQ (Engle & Manganelli (2004)): regresia lui $I_t - 0{,}01$ pe o constantă, patru depășiri întîrziate și VaR_t.
- **Raportați**: graficul și un tabel: depășiri, rata, p Kupiec, p independență, FZ0, DM față de FHS
- **Interpretare**: de ce trece hibridul testul de acoperire și totuși este inferior FHS?

◂ Înapoi ▸ Lista exercițiilor

B3 [Rezolvat]: Rezolvare (1/3) — Pas cu pas

1. Depășirile $x = \sum_t I_t$, rata $\hat{\pi} = x/T$; Chronos-2 direct: $x = 45$, $\hat{\pi} = 1{,}67\%$ față de 26,9 așteptate: LR_{uc} dă $p = 0{,}001$, acoperirea corectă este respinsă.
2. Independența: numărați tranzițiile n_{ij} ale $I_{t-1} \to I_t$ și testați $\pi_{01} = \pi_{11}$ (depășiri grupate: Christoffersen și DQ); hibridul: $p = 0{,}021$, depășirile se grupează.
3. FZ0 din simularea istorică, FHS și pierderea FZ0, cu v = VaR 2,5%, e = ES 2,5%, $\alpha = 0{,}025$; hibrid 1,084 față de FHS 0,999, DM +2,65.
4. $\mathrm{DQ} = \hat{\beta}^\top X^\top X \hat{\beta}/(0{,}01 \cdot 0{,}99) \sim \chi^2_6$, X: cei șase regresori, $\hat{\beta}$: coeficienții lor OLS (Ordinary Least Squares): hibrid $p = 0{,}001$, FHS $p = 0{,}021$.

B3 [Rezolvat]: Rezolvare (2/3) — Grafic



- Depășirile VaR 1% cumulate ale celor șase modele; negru întrerupt: 1% din zile
- Chronos-2 direct și GARCH-t sînt deasupra dreptei așteptate (prea înguste)
  - Hibrizii o urmează în medie, dar în trepte în 2020 și 2022 (grupări)
- Interpretare: hibridul preia coada din date prin FHS, deci trece testul de acoperire
  - Volatilitatea prognozată este zgomotoasă: uneori prea mică (depășiri grupate, p independență $= 0{,}021$), alteori prea mare (FZ0 mai mare)

MFM_ch14_seminar

B3 [Rezolvat]: Rezolvare (3/3) — Tabel

Model	Depășiri	Rata (%)	p Kupiec	p indep.	FZ0	DM față de FHS
HS	37	1,37	0,065	0,001	1,309	+2,46
GARCH-t	44	1,63	0,002	0,205	1,026	+1,34
FHS	30	1,11	0,559	0,045	0,999	–
Cuantile Chronos-2	45	1,67	0,001	0,223	1,147	+3,16
FHS + vol. Chronos-2	25	0,93	0,705	0,021	1,084	+2,65
FHS + vol. TimesFM	27	1,00	0,989	0,029	1,059	+1,89

- Chronos-2 direct: 1,67% depășiri, p Kupiec = 0,001: cuantilele lui sînt prea înguste; hibridul: 0,93%, $p = 0{,}705$
- Calibrat în medie, dar cu depășiri grupate (p independență = 0,021, p DQ = 0,001) și o pierdere FZ0 semnificativ mai mare decît FHS (DM +2,65)

MFM_ch14_sem_partB

C1 [Propus]: Critica unui răspuns AI — răspunsul

- Un student îi cere unui asistent AI: „Fă backtesting pentru VaR 1% la o zi, zero-shot, cu Chronos-2 pentru S&P 500 pe ultimele 1 500 de zile, cu un context de 512 zile”
- Asistentul răspunde cu acest cod (spx: închideri zilnice, fără importuri):

```
pipe = BaseChronosPipeline.from_pretrained('amazon/chronos-2')
r = 100 * np.log(spx).diff().dropna().to_numpy()      # daily returns, %
test = np.arange(len(r) - 1500, len(r))
q01 = np.empty(len(test))
for i, t in enumerate(test):                          # forecast of r[t]
    ctx = torch.tensor(r[t - 511:t + 1], dtype=torch.float32)   # 512 days
    q, _ = pipe.predict_quantiles([ctx[None, :]], prediction_length=1,
                                  quantile_levels=[0.01])
    q01[i] = q[0][0, 0, 0].item()
x = int((r[test] < q01).sum())                        # breaches
print(f"VaR 1% = {q01.mean():.2f}%, breaches = {x}")
```

- „Rezultat: VaR 1% $= -2{,}41\%$, depășiri $= 5$, adică 0,33% față de 15 așteptate: modelul este prudent, deci trece ușor testul Kupiec”

◂ Înapoi ▸ Lista exercițiilor

C1 [Propus]: Critica unui răspuns AI — Cerință

- **Model**: A2 și B3 [Rezolvat]
- **Întrebarea**: ce trei erori apar în cod și în concluzie (discutați separat contaminarea antrenării)?
- **Date**: codul și rezultatul de pe slide-ul anterior; închiderile zilnice ale S&P 500, ultimele 1 500 de zile, Chronos-2 zero-shot, context de 512 zile
- **Cerințe**
  1. Pentru fiecare eroare, precizați ce este greșit și de ce este important.
  2. Precizați cum ați detecta-o: o verificare a indicilor, convenția VaR din Capitolul 7 sau statistica testului.
  3. Dați linia sau concluzia corectată și calculați statistica Kupiec și p-value-ul pentru 5 depășiri în 1 500 de zile.
  4. Treceți cele trei erori în `AI_ERRORS.md`: ce s-a cerut, ce a răspuns AI-ul, ce era greșit, cum ați identificat eroarea, varianta corectă.
- **Raportați**: cele trei erori cu detectarea și corectarea lor, statistica Kupiec, înregistrarea din jurnal
- **Interpretare**
  - sînt „prea puține depășiri" o veste bună pentru un manager de risc?
  - dar pentru autoritatea de supraveghere?
  - sînt ultimele 1 500 de zile un test fără contaminare pentru un model publicat în toamna lui 2025?

B4: date și direcția comparației

- Folosiți coloana close din BET.csv și adjusted close din TVBETETF.RO.csv. Aliniați întîi datele prețurilor, apoi calculați randamentele.
- Alternativa unilaterală este că LSTM-ul candidat are pierdere medie de prognoză mai mică decît prognoza zero, reperul din B1. Produceți un tabel de modele ca în B1 pentru Bitcoin și unul separat pentru BET (*Bucharest Exchange Trading*, indicele de referință al Bursei de Valori București).
- Un semnal statistic nu stabilește singur profitabilitatea netă. Pentru o afirmație de tranzacționare, specificați regula poziției și rulajul (*turnover*), calculați costul la care profitul net devine zero (*break-even*) și comparați-l cu un cost presupus explicit; tranzacționabilitatea nu rezultă din valoarea p.

