








Modelarea piețelor financiare

Seminarul 19: Recapitulare și atelier de proiect — 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

Acronime folosite în acest material (1/2)

- **AI**: Artificial Intelligence — *inteligență artificială*
- **AIC**: Akaike Information Criterion — *criteriul informațional Akaike*
- **AR**: AutoRegressive (model); in event studies: Abnormal Return — *model autoregresiv; în studiile de eveniment: randament anormal*
- **BET**: Bucharest Exchange Trading (index) — *indicele de referință al Bursei de Valori București*
- **BH**: Benjamini–Hochberg (procedure) — *procedura Benjamini–Hochberg*
- **BIC**: Bayesian Information Criterion — *criteriul informațional bayesian*
- **BVB**: Bursa de Valori București
- **CI**: Confidence Interval — *interval de încredere*
- **COVID-19**: COronaVIrus Disease 2019 — *boala provocată de coronavirus, 2019*
- **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*
- **FDR**: False Discovery Rate — *rata descoperirilor false*
- **FHS**: Filtered Historical Simulation — *simulare istorică filtrată*
- **FTSE**: Financial Times Stock Exchange (FTSE Russell, index provider) — *furnizorul de indici FTSE Russell*
- **FWER**: Family-Wise Error Rate — *probabilitatea de cel puțin o eroare de tip I în familia de teste*
- **FZ0**: Fissler–Ziegel loss, zero-homogeneous version — *funcția de pierdere Fissler–Ziegel, varianta omogenă de grad zero*

Acronime folosite în acest material (2/2)

- **GARCH**: Generalised AutoRegressive Conditional Heteroskedasticity — *heteroscedasticitate condiționată autoregresivă generalizată*
- **GJR**: Glosten–Jagannathan–Runkle (asymmetric GARCH) — *GARCH asimetric Glosten–Jagannathan–Runkle*
- **GJR-GARCH**: Glosten–Jagannathan–Runkle GARCH — *GARCH asimetric Glosten–Jagannathan–Runkle*
- **HAC**: Heteroskedasticity and Autocorrelation Consistent (standard errors) — *erori standard robuste la heteroscedasticitate și autocorelație*
- **HS**: Historical Simulation — *simulare istorică*
- **IGARCH**: Integrated GARCH — *GARCH integrat*
- **IQR**: InterQuartile Range — *intervalul intercuartilic*
- **LR**: Likelihood Ratio (test) — *testul raportului de verosimilitate*
- **MCS**: Model Confidence Set — *mulțimea de încredere a modelelor*
- **OLS**: Ordinary Least Squares — *metoda celor mai mici pătrate*
- **OUG**: Ordonanță de Urgență a Guvernului
- **QML**: Quasi-Maximum Likelihood — *cvasi-verosimilitate maximă*
- **QMLE**: Quasi-Maximum Likelihood Estimator — *estimatorul de cvasi-verosimilitate maximă*
- **SE**: Standard Error — *eroare standard*
- **STOXX50E**: Euro Stoxx 50 index (symbol of the data file) — *indicele Euro Stoxx 50 (simbolul fișierului de date)*
- **SUA**: Statele Unite ale Americii
- **VR**: Variance Ratio — *raportul variațelor*
- **WIG20**: Warszawski Indeks Giełdowy 20 — *indicele celor mai mari 20 de companii de la Bursa din Varșovia*

Este validată prognoza de risc prin backtesting?

- Întrebarea seminarului
  - **ar trece proiectul dumneavoastră verificările recapitulate în Capitolul 19?**
  - de la prețurile brute de închidere ale indicelui la un VaR 1% (valoarea la risc, *Value at Risk*) validat prin backtesting, cu inferență corectă, și de la o idee la o propunere de o pagină
- Livrabile
  - verificările A1–A2 și prognoza independentă A6 cu un model GARCH (*Generalised AutoRegressive Conditional Heteroskedasticity*, heteroscedasticitate condiționată autoregresivă generalizată)
  - un grafic și un tabel pentru B1–B4; B5–B8 sînt extensii opționale
  - codul generat de AI (inteligență artificială) și corectat în C1
  - o propoziție de interpretare pentru fiecare exercițiu
- Notebook-ul de azi: deschideți notebook-ul seminarului în Colab
  - secțiunile notebook-ului poartă aceleași etichete ca exercițiile (A1–A12, B1–B8, C1–C3)
- Seminar: prof. Daniel Traian Pele, danpele@ase.ro; drd. Antoaneta Amza, antoaneta.amza@csie.ase.ro

Ce observații putem compara?

- BET: Bucharest Exchange Trading index, indicele de referință al Bursei de Valori București (BVB)
- WIG20: indicele celor mai mari 20 de companii de la Bursa din Varșovia
- SPY: un Exchange-Traded Fund (ETF), fond tranzacționat la bursă, care urmărește indicele american S&P 500 (*Standard & Poor's*)
- USD: dolarul american
- Fiecare piață își păstrează calendarul de tranzacționare
- Bitcoin se tranzacționează și cînd bursele sînt închise
- Randamentele sînt variații logaritmice ale prețului, în procente
  - De la 100 la 110, randamentul simplu este 10%, iar cel logaritmic circa 9,53%
  - Pierderea este opusul randamentului logaritmic
- Pentru o analiză comună, aliniem întîi prețurile pe datele comune, apoi calculăm randamentele
  - Altfel, două randamente cu aceeași dată pot acoperi intervale diferite

Datele folosite

Serie	Prețul folosit	Calendar	Eșantion	Folosită în
BET	închiderea oficială	zilele de tranzacționare din România	6 ian. 2000 – 18 sept. 2026, 6 670 de randamente	A5, B1–B5
Bitcoin	închiderea în USD	în fiecare zi	18 sept. 2014 – 18 sept. 2026, 4 384 de randamente	B6–B8
SPY (ETF pe S&P 500)	închiderea ajustată	zilele de tranzacționare din SUA	2015 – 2026	B8
WIG20	închiderea	zilele de tranzacționare din Polonia	2014 – 2026	exemplul de propunere
Euro Stoxx 50	închiderea	zilele de tranzacționare locale	2002 – 2018 și ulterior	C3

- Prețurile zilnice din datele cursului, cîte o serie pe activ (cîmpul folosit: close / adjusted close); înregistrările duplicate și rîndurile datate în weekend ale indicilor sînt eliminate
- Două serii cu calendare diferite (B8, C2): aliniem întîi *prețurile* pe zilele comune, apoi calculăm randamentele

Pregătire: rulați și verificați

```
bet = 100 * log_returns('bet')
btc = 100 * log_returns('btc')
print(len(bet), bet.index[0].date(),
      bet.index[-1].date())
print(round(bet.std(), 2), len(btc))
```

- Secțiunea „Date” din notebook definește `log_returns(name)`: randamente log zilnice pe calendarul fiecărei serii
- Rezultatul așteptat
  - `6670 2000-01-06 2026-09-18`
  - `1.4 4384`
- Dacă numărul sau prima dată diferă, opriți-vă și verificați fișierul, cîmpul și regulile de curățare

Instrumentele din capitolele anterioare, exercițiu cu exercițiu

- Aproape toate instrumentele de azi au fost introduse în capitolele anterioare
  - Trei idei noi: puterea unui backtest, un interval la frontieră, capcana deplasării de anticipare (*look-ahead bias*)
- A2: corecția Holm, care controlează probabilitatea unei erori de tip I în familia de teste (*Family-Wise Error Rate*, FWER), și corecția Benjamini–Hochberg (BH), care controlează rata descoperirilor false (*False Discovery Rate*, FDR) (Capitolul 3) Holm (1979), Benjamini & Hochberg (1995)
- B1: indicele cozii (*tail index*) Hill, testul portmanteau robust Q^*, bootstrap pe blocuri suprapuse (Capitolul 1) Hill (1975), Lobato, Nankervis & Savin (2001), Künsch (1989)
- B2: GARCH(1,1) cu inovații Student-t, erori standard robuste, frontiera $\alpha + \beta = 1$ (Capitolul 5) Bollerslev & Wooldridge (1992)
- B3, B4 (Capitolul 8) Kupiec (1995), Christoffersen (1998), Engle & Manganelli (2004)
  - VaR prin simulare istorică (*Historical Simulation*, HS), Normal, GARCH-t și simulare istorică filtrată (*Filtered Historical Simulation*, FHS); testele Kupiec, Christoffersen și testul cuantilei dinamice (*Dynamic Quantile*, DQ)
  - Scorul Fissler–Ziegel (FZ0), testele Diebold–Mariano (DM), mulțimea de încredere a modelelor (*Model Confidence Set*, MCS), backtest-ul comparativ pentru pierderea așteptată în coadă (*Expected Shortfall*, ES)

Momentele, autocorelația și testul Ljung–Box (B1)

- Momentele de selecție ale lui $r_1, \ldots, r_T$, cu $x_t = r_t - \bar{r}$ și s abaterea standard

$$\widehat{\text{skew}} = \frac{1}{T}\sum_{t=1}^{T}\frac{x_t^3}{s^3}, \qquad \hat{\kappa} = \frac{1}{T}\sum_{t=1}^{T}\frac{x_t^4}{s^4} - 3$$

  - asimetria < 0: coada stîngă mai lungă; excesul de kurtosis $\hat{\kappa} > 0$: cozi mai groase decît la distribuția Normală (pentru care este 0)
  - zilele extreme: sub distribuția Normală se așteaptă $T \cdot 2\Phi(-4)$ zile cu $|x_t| > 4s$; Φ: funcția de repartiție Normală standard

- Autocorelația de selecție la lagul k și testul **Ljung–Box** al lipsei autocorelației pînă la lagul m

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^{T} x_t x_{t-k}}{\sum_{t=1}^{T} x_t^2}, \qquad Q(m) = T(T+2)\sum_{k=1}^{m}\frac{\hat{\rho}_k^2}{T-k} \;\sim\; \chi^2_m \text{ sub } H_0$$

  - respingem la 5% dacă $Q(m) > \chi^2_{m;0,95}$ (18,31 pentru $m = 10$); aplicat lui r_t^2, detectează gruparea volatilității (*volatility clustering*)
  - $Q(m)$ presupune randamente i.i.d.; cu gruparea volatilității respinge prea des, de aici varianta robustă $Q^*(m)$ (slide-ul „Testul Q^* robust și bootstrap-ul pe blocuri")

Indicele cozii și estimatorul Hill (B1)

- Coadă de tip putere (Pareto) a pierderilor L, cu **indicele cozii** $\alpha > 0$

$$P(L > u) \approx c\, u^{-\alpha} \quad \text{pentru } u \text{ mare}$$

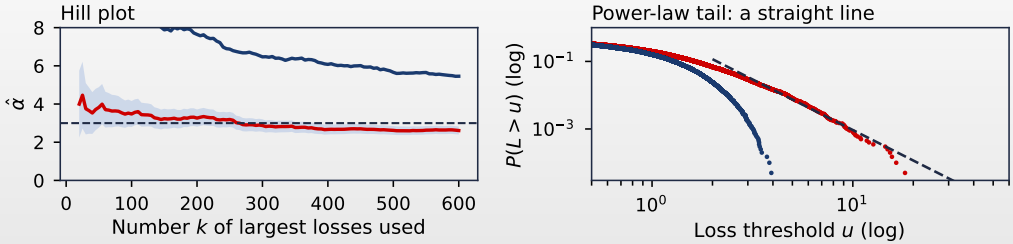
  - $c > 0$: o constantă de scală; α mai mic: coadă mai groasă; $\mathbb{E}|L|^p < \infty$ doar dacă $p < \alpha$ (dispersia: $\alpha > 2$; kurtosis: $\alpha > 4$)

- **Estimatorul Hill** Hill (1975) pe cele mai mari k pierderi $L_{(\mathbf{1})} \geq L_{(\mathbf{2})} \geq \cdots \geq L_{(k+\mathbf{1})}$

$$\hat{\alpha} = \Big(\frac{1}{k}\sum_{i=\mathbf{1}}^{k} \ln L_{(i)} - \ln L_{(k+\mathbf{1})}\Big)^{-\mathbf{1}}, \qquad \mathrm{CI}_{\mathbf{95\%}} = \hat{\alpha}\Big(1 \pm \frac{1{,}96}{\sqrt{k}}\Big)$$

  - media excesului logaritmic peste pragul $L_{(k+\mathbf{1})}$; intervalul de încredere (*Confidence Interval*, CI) presupune pierderi i.i.d. (B1 îl compară cu un bootstrap pe blocuri)
  - k echilibrează deplasarea (k mare intră în centrul distribuției) și varianța (k mic): căutăm o zonă plată a graficului Hill