B4 [Propus]: Bitcoin și BET: căutarea unui semnal — Cerință

- **Model**: B1 [Rezolvat]
- **Întrebările**
  - se mențin rezultatele din B1 pentru Bitcoin și pentru indicele românesc BET?
  - rămîne vreun semnal BET după controlul pentru prețurile neactualizate (*stale*)?
- **Date**: `ch14_returns_btc.csv` (2 818 zile calendaristice din 01.01.2019) și `ch14_returns_bet.csv` (2 675 de zile din 04.01.2016), aceleași coloane ca în B1; ETF-ul (*Exchange-Traded Fund*, fond tranzacționat la bursă) pe BET `TVBETETF.RO` (preț ajustat)
- **Cerințe**
  1. Repetați pașii 1–2 din B1 pentru ambele piețe și reprezentați intervalele R^2 alăturat.
  2. Pentru LSTM pe BET, calculați valorile p ale DM unilateral și Clark–West, Romano–Wolf și SPA pe cele șase modele BET și p ajustat Holm pe cele 18 teste DM unilaterale (șase modele $\times$ trei piețe).
  3. Calculați autocorelația de ordinul 1 a BET și corelația prognozei LSTM cu r_{t-1} (Lo & MacKinlay (1990)).
  4. Aliniați prețurile BET și ETF pe zilele comune, apoi aplicați semnalul LSTM randamentului ETF din aceeași zi și din ziua următoare (R^2, p Clark–West).
- **Raportați**: graficul cu două panouri, cele două tabele și cifrele de la pașii 2–4
- **Interpretare**
  - ce piață are randamente zilnice predictibile?
  - este semnalul tranzacționabil după costuri (Capitolul 2)?

◂ Înapoi ▸ Lista exercițiilor

Pasul pentru proiectul de echipă

- Un activ nediscutat la seminar (o acțiune BVB, un criptoactiv, un curs valutar sau o marfă)
- Volatilitatea sau VaR 1% al lui: un model fundațional zero-shot față de un reper GARCH; contextul se încheie în ziua dinaintea prognozei; perioada de după publicare raportată separat
- Testul DM și, pentru VaR, testele de acoperire și independență
- Un folder Quantlet (cod, Metainfo, grafice); utilizarea AI în `AI_USE.md`, erorile identificate în `AI_ERRORS.md`; pachetele de extindere sînt opționale
- Criterii de evaluare: întrebarea, datele și convențiile, metoda, incertitudinea (intervale, robustețe), limitele

Pachete de extindere: ce parcurgeți înainte

Pachetul	Exerciții	Condiții prealabile
1. Memoria rețelelor recurente	A5	A4
2. Tokeni și grile de cuantile	A6, A7	A1, A3
3. Riscul de coadă și calibrare	B5, A8, B6 [Rezolvat], B7	A1, B3
4. Puterea testelor	A9, B8	A2, B2
5. Prognoza volatilității	A10 [Rezolvat], B9, B10	B2
6. Practica deep learning	B11 [Rezolvat]	B1
7. Proiecte	C2, C3	B3; studiul de caz din curs

- Parcurgeți exercițiile în ordine în cadrul fiecărui pachet; fiecare exercițiu indică modelul rezolvat, iar în notebook secțiunea poartă numele exercițiului
- C2 și C3 sînt posibile puncte de plecare pentru proiectul de echipă

A5 [Propus]: Starea celulei LSTM ca AR(1) — Cerință

- **Model**: A4 [Rezolvat]
- **Întrebările**
  - cît timp își amintește un LSTM cu poarta de uitare constantă?
  - poate reproduce memoria lungă?
- **Date**: celula $c_t = fc_{t-1} + ig_t$, $f = \sigma(b_f)$ și i constante, g_t i.i.d. cu media 0 și varianța σ_g^2 (distinctă de funcția sigmoid σ); funcția de autocorelație ACF (*AutoCorrelation Function*) a lui $|r_t|$ pentru S&P 500, din 2000: 0,289 (lag 1), 0,132 (lag 50), 0,059 (lag 150)
- **Cerințe**
  1. Folosiți soluția staționară (inițializați c_0 din distribuția invariantă). Arătați că c_t este un AR(1) staționar pentru $0 < f < 1$ și derivați $\mathrm{Var}(c_t)$ și $\mathrm{Corr}(c_t, c_{t-k})$.
  2. Calculați timpul de înjumătățire $\ln 0{,}5 / \ln f$ pentru $b_f = 1$ și pentru $f = 0{,}99$, termenul liber care dă $f = 0{,}99$ și f^{250} pentru $b_f = 5$.
  3. Ajustați $\rho_k = \rho_1 f^{k-1}$ cu $f^{49} = \rho_{50}/\rho_1$ și prognozați ρ_{150}.
- **Raportați**: derivările și cifrele de la pașii 2–3
- **Interpretare**
  - de ce niciun $f < 1$ fix nu poate reproduce memoria lungă?
  - ce structură alternativă îi trebuie rețelei?

◂ Înapoi ▸ Lista exercițiilor

A6 [Propus]: O zi de crah în context — Cerință

- **Model**: A3 [Rezolvat]
- **Întrebarea**: face un crah ca zilele obișnuite să pară calme pentru Chronos?
- **Date**: scalarea din A3; S&P 500 la 16 martie 2020: randamentul de $-12{,}77\%$ intră într-un context de 512 zile și iese o zi de 0,51%; înaintea crahului $s_{\text{vechi}} = 0{,}741\%$
- **Cerințe**
  1. Înlocuiți ultima valoare din contextul A3 cu $-9{,}0$ și recalculați s, z, j și factorul $s_{\text{vechi}}/s_{\text{nou}}$.
  2. Aplicați formula din A3 cazului S&P 500 și calculați s_{nou} și factorul.
  3. Calculați valoarea scalată a unei zile obișnuite de 0,4% înainte și după crah, în ambele cazuri.
- **Raportați**: s, factorul și z al zilei de 0,4% în ambele cazuri
- **Interpretare**: face un context de 512 zile ca zilele obișnuite să pară calme după un crah?

◂ Înapoi ▸ Lista exercițiilor

Niveluri de cuantilă și praguri de pierdere

- La pasul 1 din A7, u este probabilitatea care indexează q_u.
- La pașii 2–3, u este un prag de pierdere în puncte procentuale, iar N_u este numărul pierderilor peste prag.
- Date pentru estimarea cozii: `GSPC.INDX.csv`, coloana `close`. Calculați $r_t = 100 \ln(P_t/P_{t-1})$ fără rotunjire; folosiți ultimele 1 000 de randamente pînă la data indicată și $L = -r$.
- Folosiți percentila empirică 95 cu interpolare liniară (`numpy.quantile`, `method="linear"`); estimați excesele strict pozitive cu `scipy.stats.genpareto.fit`, fixînd `floc=0`.
- Pragul este exprimat în unități de randament, nu este probabilitatea 0,95. De exemplu, un prag de pierdere de 2% și o pierdere de 3% dau un exces de un punct procentual.