Un grafic Hill pe pierderi simulate



- $T = 20\,000$ de randamente simulate: Student-$t(3)$, cu $\alpha = 3$ adevărat, și Normale (fără coadă de tip putere)
  - stînga: pentru Student-t, $\hat{\alpha}$ este 3,48 la $k = 100$ și 2,67 la $k = 400$; pentru distribuția Normală se deplasează continuu cu k
  - dreapta: pe scale logaritmice o coadă de tip putere este o dreaptă cu panta $-\alpha$; coada Normală se curbează în jos

Testul Q^* robust și bootstrap-ul pe blocuri (B1)

- Portmanteau robust Lobato, Nankervis & Savin (2001)
  - $Q^*(m) = T \sum_{k=1}^{m} \hat{\rho}_k^2 / \hat{\tau}_k, \;\; \hat{\tau}_k = \overline{x_t^2 x_{t-k}^2} / \hat{\gamma}_0^2, \;\; x_t = r_t - \bar{r}, \;\; \hat{\gamma}_0 = \overline{x_t^2}$
  - $\hat{\rho}_k$: autocorelația de selecție la lagul k; bara: media după t
  - Calibrarea χ_m^2 cere momente de ordinul patru finite
- Bootstrap pe blocuri suprapuse Künsch (1989)
  - Reeșantionăm blocuri de 20 de zile consecutive (999 de reeșantionări): dependența din blocuri se păstrează

GARCH(1,1)-t (B2)

- GARCH(1,1) Bollerslev (1986): varianța de azi este o sumă ponderată dintre o constantă, pătratul șocului de ieri și varianța de ieri

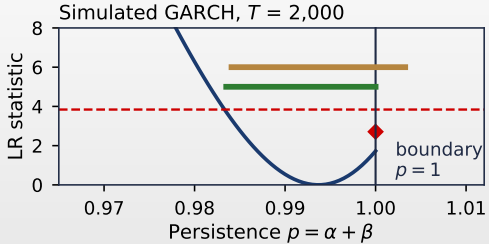
$$r_t = \mu + \varepsilon_t, \qquad \varepsilon_t = \sigma_t z_t, \qquad \sigma_t^2 = \omega + \alpha\,\varepsilon_{t-1}^2 + \beta\,\sigma_{t-1}^2$$

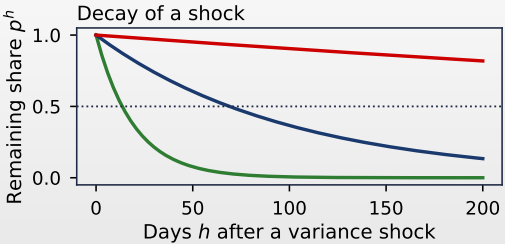
- Notațiile
  - μ: randamentul mediu; σ_t: volatilitatea condiționată, cunoscută în $t-1$
  - z_t: inovație i.i.d. cu media 0 și varianța 1: o Student-t cu ν grade de libertate, rescalată cu $\sqrt{(\nu-2)/\nu}$
  - $\omega > 0$; $\alpha \geq 0$: reacția la știri; $\beta \geq 0$: memoria; ν mic: cozi groase
- Persistența $p = \alpha + \beta$
  - Un șoc de varianță scade cu factorul p în fiecare zi; $p < 1$: staționaritate în covarianță

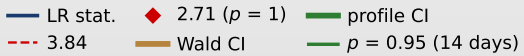
Un interval la frontieră (B2)

- Un interval Wald $\hat{p} \pm 1{,}96\,\mathrm{se}(\hat{p})$ poate ieși din spațiul parametrilor
  - Persistența GARCH $p = \alpha + \beta \leq 1$, dar $0{,}99 \pm 0{,}02$ dă [0,97; 1,01]
  - Lîngă frontieră $\hat{p}$ nu are distribuție Normală Andrews (1999), iar intervalul conține valori în afara regiunii impuse $p \leq 1$
- **Verosimilitatea profil**: $\ell_P(p)$ = maximul log-verosimilității în raport cu ceilalți parametri, cu p fixat
  - Intervalul de 95%: toți p cu $2(\ell_{\max} - \ell_P(p)) \leq 3{,}84$; $\ell_{\max}$: log-verosimilitatea maximizată
  - Rămîne în interiorul $p \leq 1$
  - La $p = 1$ statistica are distribuția $\frac{1}{2}\chi^2_0 + \frac{1}{2}\chi^2_1$ (χ^2_0: o masă punctuală în 0), valoarea critică 2,71 (Capitolul 5)
- Timpul de înjumătățire al unui șoc: $h = \ln 0{,}5 / \ln p$
  - 14 zile la $p = 0{,}95$, 69 la $p = 0{,}99$, 693 la $p = 0{,}999$, infinit la $p = 1$

Intervalele Wald și profil; timpul de înjumătățire al unui șoc







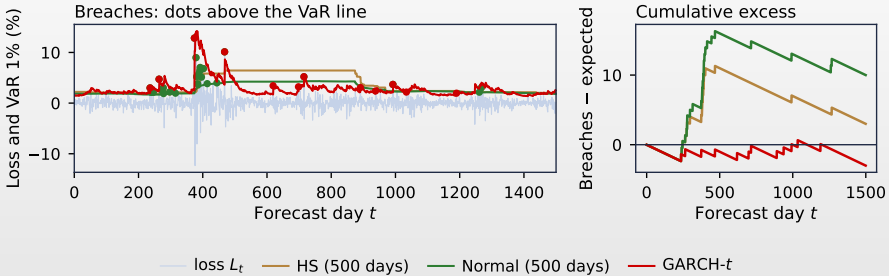


- Stînga: GARCH(1,1) simulat, $p = 0{,}995$ adevărat, $T = 2\,000$; estimarea $\hat{p} = 0{,}994$
  - Wald [0,984; 1,003] trece de frontieră; profil [0,984; 1]: $2(\ell_{\max} - \ell_P(1)) = 1{,}73 < 2{,}71$, deci $p = 1$ nu este respins
- Dreapta: partea p^h rămasă dintr-un șoc de varianță după h zile; timpul de înjumătățire $\ln 0{,}5 / \ln p$
  - lîngă $p = 1$ o mică schimbare a lui p mută timpul de înjumătățire cu luni întregi: raportați-l ca interval

Patru prognoze VaR

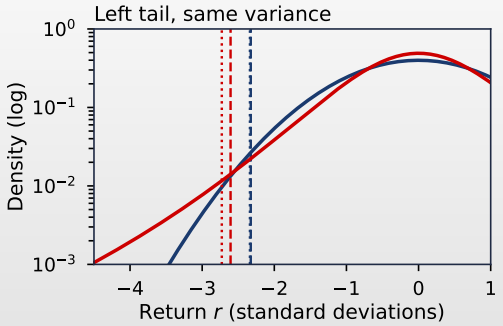
- VaR 1% pe o zi pentru ziua t, din datele pînă la $t-1$
  - **HS**: cuantila empirică de 1% a pierderilor din ultimele 500 de zile
  - **Normal**: $\mathrm{VaR} = -(\hat{\mu} + \hat{\sigma} z_{0,01})$, $\hat{\mu}$, $\hat{\sigma}$ din ultimele 500 de zile, $z_{0,01} = -2{,}326$
- **GARCH-t**: $\mathrm{VaR}_t = -(\hat{\mu} + \hat{\sigma}_t \, \tilde{t}_{\hat{\nu}}^{-1}(0{,}01))$
  - $\tilde{t}_{\nu}^{-1}(0{,}01) = t_{\nu}^{-1}(0{,}01)\sqrt{(\nu-2)/\nu}$: cuantila de 1% a distribuției Student-t standardizate
- **FHS** (simulare istorică filtrată) McNeil & Frey (2000)
  - Aceeași formulă cu $\hat{q}_{0,01}(z)$ în locul cuantilei Student-t
  - $\hat{q}_{0,01}(z)$: cuantila empirică de 1% a reziduurilor standardizate $\hat{z}_t = (r_t - \hat{\mu})/\hat{\sigma}_t$

Prognoze VaR pe fereastră mobilă și depășirile lor



- O traiectorie GARCH-t(5) simulată, 1 500 de zile de prognoză, VaR 1% din datele pînă la $t-1$; 15 depășiri așteptate
  - Întrebare pentru sală: care dintre cele trei modele din grafic (HS, Normal, GARCH-t) va avea cele mai multe depășiri?
  - Răspuns: Normal, 25 față de 15 așteptate; GARCH-t, modelul adevărat, are 12
  - HS are 18 depășiri; HS și Normal reacționează tîrziu, iar depășirile lor se grupează după creșterea volatilității
  - dreapta: depășirile cumulate minus cele așteptate, 0,01 t; un model corect oscilează în jurul lui zero

VaR și ES sub distribuția Normală și Student-t



- Randamente $N(\mu, \sigma^2)$, probabilitatea cozii α

$$\mathrm{VaR}_\alpha = -(\mu + \sigma z_\alpha), \qquad \mathrm{ES}_\alpha = -\mu + \sigma \frac{\varphi(z_\alpha)}{\alpha}$$

  - $z_\alpha = \Phi^{-1}(\alpha)$; φ: densitatea Normală standard
- Inovații Student-t cu ν grade de libertate, scalate la varianța 1
  - cuantila $\tilde{t}_\nu^{-1}(\alpha)$ din slide-ul „Patru prognoze VaR" în locul lui z_α; ES are propria formulă închisă (A8)
- Graficul (coada stîngă, scală logaritmică, varianța 1)
  - Distribuția Normală: VaR 1% = 2,33, ES 2,5% = 2,34, aproape egale
  - Student-t(5): calculați cifrele în A8

Testele de independență și DQ

- Testul de independență Christoffersen Christoffersen (1998): este o depășire mai probabilă după o depășire?

$$LR_{ind} = -2 \ln \frac{(1-\hat{\pi})^{n_{00}+n_{10}}\, \hat{\pi}^{\,n_{01}+n_{11}}}{(1-\hat{\pi}_{01})^{n_{00}}\, \hat{\pi}_{01}^{n_{01}}\, (1-\hat{\pi}_{11})^{n_{10}}\, \hat{\pi}_{11}^{n_{11}}} \;\sim\; \chi^2_1$$

  - n_{ij}: numărul zilelor în starea j după o zi în starea i ($1 =$ depășire); $\hat{\pi}_{01} = n_{01}/(n_{00}+n_{01})$, $\hat{\pi}_{11} = n_{11}/(n_{10}+n_{11})$
  - $\hat{\pi} = (n_{01}+n_{11})/(n_{00}+n_{01}+n_{10}+n_{11})$: rata comună a depășirilor (numărătorul); numitorul lasă rata să depindă de ziua precedentă; LR_{ind} mare: depășiri grupate

- Testul cuantilei dinamice (DQ) Engle & Manganelli (2004): poate fi prezisă depășirea de azi?

$$I_t - \alpha = \delta_0 + \sum_{k=1}^{4} \delta_k \left(I_{t-k} - \alpha\right) + \delta_5 \,\mathrm{VaR}_t + u_t, \qquad DQ = \frac{\hat{\delta}^\top X^\top X\, \hat{\delta}}{\alpha(1-\alpha)} \;\sim\; \chi^2_6$$

  - I_t: indicatorul depășirii; X: matricea celor 6 regresori; $\hat{\delta}$: coeficienții estimați prin metoda celor mai mici pătrate (*Ordinary Least Squares*, OLS); H_0: toți $\delta = 0$
  - DQ mare: depășirile sînt previzibile din depășirile trecute sau din nivelul VaR

Scoruri și compararea modelelor

- Scorul FZ0 Patton, Ziegel & Chen (2019): evaluează o pereche (VaR, ES) în ziua t; mai mic este mai bine

$$\mathrm{FZ0}(v, e; y) = -\frac{1\{y \le v\}(v - y)}{\alpha e} + \frac{v}{e} + \ln(-e) - 1$$

  - $\alpha = 2{,}5\%$, $v = -\mathrm{VaR}$, $e = -\mathrm{ES}$, y: randamentul realizat; $1\{\cdot\}$: 1 dacă este adevărat, altfel 0

- Două modele: statistica Diebold–Mariano $t = \bar{d}/\sqrt{\hat{\Omega}/T}$ Diebold & Mariano (1995)
  - d_t: diferența FZ0 în ziua t, $\bar{d}$: media ei; $\hat{\Omega}$: dispersia ei de termen lung Newey–West, cu $L = \lfloor 4(T/100)^{2/9} \rfloor$ laguri; T: numărul zilelor de prognoză comune
  - FZ0 se calculează pe randamente procentuale: comparați modele pe același activ și în aceleași unități
  - $t \approx N(0, 1)$ la acuratețe egală
- Mai multe modele: mulțimea de încredere a modelelor la 10% Hansen, Lunde & Nason (2011)
  - Păstrează fiecare model cu $p_{MCS} \ge 0{,}10$; conține cel mai bun model cu probabilitatea 90%

Ce probabilitate numim p aici?

- În discuția GARCH, p este persistența
  - suma celor doi coeficienți dinamici, $\alpha + \beta$
  - o valoare apropiată de 1: un șoc se stinge lent
- În testul Kupiec și în A1, p este probabilitatea nominală de depășire
  - fixată la 1%
  - corespunde probabilității de coadă α din formulele VaR, ES, DQ și FZ0 și din Capitolul 19
- Proporția observată a depășirilor estimează probabilitatea reală de depășire, π
  - H_0: probabilitatea reală este egală cu cea nominală
- Nu confundați cele două sensuri: persistența GARCH nu intră în statistica Kupiec
  - la fel, α notează și coeficientul ARCH (*AutoRegressive Conditional Heteroskedasticity*, heteroscedasticitate condiționată autoregresivă) al GARCH și indicele cozii Hill: sensul reiese din formulă

Nivelul, puterea și testul Kupiec (A1)

- **Nivelul** (*size*) și **puterea**
  - Nivelul: probabilitatea de a respinge un model corect; puterea: probabilitatea de a respinge un model greșit
  - Un backtest este util doar dacă are putere mare pentru erorile care contează
- Cu depășiri independente, numărul lor X în T zile este binomial, cu rata reală π

$$X \sim \text{Bin}(T, \pi), \qquad P(X = x) = \binom{T}{x} \pi^x (1-\pi)^{T-x}$$

- Testul de acoperire Kupiec Kupiec (1995): rata nominală $p = 1\%$ față de rata observată $\hat{\pi} = x/T$, cu statistica raportului de verosimilitate (*Likelihood Ratio*, LR)

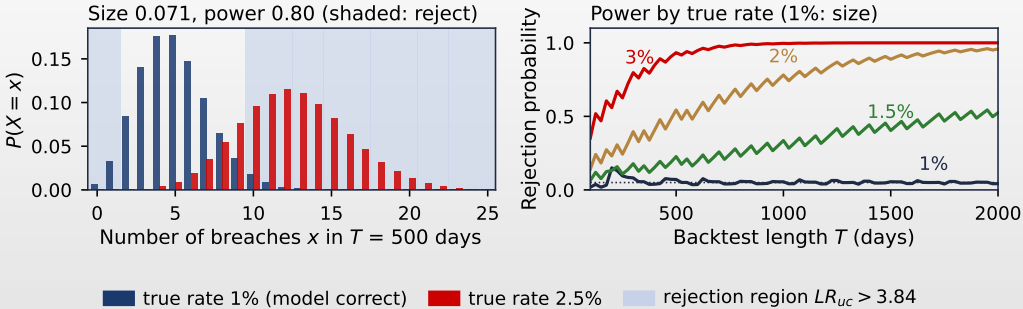
$$LR_{uc}(x) = -2 \ln \frac{(1-p)^{T-x} p^x}{(1-\hat{\pi})^{T-x} \hat{\pi}^x} \sim \chi^2_1 \text{ sub } H_0 : \pi = p$$

  - x: depășirile observate; $LR_{uc} = 0$ cînd $\hat{\pi} = p$, mai mare cînd $\hat{\pi}$ se îndepărtează de p; convenția $0 \ln 0 = 0$
  - Respingem la 5% cînd $LR_{uc}(x) > 3{,}84$, cuantila de 95% a lui χ^2_1

Nivelul și puterea exacte (A1)

- Nivelul și puterea exacte din distribuția binomială
  - Regiunea de respingere R = numerele x cu $LR_{uc}(x) > 3{,}84$
  - Nivelul exact $= \sum_{x \in R} P(X = x \mid \pi = 1\%)$; puterea exactă $= \sum_{x \in R} P(X = x \mid \pi)$ la π adevărat
  - X este discret, deci nivelul exact nu este chiar 5%
- Exemplu: $T = 100$ de zile (1 depășire așteptată)
  - $LR_{uc}(0) = -2 \cdot 100 \ln 0{,}99 = 2{,}01$, $LR_{uc}(3) = 2{,}63$, $LR_{uc}(4) = 5{,}18$: $R = \{x \geq 4\}$
  - Zero depășiri nu duc niciodată la respingere
  - Nivelul $P(X \geq 4 \mid 1\%) = 0{,}018$; puterea 0,14 dacă rata reală este 2%, 0,74 dacă este 5%
- O nerespingere cu putere mică înseamnă „nicio dovadă", nu „nicio problemă"
  - Precizați puterea înaintea testului (A1; detalii: Capitolul 19)

Nivelul și puterea testului Kupiec



- Stînga: depășirile în $T = 500$ de zile sub rata nominală 1% și sub o rată reală de 2,5%; numerele hașurate duc la respingere ($x \leq 1$ sau $x \geq 10$)
  - nivelul 0,071 (nu exact 5%: X este discret); puterea 0,80 față de 2,5%
  - dreapta: puterea crește cu T și cu distanța ratei reale față de 1%; la 1,5% chiar și 2 000 de zile dau doar aproximativ 50%

Multe teste simultan: Bonferroni, Holm și BH (A2)

- m teste, valorile p (*p-values*) ordonate $p_{(1)} \le \cdots \le p_{(m)}$; V: respingerile false, R: toate respingerile

$$\mathrm{FWER} = P(V \ge 1), \qquad \mathrm{FDR} = \mathbb{E}\Big[\frac{V}{\max(R,1)}\Big]$$

  - FWER: probabilitatea a cel puțin unei descoperiri false; FDR: proporția așteptată a descoperirilor false printre toate descoperirile

- **Bonferroni**: respingem $H_{(k)}$ dacă $p_{(k)} \le 0{,}05/m$; controlează FWER, conservator
- **Corecția descendentă** (*step-down*) Holm (1979): FWER $\le 5\%$ sub orice dependență, niciodată mai puține respingeri decît Bonferroni

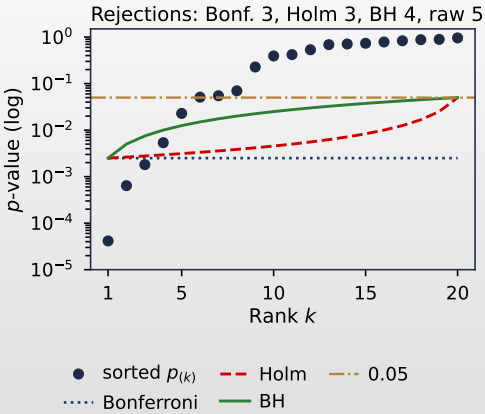
$$\text{respingem } H_{(k)} \text{ cît timp } p_{(k)} \le \frac{0{,}05}{m-k+1},\ k=1,2,\ldots; \text{ ne oprim la primul eșec}$$

- **Procedura ascendentă** (*step-up*) Benjamini & Hochberg (1995): FDR $\le 5\%$ pentru valori p independente sau pozitiv dependente

$$k^* = \max\Big\{k : p_{(k)} \le \frac{0{,}05\,k}{m}\Big\}; \quad \text{respingem } H_{(1)},\ldots,H_{(k^*)}$$

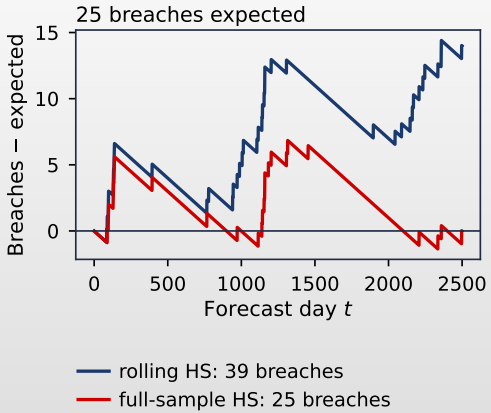
  - BH (Benjamini–Hochberg) respinge cel puțin la fel de multe ipoteze ca Holm

Cele trei praguri pe valori p simulate



- 20 de teste simulate, 6 cu un efect real
  - punctele: $p_{(k)}$ ordonate; un punct sub o linie este respins de acea regulă (Holm și BH ca pe slide-ul anterior)
- Respingerile
  - necorectat la 5%: 5; Bonferroni: 3; Holm: 3; BH: 4
  - linia Holm urcă pînă la 0,05 la ultimul rang; linia BH urcă liniar, deci admite mai multe descoperiri
- Ce alegem
  - selecția dintre mulți candidați: BH (cîteva piste false sînt acceptabile)
  - alocarea de capital unui singur rezultat: controlul FWER (Holm)

Deplasarea de anticipare (B4)



- **Deplasarea de anticipare**: o prognoză pentru ziua t care folosește informație de după ziua $t-1$
  - de ex. parametri sau o cuantilă estimate pe tot eșantionul
  - prognoza pare mai bună decît orice prognoză care ar fi putut fi făcută în timp real
- Graficul: pierderi GARCH-t simulate, 2 500 de zile de prognoză, 25 de depășiri așteptate
  - HS pe fereastră mobilă (500 de zile): 39 de depășiri; HS pe tot eșantionul: 25, exact 1% prin construcție
- Testul pentru codul dumneavoastră
  - orice prognoză pentru ziua t trebuie să poată fi calculată dintr-un fișier de date trunchiat la $t-1$ (B4)

Erori robuste și prognoze GARCH

- Estimarea prin cvasi-verosimilitate maximă (*Quasi-Maximum Likelihood Estimation*, QMLE; prescurtat și QML) Bollerslev & Wooldridge (1992) (A3): verosimilitate Gaussiană, orice distribuție a inovațiilor cu $\mathbb{E}z_t^4 < \infty$

$$\widehat{\mathrm{Var}}(\hat\theta) = \frac{1}{T} A^{-1} B A^{-1}, \qquad A = -\mathbb{E}\Big[\frac{\partial^2 \ell_t}{\partial\theta\,\partial\theta^\top}\Big], \quad B = \mathbb{E}\big[s_t s_t^\top\big]$$

  - ℓ_t: log-verosimilitatea zilei t; $s_t = \partial\ell_t/\partial\theta$: scorul; cu inovații Normale $A = B$, iar sandvișul se reduce la A^{-1}/T
- Metoda delta (A4): eroarea standard a unei funcții netede $g(\hat p)$
  - $\mathrm{se}(g(\hat p)) \approx |g'(\hat p)|\,\mathrm{se}(\hat p)$; nefiabilă când g' explodează, ca pentru timpul de înjumătățire lîngă $p = 1$
- Prognozele de varianță GARCH(1,1) (A5, A6): revenire geometrică la nivelul de termen lung

$$\sigma^2_{T+h} = \bar\sigma^2 + (\alpha+\beta)^{h-1}\big(\sigma^2_{T+1} - \bar\sigma^2\big), \qquad \bar\sigma^2 = \frac{\omega}{1-\alpha-\beta}$$

  - varianța pe h zile este suma $\sum_{j=1}^{h} \sigma^2_{T+j}$, nu $h\,\sigma^2_{T+1}$ (regula rădăcinii timpului)

Harta exercițiilor

Exercițiu	Întrebare	Tip	Model
A1	cît de des detectează un an de date un model VaR slab calibrat?	Rezolvat	–
A2	ce strategii rămîn semnificative după corecția pentru testare multiplă?	Rezolvat	–
A6	prognoză GARCH cu parametri dați	Propus, independent	noțiunile GARCH
B1	cît de groase sînt cozile BET?	Rezolvat	–
B2	ce interval pentru persistența GARCH?	Rezolvat	–
B3	ce model VaR 1% ar fi fost fiabil?	Rezolvat	–
B4	cît de mult avantajează artificial deplasarea de anticipare un backtest?	Rezolvat	–
C1	ce este greșit într-un backtest scris de AI?	Propus	A1, B3, B4
C2	propunerea echipei (temă)	Propus	exemplul de propunere
Pachete de extindere	A3–A5, A7–A12, B5–B8, C3 (opționale, după curs)	Rezolvate și propuse (indicat la fiecare cerință)	indicat la fiecare cerință

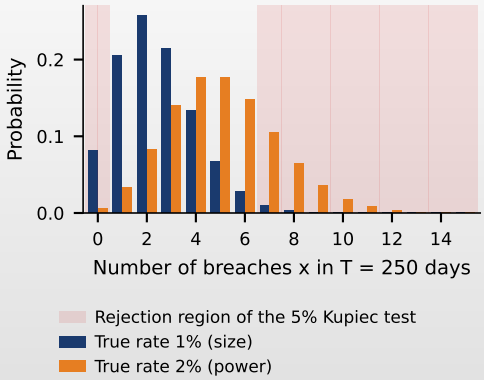
- **[Rezolvat]**: rezolvarea completă în slide-uri și în notebook, un model de urmat; **[Propus]**: îl rezolvați dumneavoastră, rezolvarea se discută la seminar

A1 [Rezolvat]: Puterea testului Kupiec — Cerință

- **Întrebare**: Cît de des detectează un an de date un model VaR 1% a cărui rată reală de depășire este 2%?
- **Date și ipoteze**
  - Un model VaR 1% este verificat pe T zile; numărul de depășiri este $X \sim \mathrm{Bin}(T, \pi)$; $H_0 : \pi = 0{,}01$.
  - Statistica Kupiec $LR_{uc}(x)$ cu rata nominală $p = 0{,}01$ (formula: slide-ul „Nivelul, puterea și testul Kupiec"); respingem la 5% dacă $LR_{uc} > 3{,}84$ Kupiec (1995).
- **Cerințe**
  - (a) Pentru $T = 250$, calculați $LR_{uc}(x)$ pentru $x = 0, 1, 2, \ldots$ și găsiți regiunea de respingere R; repetați pentru $T = 1\,000$.
  - (b) Calculați nivelul exact $P(X \in R \mid \pi = 1\%)$ pentru ambele valori T.
  - (c) Calculați puterea exactă $P(X \in R \mid \pi = 2\%)$ pentru ambele valori T.
- **Raportați**
  - Cele două regiuni de respingere, nivelurile exacte și puterile exacte.
- **Interpretare**
  - Cît de des detectează un an de date un model a cărui rată reală a depășirilor este dublă față de cea nominală?
  - De ce adaugă Basel multiplicatorul semaforului (zonele verde, galbenă și roșie; Capitolul 8)?