A7 [Propus]: Coada limitată a grilei; o coadă GPD — Cerință

- **Model**: A1 [Rezolvat]
- **Întrebările**
  - în ce sens deplasează grila ES?
  - ce indică o coadă groasă estimată?
- **Date**: grila din A1; pierderile S&P 500 $L = -r$ pe ultimele 1 000 de zile pînă la 18.09.2026; depășiri peste prag: u = cuantila de 95% a lui L, excesele $L - u \sim \text{GPD}(\xi, \beta)$ din instrumente pentru pachetele de extindere, $\xi < 1$ (McNeil & Frey (2000)); n: mărimea eșantionului, N_u: numărul de excese
- **Cerințe**
  1. Demonstrați că păstrarea $q_u = q_{0,01}$ pentru $u < 0{,}01$ nu mărește niciodată ES 2,5% și arătați că interpolarea liniară pe $[0{,}01; 0{,}025]$ poate produce erori în ambele sensuri.
  2. Derivați $\text{ES}_\alpha = \frac{\text{VaR}_\alpha}{1-\xi} + \frac{\beta - \xi u}{1-\xi}$ din funcția excesului mediu a GPD.
  3. Estimați ξ, β prin verosimilitate maximă și calculați $\text{VaR}_\alpha = u + \frac{\beta}{\xi}[(n\alpha/N_u)^{-\xi} - 1]$ pentru 1% și 2,5%, și ES 2,5%.
- **Raportați**: u, N_u, $\hat{\xi}$, $\hat{\beta}$, VaR, ES și ES/VaR
- **Interpretare**: de ce este ES/VaR mai mare aici decît în A1?

Înapoi Lista exercițiilor

B5 [Propus]: Bitcoin și VaR 10% pe toate cele trei piețe — Cerință

- **Model**: B3 [Rezolvat]
- **Întrebările**
  - se mențin concluziile din B3 pentru Bitcoin?
  - sînt cuantilele modelelor fundaționale prea înguste chiar și la 10%?
- **Date**: `ch14_risk_btc.csv` (2 818 zile din 01.01.2019); coloanele `<model>|VaR10` pentru HS, FHS și cuantilele directe Chronos-2, Chronos-Bolt și TimesFM-2.5 din `ch14_risk_<piață>.csv`
- **Cerințe**
  1. Repetați pașii 1–3 din B3 pentru Bitcoin.
  2. Pentru VaR 10%, calculați ratele de depășire și testul Kupiec cu $\alpha = 10\%$ pentru cele cinci modele pe cele trei piețe.
  3. Reprezentați ratele față de intervalul de 95% al unui model corect.
- **Raportați**: tabelul pentru Bitcoin, graficul și un tabel 3×5 cu rate și p Kupiec
- **Interpretare**: este îngustimea cuantilelor modelelor fundaționale specifică nivelului de 1%?

◂ Înapoi ▸ Lista exercițiilor

Riscul de coadă pentru distribuția Normală cu medie nenulă

- Pentru randamentul predictiv $r = m + \sigma Z$, cu Z distribuit Normal standard, pierderea este $L = -r$. O medie m pozitivă deplasează distribuția pierderii spre stînga.
- Notăm $z_{1-\alpha} = \Phi^{-1}(1-\alpha) > 0$. Atunci $\mathrm{VaR}_\alpha = -m + \sigma z_{1-\alpha}$ și $\mathrm{ES}_\alpha = -m + \sigma\varphi(z_{1-\alpha})/\alpha$.
- Formula ES rezultă din media cozii superioare Normale: integrala lui $z\varphi(z)$ deasupra lui $z_{1-\alpha}$ este $\varphi(z_{1-\alpha})$; împărțim la probabilitatea cozii α.
- În A8, mediana furnizează m, iar distanța dintre mediană și prima decilă furnizează σ. Folosiți B5 pentru comparația ratelor empirice de depășire.

A8 [Propus]: Un „VaR 1%" din decile — Cerință

- **Model**: A1 [Rezolvat]; după B5
- **Întrebarea**: ce raportează un model care furnizează doar decile cînd i se cere VaR 1%?
- **Date**: Chronos-Bolt (small), aceeași zi și același context ca în A1: $q_{0,1} = -0{,}691\%$, $q_{0,5} = 0{,}071\%$
- **Cerințe**
  1. Precizați valoarea returnată cînd se cere VaR 1% dintr-o grilă de decile limitată la cel mai mic nivel.
  2. Presupuneți o distribuție predictivă Normală cu mediana $q_{0,5}$ și $\hat{\sigma} = (q_{0,5} - q_{0,1})/z_{0,9}$, cu $z_{0,9} = 1{,}2816$ cuantila de 90% a distribuției Normale standard; calculați VaR 1% și ES 2,5%.
  3. Comparați cu Chronos-2 din A1 și cu ratele de depășire ale VaR 10% din B5.
- **Raportați**: cele trei cifre și comparația
- **Interpretare**: ce test ar decide dacă extrapolarea Normală funcționează?

Înapoi Lista exercițiilor

B6 [Rezolvat]: Calibrarea întregii distribuții predictive — Cerință

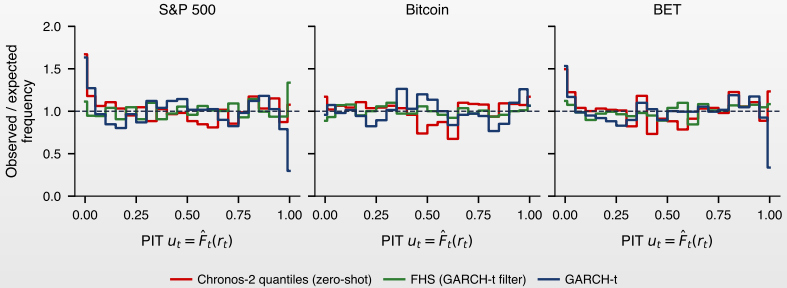
- **Întrebarea**: este calibrată întreaga distribuție predictivă Chronos-2 sau doar centrul ei?
- **Date**: S&P 500, 2 693 de zile; grila Chronos-2 directă de 21 de cuantile (1%, 5%, . . . , 99%) din `ch14_returns_sp500.csv`; FHS și GARCH-t evaluate pe aceeași grilă
- **Cerințe**
  1. Calculați PIT u_t din instrumente pentru pachetele de extindere prin interpolare liniară, cenzurat în afara $[q_{1\%}, q_{99\%}]$, și aplicați χ^2 Pearson frecvențelor din cele 22 de intervale.
  2. Aplicați testul LR Berkowitz (Berkowitz (2001)) pentru $\mu = 0$, $\sigma = 1$ lui $z_t = \Phi^{-1}(u_t)$, cu o verosimilitate Normală cenzurată (χ^2_2).
  3. Calculați scorul cuantilic QS_t din instrumente pentru pachetele de extindere și DM pentru Chronos-2 minus FHS, total, în centru și în cozi.
- **Raportați**: histograma PIT și un tabel: ponderile din cozi, $\hat{\mu}/\hat{\sigma}$, p Berkowitz, p χ^2, QS
- **Interpretare**
  - este eșecul în centru sau în cozi?
  - de ce cenzurăm în loc să extrapolăm?