◂ Înapoi ▸ Lista exercițiilor

A1 [Rezolvat]: Rezolvare (1/2)



- (a) Pasul 1 ($T = 250$, 2,5 depășiri așteptate): $LR_{uc}(0) = -2 \cdot 250 \ln 0{,}99 = 5{,}03 > 3{,}84$, deci $x = 0$ este respins; $LR_{uc}(1) = 1{,}18$.
- Pasul 2: $LR_{uc}(6) = 3{,}56 < 3{,}84$ și $LR_{uc}(7) = 5{,}50 > 3{,}84$, deci $R = \{0\} \cup \{x \geq 7\}$.
- Pasul 3 ($T = 1\,000$, 10 depășiri așteptate): același calcul dă $R = \{x \leq 4\} \cup \{x \geq 17\}$.
- (b) Pasul 4: la $T = 250$, nivelul $= P(X = 0) + P(X \geq 7) = 0{,}0811 + 0{,}0137 = 0{,}095$ pentru $\pi = 1\%$; la $T = 1\,000$ nivelul este 0,055.
- Observație: X este discret, deci nivelul exact nu este 5%; cu puține depășiri așteptate poate fi aproape dublul nivelului nominal.

MFM_ch19_seminar

A1 [Rezolvat]: Rezolvare (2/2)

- (c) Pasul 5: la $T = 250$, puterea $= P(X = 0) + P(X \geq 7) = 0{,}0064 + 0{,}2363 = 0{,}243$ pentru $\pi = 2\%$.
- Pasul 6: la $T = 1\,000$, $P(X \leq 4) + P(X \geq 17) = 0{,}782$ pentru $\pi = 2\%$.

	$T = 250$	$T = 1\,000$
Regiunea de respingere	$x = 0$ sau $x \geq 7$	$x \leq 4$ sau $x \geq 17$
Nivel exact ($\pi = 1\%$)	0,095	0,055
Putere ($\pi = 2\%$)	0,243	0,782

- Interpretare: un an de date respinge un model cu rata dublă a depășirilor doar aproximativ o dată din patru; sînt necesari patru ani pentru o putere de aproape 80%; de aceea Basel adaugă multiplicatorul semaforului în loc să se bazeze doar pe test.

A2 [Rezolvat]: Holm și Benjamini–Hochberg — Cerință

- **Întrebare**: Care dintre zece strategii de tranzacționare rămîn semnificative după o corecție pentru testarea multiplă?
- **Date și ipoteze**
  - Zece strategii de tranzacționare, fiecare testată față de un randament în exces nul; valorile p ordonate $p_{(1)}, \ldots, p_{(10)}$: 0,001, 0,004, 0,012, 0,019, 0,028, 0,041, 0,09, 0,20, 0,46, 0,73; $m = 10$; nivelul țintă 5%.
  - FWER, FDR și regulile Bonferroni, Holm și BH: slide-ul „Multe teste simultan".
- **Cerințe**
  - (a) Numărați respingerile fără corecție ($p \leq 0{,}05$) și cu Bonferroni ($p \leq 0{,}05/m$).
  - (b) Aplicați Holm la FWER 5% Holm (1979): comparați $p_{(i)}$ cu $0{,}05/(m - i + 1)$ pentru $i = 1, 2, \ldots$ și opriți-vă la primul $p_{(i)}$ peste pragul său.
  - (c) Aplicați Benjamini–Hochberg la FDR 5% Benjamini & Hochberg (1995): găsiți cel mai mare k cu $p_{(k)} \leq 0{,}05k/m$ și respingeți $H_{(1)}, \ldots, H_{(k)}$.
- **Raportați**
  - Un tabel cu $p_{(i)}$ și ambele praguri, precum și numărul de respingeri al fiecărei proceduri.
- **Interpretare**
  - Ce garantează fiecare procedură?
  - Pe care ați folosi-o pentru a selecta strategii de tranzacționare?

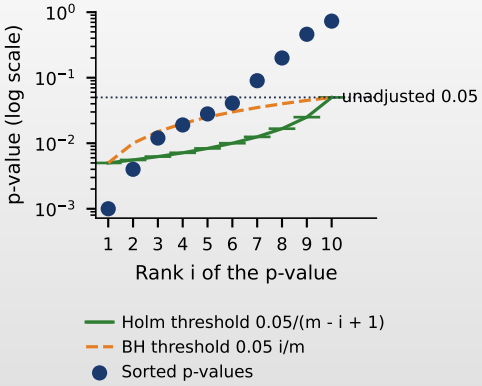
◂ Înapoi ▸ Lista exercițiilor

A2 [Rezolvat]: Rezolvare (1/2)

i	1	2	3	4	5	6	7	8	9	10
$p_{(i)}$	0,001	0,004	0,012	0,019	0,028	0,041	0,09	0,20	0,46	0,73
Holm	0,0050	0,0056	0,0063	0,0071	0,0083	0,0100	0,0125	0,0167	0,0250	0,0500
BH	0,005	0,010	0,015	0,020	0,025	0,030	0,035	0,040	0,045	0,050

- (a) Fără corecție: 6 valori p sub 5%; Bonferroni ($p \leq 0{,}005$): 2 respingeri.
- (b) Holm, verificare succesivă în ordinea crescătoare a valorilor p: $0{,}001 \leq 0{,}0050$ și $0{,}004 \leq 0{,}0056$ trec; $0{,}012 > 0{,}0063$ oprește procedura: 2 respingeri.

A2 [Rezolvat]: Rezolvare (2/2)



- (c) BH, verificare pornind de la cea mai mare valoare p: $0{,}041 > 0{,}030$, $0{,}028 > 0{,}025$, dar $0{,}019 \leq 0{,}020$, deci $k = 4$: 4 respingeri (nu mai trebuie verificat $0{,}012 \leq 0{,}015$, deoarece BH respinge toate $H_{(i)}$ cu $i \leq k$).
- Rezumat: fără corecție 6, Bonferroni 2, Holm 2, BH 4.
- Holm garantează: probabilitatea oricărei descoperiri false $\leq 5\%$ (FWER), uniform mai puternic decît Bonferroni.
- BH garantează: ponderea așteptată a descoperirilor false printre cele declarate $\leq 5\%$ (FDR) cînd valorile p sînt independente sau pozitiv dependente prin regresie de ipotezele nule adevărate Benjamini & Yekutieli (2001); respinge mai mult.
- Interpretare: pentru selectarea multor strategii candidate, BH este alegerea obișnuită (cîteva piste false sînt acceptabile); înainte de a aloca capital unei strategii, controlul FWER (Holm) este garanția mai sigură.

MFM_ch19_seminar

A6 [Propus]: Prognoza GARCH cu parametri dați — Cerință

Rezolvarea se discută la seminar

- **Întrebare**: Ce devine prognoza cînd volatilitatea de mîine este peste nivelul de termen lung?
- **Date și ipoteze**
  - Un GARCH(1,1) pentru randamente zilnice în %, cu $\omega = 0{,}02$, $\alpha = 0{,}10$, $\beta = 0{,}88$; varianța de mîine $\sigma^2_{T+1} = 4$ ($\sigma_{T+1} = 2\%$).
  - Formula de prognoză și suma geometrică: slide-ul „Erori robuste și prognoze GARCH"; 252 de zile de tranzacționare pe an.
- **Cerințe**
  - (a) Calculați persistența, varianța de termen lung $\bar{\sigma}^2$ și volatilitatea anualizată de termen lung.
  - (b) Calculați timpul de înjumătățire al unui șoc.
  - (c) Prognozați varianța în ziua 5, σ^2_{T+5}.
  - (d) Calculați varianța randamentului pe 5 zile și volatilitatea pe 5 zile și comparați-le cu $5\sigma^2_{T+1} = 20$ și $\sqrt{5}\,\sigma_{T+1}$.
- **Raportați**
  - Fiecare cifră de la (a)–(d), cu formula în care s-au înlocuit valorile.
- **Interpretare**
  - Este volatilitatea pe 5 zile peste sau sub $\sqrt{5}\,\sigma_{T+1}$?
  - De ce?

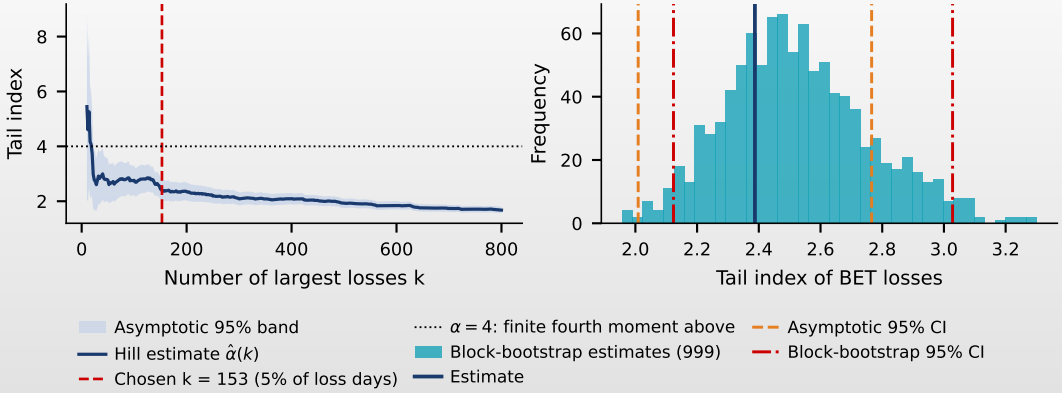
înapoi Lista exercițiilor

B1 [Rezolvat]: Faptele stilizate ale BET cu inferență robustă — Cerință

- **Întrebări**
  - Cît de groase sînt cozile BET?
  - Rămîne semnificativă autocorelația lui după ce ținem cont de gruparea volatilității?
- **Date și ipoteze**
  - Închiderile zilnice oficiale BET (datele cursului), 6 ian. 2000 – 18 sept. 2026: $T = 6\,670$ de randamente log r_t în %; pierderile $L_t = -r_t$; bootstrap pe blocuri suprapuse, 999 de reeșantionări cu blocuri de 20 de zile; toate testele la 5%.
- **Cerințe**
  - (a) Calculați abaterea standard, asimetria și excesul de kurtosis ale lui r_t.
  - (b) Estimați indicele cozii Hill al pierderilor cu $k = 5\%$ din zilele cu pierdere, cu CI asimptotic de 95% din slide-ul „Indicele cozii și estimatorul Hill" și CI percentil de 95% din bootstrap-ul pe blocuri.
  - (c) Calculați Ljung–Box $Q(10)$ pe r_t și pe r_t^2 și statistica robustă $Q^*(10)$ din slide-ul „Testul Q^* robust și bootstrap-ul pe blocuri"; comparați fiecare cu $\chi^2_{\mathbf{10;0,95}} = 18{,}31$.
  - (d) Numărați zilele cu $|r_t - \bar{r}| > 4s$ și comparați cu numărul așteptat pentru distribuția Normală (slide-ul „Momentele, autocorelația și testul Ljung–Box").
- **Raportați**
  - Statisticile de la (a), $\hat{\alpha}$ cu ambele intervale și un grafic Hill, cele trei statistici portmanteau cu valorile p și deciziile, numărul de zile de la (d).
- **Interpretare**
  - De ce este intervalul pe blocuri pentru indicele cozii mai larg decît cel asimptotic?
  - Rămîne autocorelația lui r_t semnificativă în testul robust?
  - Cît de fiabilă este valoarea p, dat fiind indicele cozii estimat?

◂ Înapoi ▸ Lista exercițiilor

B1 [Rezolvat]: Rezolvare (1/3)



MFM_ch19_seminar

B1 [Rezolvat]: Rezolvare (2/3)

- (a) $s = 1{,}40\%$, asimetria $-0{,}65$, excesul de kurtosis 11,1: cozi groase, asimetrice la stînga.
- (b) Pasul 1: ordonăm pierderile; $k = 153$ (5% din zilele cu pierdere); $\hat{\alpha} = 2{,}39$.
- Pasul 2 (asimptotic): $2{,}39(1 \pm 1{,}96/\sqrt{153}) = [2{,}01;\ 2{,}77]$, eroarea standard 0,193.
- Pasul 3 (bootstrap pe blocuri): reconstruim 999 de serii din blocuri aleatoare de 20 de zile, reestimăm $\hat{\alpha}$ cu același k, luăm percentilele 2,5% și 97,5%: $[2{,}12;\ 3{,}03]$, eroarea standard 0,233.
- Observație: ambele intervale sînt sub 4, ceea ce susține un moment de ordinul patru infinit în modelul presupus al cozii cu variație regulată, iar kurtosis-ul de selecție (11,1) nu este raportat ca estimare (A11).

MFM_ch19_seminar

B1 [Rezolvat]: Rezolvare (3/3)

- (c) $Q(10)$ pe r_t: 108,0 (p <0,001) > 18,31: respingem lipsa autocorelației; $Q^*(10) = 24{,}4$ robust (p 0,007): respingem în continuare, dar este de 4,4 ori mai mic.
- $Q(10)$ pe r_t^2: 2 440 (p <0,001): gruparea volatilității este puternică (rezultat citit descriptiv, deoarece, dacă momentul de ordinul patru este infinit, r_t^2 nu are dispersie finită); autocorelația de ordinul 1 a lui r_t: 0,109.
- (d) 51 de zile dincolo de ± 4 abateri standard, față de $6\,670 \cdot 2\Phi(-4) = 0{,}42$ așteptate pentru distribuția Normală.
- Interpretare 1: valorile extreme se grupează în crize; blocurile păstrează grupările, deci intervalul pe blocuri ține cont de dependența pe care formula asimptotică o ignoră.
- Interpretare 2: testul robust încă respinge, dar cu $\hat{\alpha} < 4$ calibrarea lui χ^2 nu este garantată: autocorelațiile de selecție ale randamentelor de tip GARCH cu cozi groase converg mai lent și către limite non-Normale Davis & Mikosch (1998); interpretați rezultatul ca indiciu al unei autocorelații mici; prețurile neactualizate (*stale prices*) ale acțiunilor mai puțin lichide (Capitolele 2, 10) sînt o explicație plauzibilă, care cere un test propriu.

MFM_ch19_seminar

B2 [Rezolvat]: GARCH(1,1)-t pentru BET: un interval la frontieră — Cerință

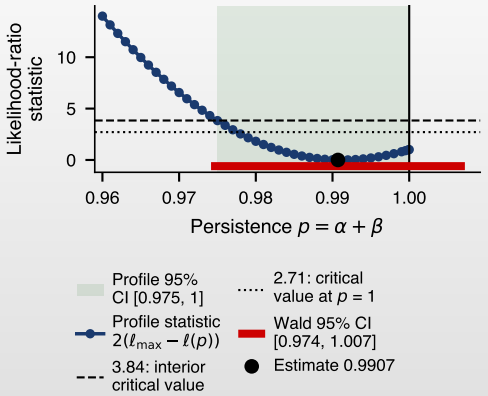
- **Întrebare**: Ce interval pentru persistența GARCH poate fi raportat cînd estimarea este aproape de 1?
- **Date și ipoteze**
  - Randamentele log zilnice BET în % (B1); medie constantă; modelul definit în slide-ul „GARCH(1,1)-t"; erori standard robuste Bollerslev & Wooldridge (1992); nivelul 95%.
  - Persistența $p = \alpha + \beta$: o dispersie necondiționată finită cere $p < 1$; pentru staționaritatea strictă ajunge $\mathbb{E}\ln(\alpha z_t^2 + \beta) < 0$ Nelson (1990), cu $z_t = \varepsilon_t/\sigma_t$.
  - $\hat{p}$ se află lîngă frontiera $p = 1$ a spațiului impus al parametrilor $p \leq 1$ Andrews (1999).
- **Cerințe**
  - (a) Estimați modelul prin verosimilitate maximă; raportați $\hat{\alpha}$, $\hat{\beta}$, $\hat{\nu}$ cu erori standard robuste și $\hat{p}$.
  - (b) Calculați $\mathrm{se}(\hat{p})$ din covarianța robustă a lui $(\hat{\alpha}, \hat{\beta})$ și CI Wald $\hat{p} \pm 1{,}96\,\mathrm{se}(\hat{p})$.
  - (c) Verosimilitatea profil: pentru fiecare p din $0{,}960, 0{,}961, \ldots, 0{,}999, 0{,}9995, 0{,}9999, 1$ maximizați log-verosimilitatea după μ, ω, ν și $s = \alpha/p$ ($\alpha = sp$, $\beta = (1-s)p$); păstrați $p < 1$ dacă $2(\ell_{\max} - \ell(p)) \leq 3{,}84$ și $p = 1$ dacă statistica este $\leq 2{,}71$ (slide-ul „Un interval la frontieră").
  - (d) Transformați estimarea și ambele intervale în timpi de înjumătățire (slide-ul „Un interval la frontieră").
  - (e) Comparați cuantila Student-t standardizată $\tilde{t}_{\hat{\nu}}^{-1}(0{,}01)$ (slide-ul „Patru prognoze VaR") cu $z_{0{,}01}$.
- **Raportați**
  - Estimările cu erorile standard; ambele CI pentru p și pentru timpul de înjumătățire; curba profil; cele două cuantile.
- **Interpretare**
  - Ce interval puteți raporta?
  - De ce nu este celălalt informativ sub această restricție?
  - Ce înseamnă $\hat{\nu} \approx 5$ pentru VaR?

◂ Înapoi ▸ Lista exercițiilor

B2 [Rezolvat]: Rezolvare (1/2)

- (a) $\hat{\alpha} = 0{,}192$ (se 0,025), $\hat{\beta} = 0{,}799$ (se 0,024), $\hat{\nu} = 5{,}01$ (se 0,31); $\hat{\rho} = 0{,}9907$.
- (b) Cu covarianța robustă, $\mathrm{se}(\hat{\rho}) = 0{,}0084$ (covarianța negativă dintre $\hat{\alpha}$ și $\hat{\beta}$ o face mai mică decît fiecare eroare standard).
- CI Wald: $0{,}9907 \pm 1{,}96 \cdot 0{,}0084 = [0{,}974;\ 1{,}007]$: capătul superior este peste 1, în afara spațiului impus al parametrilor $p \leq 1$.
- De ce nu este utilizabil sub restricția impusă: la frontieră distribuția lui $\hat{\rho}$ nu este Normală Andrews (1999), iar intervalul conține valori $p > 1$, pentru care dispersia necondiționată nu este finită și timpul de înjumătățire nu este definit.
- (d) Timpul de înjumătățire $\ln 0{,}5 / \ln 0{,}9907 = 74$ de zile; din intervalul Wald: 26 de zile la capătul inferior și nicio valoare la cel superior.

B2 [Rezolvat]: Rezolvare (2/2)



- (c) Statistica profil $2(\ell_{\max} - \ell(p))$: 4,32 la $p = 0{,}974$ (exclus), 3,83 la $p = 0{,}975$ (inclus), 0,99 la $p = 0{,}9999$ (inclus); la $p = 1$: $1{,}01 < 2{,}71$, valoarea p la frontieră 0,16 (inclus).
- CI profil [0,975; 1]: verosimilitatea nu exclude nicio valoare de la 0,975 pînă la GARCH integrat (IGARCH) inclusiv, $p = 1$.
- (d) Intervalul timpului de înjumătățire: cel puțin $\ln 0{,}5 / \ln 0{,}975 = 27$ de zile, fără o limită superioară finită, deoarece $p = 1$ este în interval.
- (e) $t_{\hat{\nu}}^{-1}(0{,}01)\sqrt{(\hat{\nu}-2)/\hat{\nu}} = -2{,}606$ față de $z_{\mathbf{0{,}01}} = -2{,}326$: la același σ_t, VaR 1% Student-t este de 1,12 ori cel Normal (medie zero).
- Interpretare 1: raportați intervalul profil; efectul șocurilor persistă cel puțin o lună, dar datele nu pot distinge o lună de cîțiva ani, iar IGARCH ($p = 1$) nu este respins.
- Interpretare 2: $\hat{\nu} \approx 5$ înseamnă cozi groase chiar și condiționat de σ_t; un VaR 1% GARCH Normal ar fi prea mic.

MFM_ch19_seminar

B3 [Rezolvat]: Backtesting VaR 1% și ES 2,5% pentru BET, 2005–2026 — Cerință

- **Întrebare**: Care dintre patru modele VaR 1% pentru BET ar fi fost fiabil în timp real, în 2005–2026?
- **Date și ipoteze**
  - Pierderile zilnice BET $L_t = -r_t$, evaluare 3 ian. 2005 – 18 sept. 2026, $T = 5\,432$ de zile; fiecare prognoză folosește doar datele pînă la $t-1$.
  - Patru modele pentru o zi McNeil & Frey (2000): HS (500 de zile), Normal (500 de zile), GARCH-t (fereastră de 1 000 de zile, reestimare la 21 de zile), FHS (filtrul GARCH-t), ca în slide-ul „Patru prognoze VaR"; VaR 1% și perechea (VaR 2,5%, ES 2,5%).
- **Cerințe**
  - (a) Numărați depășirile $L_t > \mathrm{VaR}_t$; calculați rata, p_{uc} Kupiec și p_{ind} Christoffersen Kupiec (1995), Christoffersen (1998).
  - (b) Aplicați testul DQ din slide-ul „Testele de independență și DQ" (depășirile din ultimele 4 zile și VaR_t) Engle & Manganelli (2004).
  - (c) Calculați scorul FZ0 mediu al perechii (VaR 2,5%, ES 2,5%), testele Diebold–Mariano (Newey–West, 9 laguri) și mulțimea de încredere a modelelor la 10% (slide-ul „Scoruri și compararea modelelor").
  - (d) Verificarea comparativă a ES față de HS Nolde & Ziegel (2017): cu t statistica DM pentru FZ0(model) − FZ0(HS), roșu dacă $t > 1{,}645$, verde dacă $t < -1{,}645$, galben în rest.
- **Raportați**
  - Tabelul de acoperire (a)–(b), depășirile cumulate, scorurile, valorile p MCS, zonele ES, VaR-ul mediu.
- **Interpretare**
  - Ce model îndeplinește toate criteriile acestui exercițiu?
  - Pe care l-ar prefera un trader?
  - Confirmă comparațiile formale backtest-urile?

◂ Înapoi ▸ Lista exercițiilor

B3: cuantile empirice și medii de coadă

- HS: cele mai recente 500 de pierderi zilnice
  - cuantila: interpolare liniară între statistici de ordine vecine
  - ES empiric: estimatorul din Capitolul 7; $500 \times 2{,}5\% = 12{,}5$, deci cele mai mari 12 pierderi plus jumătate din a 13-a, împărțite la 12,5
- FHS: cele 1 000 de reziduuri standardizate din fereastra de estimare GARCH
  - media cozii: același estimator; $1\,000 \times 2{,}5\% = 25$, deci media celor mai mici 25 de reziduuri
  - transformați-o cu media și volatilitatea condiționate prognozate, apoi inversați semnul pentru o pierdere
- Cînd $n\alpha$ nu este întreg, observația de la frontieră primește ponderea fracționară $n\alpha - \lfloor n\alpha \rfloor$
- Parametrii GARCH și fondul de reziduuri FHS se reestimează la 21 de zile de tranzacționare
- Volatilitatea condiționată se actualizează zilnic, numai cu informație trecută

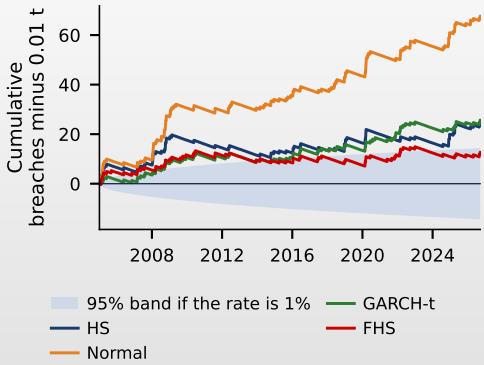
B3 [Rezolvat]: Rezolvare (1/3)

- (a)–(b) Pași: $\mathrm{Hit}_t = 1\{L_t > \mathrm{VaR}_t\}$; depășiri așteptate $0{,}01 \cdot 5\,432 \approx 54$; testele la 5%.