◂ Înapoi ▸ Lista exercițiilor

B6 [Rezolvat]: Rezolvare (1/3) — Pas cu pas

1. Dacă $q_{1\%,t} \le r_t \le q_{99\%,t}$, u_t interpolează nivelurile între cele două cuantile vecine; altfel se notează doar „sub" sau „peste"; $\chi^2 = \sum_{j=1}^{22} (O_j - Tp_j)^2/(Tp_j)$, O_j = frecvența observată și p_j = lățimea intervalului j.
2. $\ell(\mu,\sigma) = \sum_{\text{interior}} \log \frac{1}{\sigma}\varphi\left(\frac{z_t-\mu}{\sigma}\right) + n_L \log \Phi\left(\frac{c_L-\mu}{\sigma}\right) + n_H \log\left[1-\Phi\left(\frac{c_H-\mu}{\sigma}\right)\right]$, $c_L = \Phi^{-1}(0{,}01)$, $c_H = \Phi^{-1}(0{,}99)$; φ, Φ: densitatea și funcția de repartiție ale distribuției Normale standard; n_L, n_H: zilele sub și peste grilă; $\mathrm{LR} = 2[\ell(\hat\mu,\hat\sigma) - \ell(0,1)]$.
3. QS pe toate cele 21 de niveluri, pe cele 17 niveluri centrale și pe cele 4 niveluri din cozi; DM cu variantă Newey–West.

B6 [Rezolvat]: Rezolvare (2/3) — Grafic



- Frecvența observată raportată la cea așteptată în cele 22 de intervale ale grilei; 1 = calibrare perfectă; vîrfuri la capete: randamente în afara $[q_{1\%}, q_{99\%}]$
- Interpretare: Chronos-2 este mai slab decît FHS atît în centru, cît și în cozi, mai ales în cozi; cenzurarea evită inventarea unei cozi pe care grila nu o are

MFM_ch14_calibration

B6 [Rezolvat]: Rezolvare (3/3) — Tabel

Model	Sub $q_{1\%}$ (%)	Peste $q_{99\%}$ (%)	$\hat{\mu}$ / $\hat{\sigma}$	p Berkowitz	χ^2 p	QS
Chronos-2	1,67	1,08	−0,035 / 1,030	0,019	0,004	0,5101
FHS	1,11	1,34	+0,014 / 0,990	0,599	0,747	0,5062
GARCH-t	1,63	0,30	−0,041 / 0,976	0,023	<0,001	0,5063

- Chronos-2: prea multe randamente sub $q_{1\%}$, $\hat{\sigma} > 1$, ambele teste resping; FHS trece; QS Chronos-2 minus FHS: DM +3,06 (centru +2,51, cozi +3,65)

MFM_ch14_calibration

Ce cuantilă de calibrare folosim în B7?

- La fiecare prognoză, ordonăm doar cele 250 de scoruri E_t anterioare (definite în B7). Randamentul curent intră în calibrare abia după evaluarea prognozei sale.
- Inițializăm nivelul adaptiv la nivelul țintă. Alegem statistica de ordine $k = \lceil 251(1-\alpha_t) \rceil$, numărînd de la 1, fără interpolare.
- Dacă indicele depășește 250 sau nivelul adaptiv este nepozitiv, corecția este $+\infty$; dacă indicele este sub 1 sau nivelul este cel puțin unu, corecția este $-\infty$. Nu trunchiem nivelul adaptiv.
- La ținta de 1%, indicele inițial este 249: cuantila interpolată implicit de o bibliotecă ar implementa altă regulă. Actualizarea din B7 se aplică după observarea depășirii.

B7 [Propus]: Recalibrarea conformală a Chronos-2 — Cerință

- **Model**: B3 și B6 [Rezolvat]
- **Întrebarea**: poate recalibrarea conformală să facă cuantilele directe Chronos-2 un VaR 1% utilizabil fără filtru GARCH?
- **Date**: cuantilele Chronos-2 directe din B3 pe cele trei piețe; evaluare pe zilele de după primele 250 de zile de calibrare
- **Cerințe**
  1. Corectați cuantilele de 1% și 2,5% prin inferență conformală adaptivă (Gibbs & Candès (2021)): scorurile $E_t = \hat{q}_{\alpha,t} - r_t$ pe ultimele 250 de zile, cuantila corectată $\hat{q}_{\alpha,t} - \hat{Q}_{1-\alpha_t}(E)$, actualizarea $\alpha_{t+1} = \alpha_t + \gamma(\alpha - \mathrm{err}_t)$, $\gamma = 0{,}001$; $\hat{q}_{\alpha,t}$: cuantila directă; $\hat{Q}_{1-\alpha_t}(E)$: cuantila empirică $(1-\alpha_t)$ a scorurilor; $\mathrm{err}_t = 1$ cînd randamentul cade sub cuantila corectată.
  2. Repetați cu $\gamma = 0$ (calibrare pe fereastră mobilă, cu nivel fix); deplasați ES 2,5% cu corecția VaR 2,5%.
  3. Testați FHS, Chronos-2 direct și ambele variante conformale: depășiri, acoperire condiționată, DQ, FZ0 și DM față de FHS.
- **Raportați**: un grafic cu depășirile pe metode și piețe și un tabel
- **Interpretare**: de ce poate pierde comparația FZ0 un model cu acoperire corectă?

◂ Înapoi ▸ Lista exercițiilor

A9 [Propus]: Lungimea necesară a ferestrei de după publicare — Cerință

- **Model**: A2 [Rezolvat]
- **Întrebarea**: de cîte zile de după publicare are nevoie un backtest VaR 1% ca să detecteze o rată de depășire de 1,67% sau 2%?
- **Date**: testul Kupiec bilateral din A2 la 5%; rata reală de depășire $\pi = 1{,}67\%$ sau 2%
- **Cerințe**
  1. Calculați puterea exactă față de $\pi = 1{,}67\%$ pentru $T = 1\,000$ și $T = 2\,693$.
  2. Găsiți cel mai mic T cu puterea de cel puțin 80% și T-ul de la care rămîne peste 80%.
  3. Repetați pentru $\pi = 2\%$: puterea la $T = 220$ și T-ul pentru o putere de 80%.
- **Raportați**: un grafic al puterii în funcție de T și cifrele
- **Interpretare**
  - de ce nu este puterea monotonă în T?
  - cîți ani cere un backtest fără contaminare?

◂ Înapoi ▸ Lista exercițiilor

B8: varianța diferenței înainte/după publicare

- Estimați separat variațele pe termen lung înainte și după publicare. Pentru planificare, tratați perioadele ca blocuri independente; nu combinați variațele lor.
- Cu variațele pe termen lung $v_{\mathrm{pre}}, v_{\mathrm{post}}$, varianța diferenței este aproximată prin $v_{\mathrm{pre}}/n_{\mathrm{pre}} + v_{\mathrm{post}}/n_{\mathrm{post}}$.
- Pentru limita de la pasul 3, folosiți aceeași sumă de cuantile $z_{0,975} + z_{0,8}$ ca pentru efectul minim detectabil (MDE, *Minimum Detectable Effect*).

B8 [Propus]: Un test formal înainte/după publicare — Cerință

- **Model**: B2 și A2 [Rezolvat]
- **Întrebarea**: scade avantajul Chronos-2 față de log-HAR după publicarea ponderilor, așa cum ar implica ipoteza memorării?
- **Date**: `ch14_rv.csv` (B2); data-limită 3 noiembrie 2025, prima zi de luni după ce toate cele trei seturi de ponderi erau publice (Chronos-2: 30.10.2025); 761 de zile înainte, 218 după
- **Cerințe**
  1. Calculați $d_t = \text{QLIKE(Chronos-2)} - \text{QLIKE(log-HAR)}$ și mediile lui înainte și după data-limită.
  2. Regresați d_t pe o constantă și $1\{t \geq 3.11.2025\}$ cu erori Newey–West și testați $\beta_1 = 0$ la 5%.
  3. Calculați efectul minim detectabil $(z_{0,975} + z_{0,8})\,\mathrm{se}_{HAC}(\hat{\beta}_1)$ (z_p: cuantile ale distribuției Normale standard, 1,96 și 0,84: un test bilateral la 5% cu puterea 80%) și limita lui cînd fereastra de după publicare crește nelimitat.
- **Raportați**: un grafic al lui d_t cu ambele medii și intervalele lor; mediile, $\hat{\beta}_1$, t, p, MDE
- **Interpretare**: arată un $\hat{\beta}_1$ nesemnificativ că Chronos-2 nu a memorat perioada de test?

◂ Înapoi ▸ Lista exercițiilor

O grilă de cuantile nu este o distribuție completă

- Pentru media calculată din grila de cuantile, extindeți nivelurile cuantilelor la 0 și 1 păstrînd constante cuantilele extreme disponibile. Astfel fixați cozile neobservate; este o ipoteză, nu informație furnizată de model.
- Aplicați mai întîi transformarea țintei fiecărei cuantile, apoi integrați valorile transformate după nivelurile cuantilelor prin regula trapezelor. Pentru ținte în log-volatilitate, exponențierea unei prognoze log medii ar fi un alt calcul.
- Media simplă a cuantilelor acordă aceeași pondere fiecărui punct al grilei. În general diferă de integrarea după probabilitate, mai ales pe o grilă neuniformă; precizați regula pentru cozi și pentru integrare.

A10 [Rezolvat]: Funcționala elicitată de QLIKE — Cerință

- **Întrebările**
  - ce prognoză punctuală a variației răsplătește QLIKE?
  - ce pierdere produce folosirea medianei?
- **Date**: QLIKE $L(y, h)$ din Notația (2/2), cu $y = \mathrm{RV}$, h prognoza; Chronos-2 pe $\log \mathrm{RV}$ furnizează cuantilele q_u; raportul mediu dintre $\exp(q_{0,5})$ și $\int_0^1 \exp(q_u)\, du$ pe cele 979 de zile din B2: 0,8095
- **Cerințe**
  1. Minimizați $E[L(Y, h) \mid \mathcal{F}]$ după $h > 0$ și arătați că minimul este atins în $E[Y \mid \mathcal{F}]$.
  2. Presupuneți $\log Y \mid \mathcal{F} \sim N(\mu, s^2)$; derivați mediana, media, raportul lor și excesul de QLIKE așteptat al prognozei prin mediană.
  3. Din raportul Chronos-2 deduceți un s comun efectiv și excesul de QLIKE.
- **Raportați**: derivările, s și excesul
- **Interpretare**: de ce este exp(mediana) lui log RV o prognoză deplasată a volatilității?

◂ Înapoi ▸ Lista exercițiilor

A10 [Rezolvat]: Rezolvare

1. $E[L] = E[Y]/h - E[\log Y] + \log h - 1$; $\partial_h E[L] = -E[Y]/h^2 + 1/h = 0 \Rightarrow h^* = E[Y \mid \mathcal{F}]$; derivata a doua $1/h^{*2} > 0$: QLIKE este consistentă pentru medie (Gneiting (2011)).
2. Mediana e^{μ}, media $m = e^{\mu + s^2/2}$, raportul $e^{-s^2/2}$; cu $h = me^{-s^2/2}$: $E[L(h)] - E[L(m)] = e^{s^2/2} - 1 - s^2/2 > 0$.
3. $s = \sqrt{-2 \ln 0{,}8095} = 0{,}650$; excesul $1/\text{raport} - 1 + \ln(\text{raport}) = 0{,}0240$; observat: QLIKE 0,249 cu mediana față de 0,223 cu media.

⊡ Interpretare: transformările monotone păstrează cuantilele, deci exp(mediana lui $\log \mathrm{RV}$) este mediana lui RV, cu circa 20% sub medie; QLIKE răsplătește media

B9 [Propus]: Mediana față de medie — Cerință

- **Model**: A10 și B2 [Rezolvat]
- **Întrebările**
  - cît schimbă alegerea funcționalei (mediana sau media) QLIKE-ul unui model fundațional?
  - cît îl schimbă lungimea contextului?
- **Date**: `ch14_rv.csv`: coloanele `Chronos-2`, `Chronos-2|median`, `TimesFM-2.5`, `TimesFM-2.5|median`; `ch14_ctx.csv`: Chronos-2 cu contexte de 32, 64, 128, 256 și 512 zile, în aceleași 979 de zile
- **Cerințe**
  1. Pentru ambele modele calculați QLIKE și raportul media prognozei / media RV pentru mediană și pentru media din grilă.
  2. Testați mediana față de medie pentru Chronos-2 cu DM pe diferențele QLIKE.
  3. Calculați QLIKE Chronos-2 pentru fiecare lungime a contextului și comparați cu log-HAR.
- **Raportați**: un grafic cu două panouri, tabelul de la pasul 1, statistica DM
- **Interpretare**: ce indică efectul contextului despre memoria volatilității?