Model	x	Rata (%)	p_{uc}	p_{ind}	p_{DQ}
HS	79	1,45	0,002	<0,001	<0,001
Normal	122	2,25	<0,001	<0,001	<0,001
GARCH-t	80	1,47	0,001	0,864	0,007
FHS	67	1,23	0,095	0,267	0,129

- HS și Normal sînt respinse de testele de acoperire, de independență și DQ: reacționează prea încet, iar depășirile lor se grupează.
- Pentru GARCH-t independența nu este respinsă, dar are prea multe depășiri, deci este respins de testele de acoperire și DQ; doar FHS nu este respins de niciunul dintre cele trei teste.

B3 [Rezolvat]: Rezolvare (2/3)



- Graficul: depășirile cumulate minus cele așteptate, 0,01 t; o linie deasupra benzii albastru deschis are prea multe depășiri pentru un model de 1%
- La final doar FHS se află în interiorul benzii (numerele: tabelul de pe slide-ul anterior)
- Linia modelului Normal urcă brusc în 2008, 2020 și 2022: depășirile lui apar grupat, tiparul detectat de testele de independență și DQ

MFM_ch19_seminar

B3 [Rezolvat]: Rezolvare (3/3)

- (c)–(d) Un FZ0 mai mic este mai bun; DM t față de HS $< -1{,}645$ înseamnă că modelul este superior HS.

Model	FZ0	p_{MCS}	DM t față de HS	Zona ES	VaR mediu (%)
HS	1,422	0,00	–	standard	3,66
Normal	1,495	0,00	2,57	roșu	2,90
GARCH-t	1,175	0,18	−5,82	verde	2,95
FHS	1,157	1,00	−6,01	verde	3,17

- Nicio diferență FZ0 detectabilă statistic între FHS și GARCH-t (DM $t = 1{,}39$); împreună formează mulțimea de încredere a modelelor {GARCH-t, FHS}; Normal este roșu, GARCH-t și FHS sînt verzi.
- Interpretare 1: FHS este singurul model care îndeplinește toate criteriile enunțate (acoperire, independență, DQ, zona ES verde), adică sinteza criteriilor de evaluare folosite în acest exercițiu (nu lista de cerințe de reglementare din standardul Basel de backtesting); un trader vrea capital mic, iar modelul Normal are cel mai mic VaR mediu (2,90% față de FHS 3,17%), deși este respins de toate testele.
- Interpretare 2: comparațiile formale confirmă backtest-urile și adaugă un lucru: GARCH-t este respins de testul de acoperire, dar scorul lui FZ0 pentru (VaR, ES) nu diferă semnificativ de cel al FHS.

MFM_ch19_seminar

B4 [Rezolvat]: Deplasarea de anticipare — Cerință

- **Întrebare**: Cu cît par mai bune rezultatele unui backtest care folosește informație din viitor?
- **Date și ipoteze**
  - Pierderile BET din B3: evaluare 3 ian. 2005 – 18 sept. 2026, $T = 5\,432$ de zile, $0{,}01T \approx 54$ depășiri așteptate; rezultatele HS și GARCH-t pe fereastră mobilă din B3.
  - Variantele cu informație din viitor folosesc date care nu erau disponibile la $t-1$; testul Kupiec la 5%.
- **Cerințe**
  - (a) HS cu informație din viitor: fixați în fiecare zi $\mathrm{VaR}_{\mathbf{1\%}} = q_{\mathbf{0{,}99}}(L)$, cuantila empirică a tuturor pierderilor din perioada de evaluare; numărați depășirile și calculați rata.
  - (b) GARCH-t cu informație din viitor: estimați o singură dată pe tot eșantionul 6 ian. 2000 – 18 sept. 2026, filtrați $\hat{\sigma}_t$ zi de zi, fixați $\mathrm{VaR}_t = -(\hat{\mu} + \hat{\sigma}_t\, \tilde{t}_{\nu}^{-\mathbf{1}}(0{,}01))$ (slide-ul „Patru prognoze VaR"); numărați depășirile și aplicați testul Kupiec.
  - (c) Comparați ambele cu variantele pe fereastră mobilă din B3, într-un singur tabel.
- **Raportați**
  - Un tabel cu depășirile, rata și p_{uc} pentru cele patru variante.
- **Interpretare**
  - De ce pare mai bună varianta cu informație din viitor?
  - Ar fi putut-o folosi cineva în 2008?

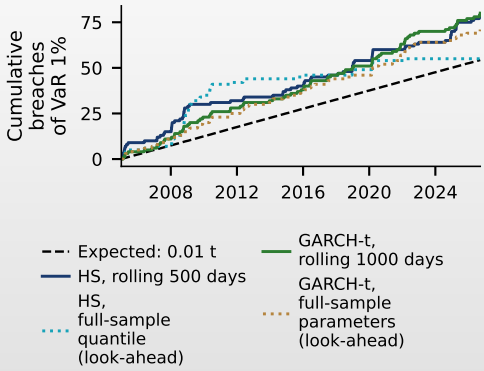
Înapoi Lista exercițiilor

B4 [Rezolvat]: Rezolvare (1/2)

- (a) HS pe tot eșantionul: $\mathrm{VaR}_{\mathbf{1\%}} = 4{,}24\%$ în fiecare zi; 55 de depășiri, rata 1,01%: aproximativ 1% prin construcție, deoarece 1% din pierderile unui eșantion depășesc propria lui cuantilă de 0,99.
- (b) GARCH-t pe tot eșantionul: 71 de depășiri, rata 1,31%, p_{uc} 0,030: încă respins la 5%, dar mult mai aproape de 1% decît modelul pe fereastră mobilă.

Model	Parametri	x	Rata (%)	p_{uc}
HS	tot eșantionul (din viitor)	55	1,01	0,93
HS	fereastră mobilă, 500 de zile	79	1,45	0,002
GARCH-t	tot eșantionul (din viitor)	71	1,31	0,030
GARCH-t	fereastră mobilă, 1 000 de zile	80	1,47	0,001

B4 [Rezolvat]: Rezolvare (2/2)



- (c) Observație: ambele variante cu informație din viitor au mai puține depășiri decît variantele pe fereastră mobilă și par mai bine calibrate.
- Explicație: tot eșantionul include 2008, 2020 și toate crahurile ulterioare, deci cuantila și parametrii au văzut deja pierderile pe care sînt testați.
- Interpretare: nimeni nu ar fi putut folosi aceste cifre în 2008, deoarece depind de pierderile din 2008–2026; doar rezultatul pe fereastră mobilă este o dovadă despre un model de prognoză.
- Verificare pentru propriul cod: testul trunchierii din slide-ul „Deplasarea de anticipare”.

MFM_ch19_seminar

C1 [Propus]: critica unui backtest scris de AI — răspunsul asistentului

Rezolvarea se discută la seminar

- **Cererea studentului**: „`bet` conține închiderile zilnice BET. Scrie un backtest pentru VaR 1% prin simulare istorică, cu fereastră de 500 de zile și testul Kupiec, și dă-mi o referință pentru rezultat."
- **Răspunsul asistentului AI**

```
import numpy as np
from scipy.stats import chi2
r = np.log(bet).diff().dropna()
var = -r.rolling(500).quantile(0.01)
hit = (-r > var)[var.notna()]
x, n, p = hit.sum(), hit.size, 0.01
pi = x / n
lr = -2 * (x*np.log(p) + (n-x)*np.log(1-p) - x*np.log(pi) - (n-x)*np.log(1-pi))
print(x, n, chi2.sf(lr, df=2))
```

- „Rata depășirilor este aproape de 1% și valoarea p a testului Kupiec este peste 5%: simularea istorică este nedeplasată pentru BET, cum au arătat Kupiec și Christoffersen (2004), *Backtesting historical simulation VaR*, Journal of Risk Model Validation, 8(2), 45–67."

Înapoi Lista exercițiilor

C1 [Propus]: critica unui backtest scris de AI — cerință

Rezolvarea se discută la seminar

- Răspunsul de pe slide-ul anterior conține cel puțin **trei** erori care apar des în codul de proiect scris de AI; evaluați și concluzia și verificați cum se comportă codul cînd nu există nicio depășire
  - Model: B3 și B4 (backtest și capcana deplasării de anticipare), A1 (testul Kupiec)
- Cerințe
  - (a) Găsiți fiecare eroare: ce rînd sau ce frază și de ce este greșită
  - (b) Explicați cum ați detecta-o fără să știți răspunsul: o verificare a informațiilor disponibile la data prognozei, numărul de restricții, o căutare după DOI (*Digital Object Identifier*, identificatorul digital al unei publicații)
  - (c) Scrieți codul corectat și indicați o referință corectă
  - (d) Documentați fiecare eroare ca în `AI_ERRORS.md`: ce a răspuns AI-ul, ce era greșit, cum v-ați dat seama, varianta corectă
- **Raportați**
  - Lista erorilor (rînd, motiv, test de detectare), codul corectat, referințele corecte și intrările `AI_ERRORS.md`.
- **Interpretare**
  - Ce afirmație despre BET poate susține backtest-ul corectat?

Exemplu de propunere (1/2)

- **Întrebare**: s-a schimbat riscul de coadă (*tail risk*) al BET după ce furnizorul de indici FTSE Russell a reclasificat BVB ca piață emergentă secundară (21 septembrie 2020)?
- **Ipoteză**: H_0: același indice al cozii Hill pentru pierderi și aceleași rate de depășire ale VaR 1% HS și FHS pe fereastră mobilă în cei șase ani de dinainte și de după
- **Date**: închideri oficiale BET, 22 sept. 2014 – 18 sept. 2026; control: WIG20 (Polonia); eliminate înregistrările duplicate (rînduri identice cu cel anterior) și rîndurile datate în weekend; prețurile BET și WIG20 aliniate întîi pe zilele comune, apoi randamentele
- **Identificare**: diferența în diferențe a indicilor cozii Hill (BET minus WIG20); date placebo $\pm$ un an
  - Ipoteze de identificare: fără eveniment, cei doi indici ai cozii s-ar fi modificat la fel; nicio propagare (*spillover*) a efectului reclasificării în Polonia; niciun șoc simultan specific uneia dintre cele două piețe
  - Diagnostic complementar de scală: un test sup-Wald de ruptură cu dată necunoscută Andrews (1993) pe $\ln|r_t|$ (detectează schimbări ale volatilității, nu ale indicelui cozii)

Exemplu de propunere (2/2)

- **Inferență**: bootstrap pe blocuri suprapuse (999 de reeșantionări, blocuri de 20 de zile)
  - În fiecare perioadă aceleași blocuri de zile pentru BET și WIG20, deci dependența dintre cele două piețe este păstrată; extrageri independente în cele două perioade disjuncte
- **Putere**: diferența în diferențe minimă detectabilă cu putere de 80% la nivelul 5%, $(z_{0,975} + z_{0,80})\,\mathrm{se} = 2{,}80\,\mathrm{se}$, cu z_p cuantila p a distribuției Normale standard și se eroarea standard a diferenței în diferențe (formula standard pentru puterea de 80%; C2(e))
- **Grila de specificații**: $k = 2{,}5\%$, 5%, 10% din zilele cu pierdere $\times$ blocuri de 10, 20, 40 de zile: 9 teste ale diferenței în diferențe, corecția Holm (A2)
- **Riscuri**: perioada de dinainte conține crahul din pandemia COVID-19 și căderea BET din 19 decembrie 2018, după anunțul taxei pe activele bancare (Ordonanța de urgență a Guvernului, OUG 114/2018); un singur eveniment, deci nicio concluzie cauzală fără control

C2 [Propus]: Propunerea echipei — Cerință (1/2)

Rezolvarea se discută la seminar

- **Întrebări**
  - Ce va studia echipa dumneavoastră?
  - Poate studiul propus să răspundă la întrebare?
- **Date și ipoteze**
  - O piață sau un activ la alegere: o serie din datele cursului (lista se află în repository-ul cursului) sau o sursă publică al cărei deținător oficial este precizat; date zilnice, dacă întrebarea nu cere altă frecvență.
  - Propunerea are secțiunile (a)–(g) de mai jos; exemplul de propunere de pe slide-urile anterioare este un model parțial; idei: Partea C din Seminarele 0–18.
- **Cerințe**
  - (a) Scrieți întrebarea într-o frază, cu H_0, H_1 și modelul de comparație pe care studiul trebuie să-l depășească.
  - (b) Numiți articolul publicat cel mai apropiat de întrebarea dumneavoastră și tabelul sau figura cu care vă comparați rezultatele, cu DOI și număr de tabel.
  - (c) Specificați datele: serii, sursă, perioadă, frecvență și reguli de curățare.
  - (d) Descrieți identificarea: piață de control, date placebo, test de ruptură cu dată necunoscută sau evaluare în afara eșantionului.
  - (e) Precizați testul principal și funcția lui de pierdere și calculați cel mai mic efect pe care studiul dumneavoastră îl detectează cu putere de 80% la T-ul dumneavoastră (ca la A1).
  - (f) Listați grila de specificații (ferestre, laguri, niveluri α, eșantioane), numărați testele și alegeți corecția pentru testele multiple (A2, A10).