◂ Înapoi ▸ Lista exercițiilor

B10: istoricul de antrenare și zilele de evaluare

- `ch14_rv.csv` conține prognozele de evaluare, nu întregul istoric necesar estimării. Prețurile SPY la cinci minute încep la 12 octombrie 2020 în fișierul intrazilnic.
- Construiți RV separat în fiecare zi din pătratele schimbărilor logaritmice de preț, cu randamente exprimate în procente; folosiți grila 09:30–15:55 (78 de bare, ora New York-ului), primul randament din deschiderea și închiderea barei de 09:30 și doar zilele cu grila completă. Obțineți varianța zilnică în procente la pătrat.
- Folosiți istoricul pentru cel mult 500 de rînduri utilizabile în regresie și cele 22 de zile anterioare cerute de media lunară; la început, folosiți toate rîndurile după primele 22 de observații. Evaluați numai pe datele din fișierul prognozelor și comparați modelele pe date identice.

B10 [Propus]: Deplasarea de anticipare în regresori — Cerință

- **Model**: B2 [Rezolvat]
- **Întrebările**
  - cît de mare este cîștigul aparent produs de o anticipare de o zi (*look-ahead*)?
  - l-ar putea egala un model corect?
- **Date**: $\log \mathrm{RV}$ al SPY ca în B2; log-HAR pe o fereastră mobilă de 500 de zile cu regresorii $\log \mathrm{RV}_{t-1}$, $\frac{1}{5}\sum_{j=1}^{5} \log \mathrm{RV}_{t-j}$, $\frac{1}{22}\sum_{j=1}^{22} \log \mathrm{RV}_{t-j}$; prognoza $\exp(\hat{y}_t + \hat{s}^2/2)$
- **Cerințe**
  1. Construiți o variantă greșită în care mediile pe 5 și 22 de zile iau $j = 0, \ldots, 4$ și $j = 0, \ldots, 21$, deci includ ziua prognozată t.
  2. Calculați QLIKE pentru log-HAR corect, log-HAR greșit și Chronos-2 pe aceleași 979 de zile.
  3. Testați varianta greșită față de cea corectă cu DM și scrieți un test pentru scurgerea de informație (*leakage*) în propriul cod.
- **Raportați**: cele trei valori QLIKE, cîștigul relativ și testul DM
- **Interpretare**: unde se poate ascunde aceeași eroare într-un flux de lucru de deep learning?

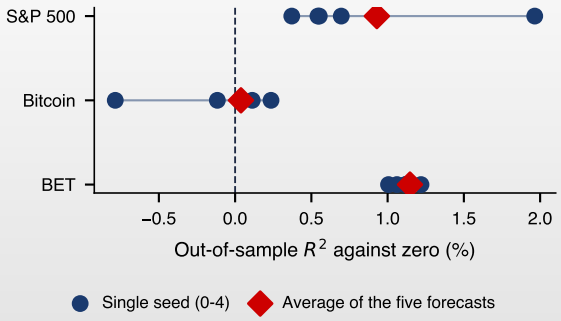
◂ Înapoi ▸ Lista exercițiilor

B11 [Rezolvat]: Variabilitatea între seed-uri — Cerință

- **Întrebarea**: cît schimbă doar seed-ul aleator R^2 în afara eșantionului al unui LSTM?
- **Date**: LSTM-ul pentru randamente din Capitolul 14 (ultimele 20 de randamente standardizate, 16 unități ascunse, MSE, Adam, oprire timpurie); cinci rețele identice cu excepția seed-urilor, coloanele `lstm_s0`, ..., `lstm_s4` și media lor `lstm` din `ch14_returns_<piață>.csv`
- **Cerințe**
  1. Calculați R^2 în afara eșantionului față de zero pentru fiecare seed, pe cele trei piețe.
  2. Calculați R^2 pentru media celor cinci prognoze.
  3. Comparați dispersia dintre seed-uri cu mărimea lui R^2.
- **Raportați**: un grafic și un tabel 3×6 cu R^2 (%)
- **Interpretare**
  - ce concluzie ar fi dat o singură rulare pe Bitcoin?
  - de ce este media seed-urilor o formă de regularizare?

Înapoi Lista exercițiilor

B11 [Rezolvat]: Rezolvare (1/2) — Grafic



- Albastru: seed-urile individuale; roșu: media celor cinci prognoze
- Interpretare: pe Bitcoin, seed-urile variază de la $-0{,}79\%$ la $+0{,}24\%$
  - O singură rulare ar putea „arăta" abilitate sau lipsa ei; pe BET toate cinci sînt de acord
  - Media reduce varianța, ca bagging-ul

MFM_ch14_seminar

B11 [Rezolvat]: Rezolvare (2/2) — Tabel

R^2 (%)	Seed 0	1	2	3	4	Media
S&P 500	+0,55	+0,37	+0,70	+0,54	+1,96	+0,93
Bitcoin	+0,11	+0,03	−0,79	−0,12	+0,24	+0,04
BET	+1,06	+1,01	+1,12	+1,22	+1,12	+1,15

1. $R_s^2 = 1 - \sum_t (r_t - \hat{r}_{t,s})^2 / \sum_t r_t^2$ pentru fiecare seed: coloanele 0–4
2. Media celor cinci prognoze, $\bar{r}_t = \frac{1}{5} \sum_s \hat{r}_{t,s}$: coloana Media; din convexitate, eroarea ei pătratică nu depășește media erorilor seed-urilor
3. Dispersia dintre seed-uri: pe Bitcoin, de la −0,79% la +0,24%, mult peste R^2 mediu (+0,04%); pe S&P 500, de la +0,37% la +1,96%; pe BET, de la +1,01% la +1,22%, mică față de R^2 = +1,15%

MFM_ch14_sem_partB

C2 [Propus]: Un model fundațional zero-shot în locul GARCH pentru BET — Cerință

- **Model**: B3 [Rezolvat]; proiect: ajustarea fină (*fine-tuning*) a unui model fundațional pe acțiuni românești
- **Întrebarea**: un manager de risc propune Chronos-2 pentru VaR 1% și ES 2,5% zilnice ale unei poziții pe BET: acceptați, respingeți sau păstrați ca model alternativ?
- **Date**: `ch14_risk_bet.csv`, 2 675 de zile din 04.01.2016; fereastra de după publicare de la 3.11.2025 (217 zile)
- **Cerințe**
  1. Testați VaR 1% pentru cele șase modele din B3: depășiri, Kupiec, independență și DQ.
  2. Comparați (VaR 2,5%, ES 2,5%) cu FZ0, DM față de FHS și MCS de 90%.
  3. Clasificați primele zece blocuri consecutive, fără suprapunere, de 250 de zile în zonele Basel, excluzînd blocul final incomplet (verde ≤ 4, galben 5–9, roșu ≥ 10 depășiri) și repetați pasul 1 după publicare.
  4. Proiect live opțional: preînregistrați data de început și Chronos-2 direct, cu 512 randamente log zilnice anterioare. Înregistrați VaR pentru ziua următoare înainte de începerea ei, timp de 20 de zile de tranzacționare, apoi testați.
- **Raportați**: o recomandare de o pagină, cu tabelele de la pașii 1–3
- **Interpretare**
  - ce dovezi v-ar face să schimbați modelul?
  - cînd ați face schimbarea?

◂ Înapoi ▸ Lista exercițiilor

C3: date, numărul trăsăturilor și calculul

- Folosiți închiderile zilnice GSPC.INDX pentru suma lunară a pătratelor randamentelor și închiderile ajustate SPY pentru randamentele totale lunare. Istoricul lung al prețurilor este distinct de fișierul prognozelor evaluate din 2016.
- P numără coloanele individuale: 12 000 de trăsături folosesc 6 000 de frecvențe, fiecare producînd o coloană sinus și una cosinus. Folosiți aceleași extrageri pentru numerele imbricate de trăsături.
- Cu numai 12 observații pe fereastră, calculați ridge prin matricea Gram de 12 pe 12, nu prin inversarea unei matrice de 12 000 pe 12 000. Începeți cu cîteva extrageri pentru verificarea dimensiunilor, apoi rulați cele 1 000 cerute.

C3: țintele lunare și observațiile lipsă

- Construiți lag-urile și ținta lunii următoare pe un index lunar complet, fără completarea valorilor lipsă. Calculați volatilitatea ultimelor 12 luni înainte de eliminarea rîndurilor incomplete.
- Indicele prețurilor de consum CPI (*Consumer Price Index*) pentru octombrie 2025 nu a fost publicat. În versiunea datelor (*vintage*) din 30 noiembrie 2025, ultima inflație cunoscută este încă cea din septembrie; în decembrie 2025 inflația din noiembrie nu se poate calcula și nu se imputează. Testul final are 354 de ținte, februarie 1997–august 2026, fără ianuarie 2026.
- Fereastra regresiei conține 12 observații complete și poate acoperi mai mult de 12 luni calendaristice. Ținta rămîne exact luna calendaristică următoare. Folosiți aceleași 1 000 de extrageri ale trăsăturilor aleatoare, seed 0, în toate comparațiile.

C3 [Propus]: Modele complexe pe datele cursului (1/2) — Cerință

- **Model**: B1 pentru etapele evaluării; aici $R^2 = 1 - \mathrm{Var}(\text{eroare})/\mathrm{Var}(R)$ (nota 40 din lucrarea citată mai jos)
- **Întrebarea**: apare pe datele cursului avantajul complexității (Kelly, Malamud & Zhou (2024))?
- **Date**
  - Randamentele totale lunare ale SPY din prețurile ajustate de sfîrșit de lună, din 1993, minus `TB3MS`/1 200 din luna precedentă (FRED, *Federal Reserve Economic Data*)
  - Șapte predictori din Welch & Goyal (2008) datați t: `tbl` = `TB3MS`/100, `lty` = `GS10`/100, `tms` = `lty` − `tbl`, `dfy` = (`BAA` − `AAA`)/100
  - `infl` (creșterea lunară a `CPIAUCSL` cunoscută la sfîrșitul lunii t: versiunea ALFRED (*ArchivaL Federal Reserve Economic Data*, arhiva versiunilor istorice ale seriilor FRED) din acea zi, fără revizii ulterioare), `svar` (suma pătratelor randamentelor zilnice S&P 500), randamentul în exces
- **Cerințe**
  1. Construiți datele lunare și standardizați ca în Secțiunea V.A: randamentele prin abaterea standard necentrată din ultimele 12 luni, predictorii prin abaterea standard pe o fereastră crescătoare, după 36 de luni de inițializare.
  2. Generați 12 000 de trăsături Fourier aleatoare cu $\gamma = 2$ (ec. 20); pentru P de la 2 la 12 000 și $\log_{10} z = -3, \ldots, 3$, estimați ridge fără termen liber pe ferestre mobile de 12 observații complete și calculați randamentul strategiei $\hat{\beta}' S_t R_{t+1}$; mediați pe 1 000 de extrageri.
  3. Estimați regresia liniară cu toți cei șapte predictori (kitchen sink), cu $z = 0^+$ și $z = 10^3$, pe aceleași ferestre.
- **Raportați**: slide-ul următor

◂ Înapoi ▸ Lista exercițiilor

C3 [Propus]: Modele complexe pe datele cursului (2/2) — raport și întrebări

- **Notații**: $R_{t+\mathbf{1}}$: randamentul în exces din luna următoare; G_t: cei șapte predictori
  - S_t: cele P trăsături Fourier aleatoare $\sin(\gamma\,\omega_i' G_t)$, $\cos(\gamma\,\omega_i' G_t)$, cu $\omega_i \sim N(0, I)$ extras o singură dată
  - ridge $\hat\beta(z) = (zI + \frac{\mathbf{1}}{T}\sum_t S_t S_t')^{-\mathbf{1}}\frac{\mathbf{1}}{T}\sum_t S_t R_{t+\mathbf{1}}$, z: penalizarea, $T = 12$: fereastra; randamentul strategiei $\hat\beta' S_t R_{t+\mathbf{1}}$; complexitatea $c = P/T$
- **Raportați**
  - Curbele pentru $\|\hat\beta\|^{\mathbf{2}}$, $R^{\mathbf{2}}$ și raportul Sharpe în funcție de $c = P/T$ (Figurile 7–8)
  - Un tabel în formatul Tabelului I pentru cele două modele liniare și RFF (trăsături Fourier aleatoare, *Random Fourier Features*) cu $c = 1\,000$, $z = 10^{\mathbf{3}}$
  - Coloanele: $R^{\mathbf{2}}$, raportul Sharpe cu t, rapoartele de informație față de SPY și față de modelul liniar cu t, pierderea lunară maximă (în abateri standard), asimetria
- **Interpretare**
  - Regresia dumneavoastră are $c = 7/12$, față de $15/12$ în lucrare: ce prezice curba fără penalizare din Figura 7 pentru $R^{\mathbf{2}}$?
  - Poate raportul Sharpe să fie pozitiv cînd $R^{\mathbf{2}} < 0$? Folosiți descompunerea erorii într-o componentă de corelație și una de scală (nota 5)
  - Cu circa 30 de ani, care este puterea unui test bilateral la 5% pentru un raport de informație de 0,3 ($t \approx$ raport $\times \sqrt{\text{ani}}$)?
- **Extindere pentru proiect**: înlocuiți trăsăturile aleatoare cu reprezentările interne (*embeddings*) ale unui model fundațional preantrenat și repetați pasul 2