Înapoi Lista exercițiilor

C2 [Propus]: Propunerea echipei — Cerință (2/2)

Rezolvarea se discută la seminar

- **Cerințe (continuare)**
  - (g) Enumerați riscurile principale: probleme de date, prea puține evenimente, supraajustare (*overfitting*).
- **Raportați**
  - Propunerea pe o pagină, cu 10–15 referințe, cel puțin 3 din ultimii cinci ani, fiecare verificată că există (DOI sau link), în termen de o săptămînă.
  - Model: exemplul de propunere
- **Verificări pentru propunere**
  - Se poate răspunde la întrebare cu date pe care le aveți deja?
  - Este modelul de comparație suficient de puternic?
  - Este funcția de pierdere fixată dinainte?
  - Folosește fiecare prognoză doar informația disponibilă la $t - 1$?

Proiectul de echipă: propunerea și repository-ul

- C2: propunerea de o pagină, în termen de o săptămînă
- README: întrebarea, datele, instrucțiunile de rulare
- Planul de analiză, fixat în scris printr-un commit înainte de a deschide perioada de test
- `AI_USE.md` și `AI_ERRORS.md`: orice utilizare AI și orice eroare detectată, ca în C1
- Pachetele de extindere sînt opționale; studiu suplimentar: Statistics of Financial Markets
- Criterii de evaluare
  - o întrebare precisă și un plan de analiză fixat înaintea perioadei de test
  - inferență corectă: erori standard HAC (*Heteroskedasticity and Autocorrelation Consistent*, robuste la heteroscedasticitate și autocorelație) sau bootstrap, testarea multiplă controlată
  - prognozele de risc validate prin backtesting în afara eșantionului
  - un repository reproductibil: codul rulează de la datele brute la tabele

Pachetele de extindere: conținut și cerințe prealabile

Pachet	Exerciții	Cerințe prealabile
1. Inferență pentru volatilitate	A3 [Rezolvat], A4, A5 [Rezolvat], B5	B2
2. Măsuri de risc și backtesting	A7 [Rezolvat], A8, A9, B6	A1, B3
3. Cozi și testare multiplă	A10, A11 [Rezolvat], B7, A12	A2, B1
4. Capcanele datelor	B8	B4
5. Erori nestandard	C3	A2, studiul de caz din curs

- Rezolvați exercițiile unui pachet în ordinea dată; fiecare indică modelul, iar secțiunile notebook-ului poartă aceleași etichete
- Util pentru proiectul de echipă: C3 poate fi punctul de plecare al unui proiect

A3 [Rezolvat]: Sandvișul QMLE pentru GARCH(1,1) — Cerință

- **Întrebare**: Cît subestimează erorile standard doar din hessiană incertitudinea GARCH cînd inovațiile au cozi groase?
- **Date și ipoteze**
  - Modelul: $r_t = \sigma_t z_t$, $\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2$, $\theta = (\omega, \alpha, \beta)^\top$; z_t i.i.d. cu $\mathbb{E}z_t = 0$, $\mathbb{E}z_t^2 = 1$, $\kappa = \mathbb{E}z_t^4 < \infty$.
  - Estimare prin cvasi-verosimilitatea Gaussiană $\ell_t(\theta) = -\frac{1}{2}(\ln \sigma_t^2 + r_t^2/\sigma_t^2)$; notația $g_t = \sigma_t^{-2}\, \partial \sigma_t^2 / \partial \theta$, cunoscută la $t-1$ Bollerslev & Wooldridge (1992).
  - Pentru (c): estimarea GARCH(1,1)-t pe BET din B2 dă $\hat{\nu} = 5{,}011$; pentru inovații Student-t, $\kappa = 3(\nu-2)/(\nu-4)$ cînd $\nu > 4$.
- **Cerințe**
  - (a) Derivați ℓ_t în raport cu θ, scrieți scorul $s_t = \partial \ell_t / \partial \theta$ în funcție de z_t și g_t și arătați că $\mathbb{E}[s_t \mid \mathcal{F}_{t-1}] = 0$.
  - (b) Calculați $A = -\mathbb{E}[\partial s_t / \partial \theta^\top]$ și $B = \mathbb{E}[s_t s_t^\top]$ folosind $\mathbb{E}z_t^2 = 1$ și $\mathbb{E}z_t^4 = \kappa$ și arătați că sandvișul este $A^{-1} B A^{-1} = \frac{\kappa - 1}{2} A^{-1}$.
  - (c) Înlocuiți $\hat{\nu} = 5{,}011$: calculați κ, factorul de corecție $(\kappa - 1)/2$ pentru dispersii și $\sqrt{(\kappa-1)/2}$ pentru erorile standard.
- **Raportați**
  - Expresiile lui s_t, A și B; valorile lui κ și ale celor doi factori de corecție.
- **Interpretare**
  - Cînd sînt corecte erorile standard doar din hessiană (rădăcinile diagonalei lui A^{-1}/T)?
  - Cu cît exagerează ele statisticile t pentru BET?
  - Ce se întîmplă cînd $\nu \leq 4$?

◂ Înapoi ▸ Lista exercițiilor

A3 [Rezolvat]: Rezolvare (1/2)

- (a) Pasul 1: $\partial \ell_t / \partial \sigma_t^2 = -\frac{1}{2}\left(\frac{1}{\sigma_t^2} - \frac{r_t^2}{\sigma_t^4}\right) = \frac{1}{2\sigma_t^2}(z_t^2 - 1)$.
- Pasul 2 (derivarea funcției compuse): $s_t = \frac{1}{2}(z_t^2 - 1)\, g_t$; g_t este cunoscut la $t-1$ și $\mathbb{E}[z_t^2 - 1] = 0$, deci $\mathbb{E}[s_t \mid \mathcal{F}_{t-1}] = 0$: scorul este o diferență de martingal.
- (b) Pasul 3: $\partial z_t^2 / \partial \theta = -z_t^2 g_t$, deci $\frac{\partial s_t}{\partial \theta^\top} = \frac{1}{2}(z_t^2 - 1)\frac{\partial g_t}{\partial \theta^\top} - \frac{1}{2} z_t^2\, g_t g_t^\top$.
- Pasul 4: primul termen are media condiționată 0 și $\mathbb{E} z_t^2 = 1$, deci $A = \frac{1}{2}\mathbb{E}[g_t g_t^\top]$.
- Pasul 5: $\mathbb{E}(z_t^2 - 1)^2 = \kappa - 2 + 1 = \kappa - 1$, deci $B = \frac{1}{4}(\kappa - 1)\mathbb{E}[g_t g_t^\top] = \frac{\kappa - 1}{2} A$.

A3 [Rezolvat]: Rezolvare (2/2)

- Pasul 6: $A^{-1}BA^{-1} = \frac{\kappa-1}{2}A^{-1}AA^{-1} = \frac{\kappa-1}{2}A^{-1}$; este egal cu matricea doar din hessiană A^{-1} numai dacă $\kappa = 3$: kurtosis-ul distribuției Normale, dar și al altor distribuții; egalitatea privește blocul parametrilor variației din QMLE Gaussian.
- (c) Pasul 7: $\kappa = 3(\hat{\nu} - 2)/(\hat{\nu} - 4) = 3 \cdot 3{,}011/1{,}011 = 8{,}93$.
- Pasul 8: dispersiile $\times\,(\kappa - 1)/2 = 3{,}97$; erorile standard $\times\sqrt{3{,}97} = 1{,}99$.
- Interpretare: cu cozi ca ale BET, erorile standard doar din hessiană sînt cam jumătate din cele corecte, deci statisticile t sînt de aproximativ două ori prea mari; raportați sandvișul Bollerslev & Wooldridge (1992); factorul se aplică QMLE Gaussian, nu hessianei estimatorului Student-t din studiul de caz.
- Cu $\nu \leq 4$, $\kappa = \infty$, B nu există, iar QMLE nu mai este $\sqrt{T}$-Normal: intervalele obișnuite nu mai sînt justificate.

A4 [Propus]: Metoda delta pentru timpul de înjumătățire — Cerință

Rezolvarea se discută la seminar

- **Întrebare**: Poate un interval obținut prin metoda delta să descrie incertitudinea timpului de înjumătățire al volatilității?
- **Date și ipoteze**
  - Timpul de înjumătățire al unui șoc de volatilitate $h(p) = \ln 0{,}5 / \ln p$, unde $p = \alpha + \beta$ este persistența unui GARCH(1,1).
  - Estimările BET din B2, cu covarianță robustă: $\hat{p} = 0{,}9907$, $\mathrm{se}(\hat{p}) = 0{,}0084$; presupuneți pentru acest calcul prin metoda delta că $\hat{p}$ este asimptotic Normal (A3); evaluați aproximarea aproape de frontieră; nivelul 95%.
- **Cerințe**
  - (a) Derivați $h'(p)$ și scrieți eroarea standard prin metoda delta $\mathrm{se}(\hat{h}) = |h'(\hat{p})| \, \mathrm{se}(\hat{p})$.
  - (b) Calculați $\hat{h}$, $|h'(\hat{p})|$, $\mathrm{se}(\hat{h})$ și intervalul Wald $\hat{h} \pm 1{,}96 \, \mathrm{se}(\hat{h})$.
  - (c) Folosind $\ln p \approx -(1-p)$, arătați că $|h'(p)| \approx \ln 2 / (1-p)^2$ cînd $p \to 1$; evaluați panta exactă și cea aproximativă la $p = 0{,}99$, $0{,}995$ și $0{,}999$.
- **Raportați**
  - $\hat{h}$ cu eroarea standard și intervalul Wald; un tabel cu $|h'(p)|$, exact și aproximativ.
- **Interpretare**
  - De ce nu are sens aici un interval Normal pentru timpul de înjumătățire?
  - Ce ar trebui raportat în schimb?
- **Model**: A3

înapoi Lista exercițiilor

A5 [Rezolvat]: Prognoza GARCH(1,1) pentru BET — Cerință

Întrebări

- Ce risc implică modelul GARCH pentru BET în următoarele 10 zile?
- Este corectă regula rădăcinii timpului?

Date și ipoteze

- Randamentele log zilnice BET în % (închideri oficiale, datele cursului), 6 ian. 2000 – 18 sept. 2026; GARCH(1,1)-t pe tot eșantionul: $\omega = 0{,}0448$, $\alpha = 0{,}1916$, $\beta = 0{,}7991$.
- Varianța de mîine, prognoza pe un pas $\sigma^2_{T+1} = \omega + \alpha\varepsilon^2_T + \beta\sigma^2_T = 3{,}11$ ($\sigma_{T+1} = 1{,}76\%$); prognozele σ^2_{T+h} și varianța de termen lung $\bar{\sigma}^2$ ca în slide-ul „Erori robuste și prognoze GARCH"; 252 de zile de tranzacționare pe an.

Cerințe

- (a) Calculați persistența $\alpha + \beta$, varianța de termen lung $\bar{\sigma}^2$, volatilitatea zilnică $\bar{\sigma}$ și volatilitatea anualizată $\sqrt{252\,\bar{\sigma}^2}$.
- (b) Calculați timpul de înjumătățire al unui șoc (slide-ul „Un interval la frontieră").
- (c) Prognozați varianța în ziua 10, σ^2_{T+10}.
- (d) Calculați varianța randamentului pe 10 zile $\sum_{h=1}^{10} \sigma^2_{T+h}$ cu forma închisă a sumei geometrice, volatilitatea pe 10 zile, și comparați-o cu regula rădăcinii timpului $\sqrt{10}\,\sigma_{T+1}$.

Raportați

- Fiecare cifră de la (a)–(d), cu formula în care s-au înlocuit valorile.

Interpretare

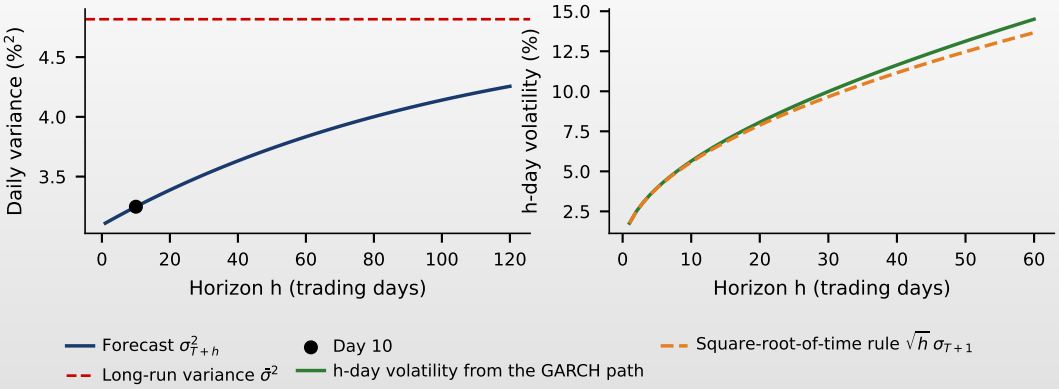
- De ce converge prognoza spre nivelul de termen lung?
- Cînd subestimează sau supraestimează regula rădăcinii timpului riscul pe 10 zile?