Acronime folosite în acest material (1/2)

- **ACF**: AutoCorrelation Function — *funcția de autocorelație*
- **ACI**: Adaptive Conformal Inference — *inferență conformală adaptivă*
- **AI**: Artificial Intelligence — *inteligență artificială*
- **AR**: AutoRegressive (model)
- **BET**: Bucharest Exchange Trading (index) — *indicele de referință al Bursei de Valori București*
- **BVB**: Bursa de Valori București
- **CI**: Confidence Interval — *interval de încredere*
- **CPI**: Consumer Price Index — *indicele prețurilor de consum*
- **CW**: Clark–West (test for nested models) — *testul Clark–West pentru modele imbricate*
- **DM**: Diebold–Mariano (test) — *testul Diebold–Mariano de comparare a prognozelor*
- **DQ**: Dynamic Quantile (test) — *testul cuantilei dinamice*
- **ES**: Expected Shortfall — *pierderea așteptată în coadă*
- **ETF**: Exchange-Traded Fund — *fond tranzacționat la bursă*
- **FHS**: Filtered Historical Simulation — *simulare istorică filtrată*
- **FRED**: Federal Reserve Economic Data — *baza de date economice a Rezervei Federale (St. Louis)*
- **FZ0**: Fissler–Ziegel loss, zero-homogeneous version — *funcția de pierdere Fissler–Ziegel, varianta omogenă de grad zero*
- **GARCH**: Generalised AutoRegressive Conditional Heteroskedasticity — *heteroscedasticitate condiționată autoregresivă generalizată*
- **GPD**: Generalised Pareto Distribution — *distribuția Pareto generalizată*

Acronime folosite în acest material (2/2)

- **HAC**: Heteroskedasticity and Autocorrelation Consistent (standard errors) — *erori standard robuste la heteroscedasticitate și autocorelație*
- **HAR**: Heterogeneous AutoRegressive (model) — *model autoregresiv heterogen*
- **HS**: Historical Simulation — *simulare istorică*
- **LR**: Likelihood Ratio (test) — *testul raportului de verosimilitate*
- **LSTM**: Long Short-Term Memory (network) — *rețea cu memorie pe termen lung și scurt*
- **MCS**: Model Confidence Set — *mulțimea de încredere a modelelor*
- **MDE**: Minimum Detectable Effect — *efectul minim detectabil*
- **MSE**: Mean Squared Error — *eroarea pătratică medie*
- **OLS**: Ordinary Least Squares — *metoda celor mai mici pătrate*
- **PIT**: Probability Integral Transform — *transformarea integrală de probabilitate*
- **QLIKE**: Quasi-LIKElihood (loss function) — *funcția de pierdere de tip cvasi-verosimilitate*
- **QS**: Quantile Score — *scorul cuantilic*
- **RFF**: Random Fourier Features — *trăsături Fourier aleatoare*
- **RNN**: Recurrent Neural Network — *rețea neuronală recurentă*
- **RV**: Realised Variance — *varianța realizată*
- **SPA**: Superior Predictive Ability (test) — *testul abilității predictive superioare*
- **SUA**: Statele Unite ale Americii
- **VAR**: Vector AutoRegression — *model vector autoregresiv*

Bibliografie (1/3)

- Ansari, A.F., Shchur, O., Küken, J., Auer, A., et al. (2025). Chronos-2: from univariate to universal forecasting. arXiv:2510.15821.
- Ansari, A.F., Stella, L., Turkmen, C., Zhang, X., et al. (2024). Chronos: learning the language of time series. *Transactions on Machine Learning Research*; arXiv:2403.07815.
- Berkowitz, J. (2001). Testing density forecasts, with applications to risk management. *Journal of Business & Economic Statistics*, 19(4), 465–474.
- Christoffersen, P.F. (1998). Evaluating interval forecasts. *International Economic Review*, 39(4), 841–862.
- Clark, T.E., West, K.D. (2006). Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis. *Journal of Econometrics*, 135(1–2), 155–186.
- Clark, T.E., West, K.D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1), 291–311.
- Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2), 174–196.
- Das, A., Kong, W., Sen, R., Zhou, Y. (2024). A decoder-only foundation model for time-series forecasting. *ICML*, PMLR 235; arXiv:2310.10688.
- Diebold, F.X., Gunther, T.A., Tay, A.S. (1998). Evaluating density forecasts with applications to financial risk management. *International Economic Review*, 39(4), 863–883.

Bibliografie (2/3)

- Diebold, F.X., Mariano, R.S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253–263.
- Engle, R.F., Manganelli, S. (2004). CAViaR: conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4), 367–381.
- Gibbs, I., Candès, E. (2021). Adaptive conformal inference under distribution shift. *NeurIPS*; arXiv:2106.00170.
- Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association*, 106(494), 746–762.
- Hansen, P.R. (2005). A test for superior predictive ability. *Journal of Business & Economic Statistics*, 23(4), 365–380.
- Hansen, P.R., Lunde, A., Nason, J.M. (2011). The model confidence set. *Econometrica*, 79(2), 453–497.
- Hochreiter, S., Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780.
- Kelly, B., Malamud, S., Zhou, K. (2024). The virtue of complexity in return prediction. *Journal of Finance*, 79(1), 459–503.
- Kupiec, P.H. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, 3(2), 73–84.

Bibliografie (3/3)

- Lo, A.W., MacKinlay, A.C. (1990). An econometric analysis of nonsynchronous trading. *Journal of Econometrics*, 45(1–2), 181–211.
- McNeil, A.J., Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7(3–4), 271–300.
- Pascanu, R., Mikolov, T., Bengio, Y. (2013). On the difficulty of training recurrent neural networks. *ICML*; arXiv:1211.5063.
- Patton, A.J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1), 246–256.
- Patton, A.J., Ziegel, J.F., Chen, R. (2019). Dynamic semiparametric models for expected shortfall (and value-at-risk). *Journal of Econometrics*, 211(2), 388–413.
- Romano, J.P., Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica*, 73(4), 1237–1282.
- Rumelhart, D.E., Hinton, G.E., Williams, R.J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536.
- Welch, I., Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4), 1455–1508.