◂ Înapoi ▸ Lista exercițiilor

A5 [Rezolvat]: Rezolvare (1/3)

- (a) Pasul 1: $\alpha + \beta = 0{,}9907$, deci $1 - \alpha - \beta = 0{,}0093$.
- Pasul 2: $\bar{\sigma}^2 = 0{,}0448/0{,}0093 = 4{,}82$, adică $\bar{\sigma} = 2{,}20\%$ pe zi.
- Pasul 3: $\sqrt{252 \cdot 4{,}82} = \sqrt{1\,214{,}6} = 34{,}9\%$ pe an.
- (b) Pasul 4: $h_{1/2} = \ln 0{,}5 / \ln 0{,}9907 = 74{,}2$ zile de tranzacționare, aproximativ 3,5 luni de cîte 21 de zile de tranzacționare.
- Observație: un șoc asupra variației își pierde jumătate din efect în aproximativ 74 de zile de tranzacționare, deci volatilitatea BET este foarte persistentă.
- Observație: $\bar{\sigma}^2$ este imprecisă, deoarece numitorul $1 - \alpha - \beta = 0{,}0093$ este mic; intervalul profil din B2 ajunge la $p = 1$, unde $\bar{\sigma}^2$ este infinită.

A5 [Rezolvat]: Rezolvare (2/3)



MFM_ch19_seminar

A5 [Rezolvat]: Rezolvare (3/3)

- (c) Pasul 5: abaterea față de nivelul pe termen lung $\sigma^2_{T+1} - \bar{\sigma}^2 = 3{,}11 - 4{,}82 = -1{,}71$; $(\alpha+\beta)^9 = 0{,}919$.
- Pasul 6: $\sigma^2_{T+10} = 4{,}82 + 0{,}919 \cdot (-1{,}71) = 3{,}25$: în prezent volatilitatea este sub nivelul de termen lung, deci prognoza crește.
- (d) Pasul 7: $\sum_{h=1}^{10} \sigma^2_{T+h} = 10\bar{\sigma}^2 + (\sigma^2_{T+1} - \bar{\sigma}^2)\frac{1-(\alpha+\beta)^{10}}{1-\alpha-\beta} = 10 \cdot 4{,}82 + (-1{,}71) \cdot 9{,}592 = 31{,}80$.
- Pasul 8: volatilitatea pe 10 zile $\sqrt{31{,}80} = 5{,}64\%$, față de $\sqrt{10} \cdot 1{,}76 = 5{,}57\%$ după regula rădăcinii timpului.
- Interpretare: varianța revine spre $\bar{\sigma}^2$ cu rata $\alpha + \beta$; regula rădăcinii timpului subestimează riscul pe 10 zile cînd volatilitatea este sub nivelul de termen lung (ca acum) și îl supraestimează după un vîrf de criză.

MFM_ch19_seminar

Merită asimetria un parametru în plus?

- Glosten–Jagannathan–Runkle (GJR) GARCH adaugă o reacție specifică șocurilor negative
  - coeficientul suplimentar este zero sub modelul GARCH simetric
  - ambele modele folosesc aceleași date
- Criteriile informaționale Akaike (*Akaike Information Criterion*, AIC) și bayesian (*Bayesian Information Criterion*, BIC) pun în balanță verosimilitatea estimată și numărul parametrilor
  - preferăm valori mai mici, comparînd modele cu aceeași variabilă dependentă și același eșantion
  - pentru un parametru suplimentar, AIC adaugă o penalizare de 2, iar BIC una de $\ln T$
  - într-un eșantion mare, aceeași îmbunătățire a verosimilității poate satisface AIC, dar nu BIC
- Testul LR verifică dacă reacția suplimentară este susținută de date
- Criteriile informaționale aleg între specificații
- Pentru prognoza riscului, niciunul nu înlocuiește backtest-ul VaR în afara eșantionului

B5 [Propus]: Asimetrie: GJR-GARCH-t pentru BET — Cerință

Rezolvarea se discută la seminar

- **Întrebare**: Cresc șocurile negative volatilitatea BET mai mult decît cele pozitive?
- **Date și ipoteze**
  - Randamentele BET ca la B2; modelul GJR GARCH(1,1)-t (slide-ul anterior): $\sigma_t^2 = \omega + (\alpha + \gamma\, 1\{\varepsilon_{t-1} < 0\})\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$; GARCH(1,1)-$t$ din B2 este modelul restricționat $\gamma = 0$.
  - $T = 6\,670$; $\mathrm{AIC} = -2\ell + 2k$, $\mathrm{BIC} = -2\ell + k \ln T$ (ℓ: log-verosimilitatea maximizată, k = numărul de parametri; o valoare mai mică este mai bună); nivelul 5%.
- **Cerințe**
  - (a) Estimați modelul GJR și raportați $\hat{\gamma}$, eroarea standard robustă și $t = \hat{\gamma}/\mathrm{se}(\hat{\gamma})$.
  - (b) Calculați $LR = 2(\ell_{\mathrm{GJR}} - \ell_{\mathrm{GARCH}})$ și valoarea p din $\chi^2(1)$.
  - (c) Calculați AIC și BIC pentru ambele modele și arătați că $\Delta\mathrm{AIC} = 2 - LR$ și $\Delta\mathrm{BIC} = \ln T - LR$ (GJR minus GARCH).
  - (d) Calculați $(\hat{\alpha} + \hat{\gamma})/\hat{\alpha}$, impactul unui șoc negativ relativ la unul pozitiv.
- **Raportați**
  - $\hat{\gamma}$ cu eroarea standard și t; LR și p; ambele criterii pentru ambele modele; raportul de la (d).
- **Interpretare**
  - Conduc testele și criteriile informaționale la aceeași alegere?
  - Ce model ați folosi pentru VaR?
  - Cum ați decide?
- **Model**: B2

◂ Înapoi ▸ Lista exercițiilor

A7 [Rezolvat]: VaR 1% și ES 2,5% sub distribuția Normală — Cerință

- **Întrebare**: Cît de mari sînt VaR 1% și ES 2,5% pentru o poziție pe BET, sub distribuția Normală?
- **Date și ipoteze**
  - O poziție de 1 milion RON (*Romanian leu*, leul românesc) în BET; randamentul log pe o zi $r \sim N(\mu, \sigma^2)$ cu $\mu = 0{,}05\%$, $\sigma = 1{,}4\%$; pierderea $L = -r$; sumele în RON folosesc aproximarea liniară $10^6 \cdot L/100$ (pierderea exactă: $10^6\,[1 - e^{r/100}]$).
  - Definițiile și formulele VaR și ES pentru distribuția Normală: slide-ul „VaR și ES sub distribuția Normală și Student-t".
- **Cerințe**
  - (a) Găsiți $z_{0{,}01}$ și calculați $\mathrm{VaR}_{1\%}$ în % și în RON.
  - (b) Găsiți $z_{0{,}025}$ și $\varphi(z_{0{,}025})$ și calculați $\mathrm{ES}_{2{,}5\%}$ în % și în RON.
  - (c) Calculați raportul $\mathrm{ES}_{2{,}5\%}/\mathrm{VaR}_{1\%}$.
- **Raportați**
  - Ambele măsuri în % și în RON, precum și raportul lor.
- **Interpretare**
  - De ce a înlocuit Comitetul de la Basel VaR 1% cu ES 2,5%, dacă pentru distribuția Normală cele două cifre sînt aproape egale?

înapoi Lista exercițiilor

A7 [Rezolvat]: Rezolvare

- (a) Pasul 1: $z_{0,01} = -2{,}326$, deci $\sigma z_{0,01} = 1{,}4 \cdot (-2{,}326) = -3{,}257$.
- Pasul 2: $\mathrm{VaR}_{1\%} = -(0{,}05 - 3{,}257) = 3{,}207\%$, adică $3{,}207\% \cdot 10^6 = 32\,069$ RON.
- (b) Pasul 3: $z_{0,025} = -1{,}960$, $\varphi(z_{0,025}) = 0{,}0584$, $\varphi(z_{0,025})/0{,}025 = 2{,}338$.
- Pasul 4: $\mathrm{ES}_{2,5\%} = -0{,}05 + 1{,}4 \cdot 2{,}338 = -0{,}05 + 3{,}273 = 3{,}223\%$, adică 32 229 RON.
- (c) Pasul 5: $\mathrm{ES}_{2,5\%}/\mathrm{VaR}_{1\%} = 3{,}223/3{,}207 = 1{,}005$: pentru distribuția Normală cele două cifre sînt aproape egale.
- Interpretare: pentru distribuția Normală, schimbarea nu modifică nivelul; ES măsoară mărimea pierderilor din coadă și este subaditiv, iar VaR marchează doar intrarea în coadă (Capitolul 7); cu cozi groase ES crește mai mult (A8).

A8 [Propus]: VaR 1% și ES 2,5% Student-t — Cerință

Rezolvarea se discută la seminar

- **Întrebare**: Cum se schimbă cele două măsuri cînd cozile sînt groase, dar dispersia este aceeași?
- **Date și ipoteze**
  - Aceeași poziție, $\mu = 0{,}05\%$ și $\sigma = 1{,}4\%$ ca la A7, dar $r = \mu + s\,T_\nu$, T_ν Student-t cu $\nu = 5$, scala $s = \sigma\sqrt{(\nu-2)/\nu}$, astfel încît dispersia este σ^2.
  - Formule: $\mathrm{VaR}_\alpha = -(\mu + s\,q_\alpha)$ cu $q_\alpha = t_\nu^{-1}(\alpha)$; $\mathrm{ES}_\alpha = -\mu + s\,\frac{f_\nu(q_\alpha)}{\alpha}\,\frac{\nu+q_\alpha^2}{\nu-1}$, cu f_ν densitatea Student-t.
- **Cerințe**
  - (a) Calculați scala s.
  - (b) Găsiți $t_5^{-1}(0{,}01)$ și calculați $\mathrm{VaR}_{1\%}$ în % și în RON.
  - (c) Găsiți $q = t_5^{-1}(0{,}025)$, calculați factorii $f_5(q)/0{,}025$ și $(5+q^2)/4$, apoi $\mathrm{ES}_{2{,}5\%}$ în % și în RON.
  - (d) Calculați rapoartele Student-t/Normal pentru VaR și pentru ES (valorile Normale din A7) și raportul ES/VaR.
- **Raportați**
  - Ambele măsuri în % și în RON și cele trei rapoarte.
- **Interpretare**
  - Care cifră se schimbă mai mult cînd cozile devin groase, VaR sau ES?
  - Ce înseamnă aceasta pentru capital?
- **Model**: A7

Înapoi Lista exercițiilor

A9 [Propus]: Puterea testului Christoffersen — Cerință

Rezolvarea se discută la seminar

- **Întrebare**: Poate un an de depășiri să arate că acestea se grupează?
- **Date și ipoteze**
  - Depășirile I_t urmează un lanț Markov cu două stări, cu rata necondiționată corectă $\pi = 1\%$, dar grupate: $\pi_{11} = P(I_t = 1 \mid I_{t-1} = 1) = 0{,}20$ (independența ar însemna $\pi_{11} = 1\%$).
  - Testul de independență Christoffersen $LR_{ind} \sim \chi^2(1)$, respingere la 5% dacă $LR_{ind} > 3{,}84$ Christoffersen (1998); o traiectorie fără nicio depășire contează ca nerespingere.
- **Cerințe**
  - (a) Derivați $\pi_{01} = P(I_t = 1 \mid I_{t-1} = 0)$ din condiția de staționaritate $\pi = \pi_{01}/(1 - \pi_{11} + \pi_{01})$.
  - (b) Simulați 20 000 de traiectorii de $T = 250$ de zile (prima zi este depășire cu probabilitatea π), calculați LR_{ind} pe fiecare traiectorie și estimați puterea ca pondere a respingerilor, cu eroarea standard Monte Carlo.
  - (c) Repetați (b) pentru $T = 1\,000$ de zile, cu 5 000 de traiectorii.
  - (d) Calculați numărul așteptat de perechi de depășiri consecutive $(T-1)\pi\pi_{11}$ pentru ambele valori T.
- **Raportați**
  - π_{01}, ambele puteri cu eroarea standard Monte Carlo, numerele așteptate de perechi.
- **Interpretare**
  - De ce aproape nu poate fi testată independența cu un an de date?
- **Model**: A1

◂ Înapoi ▸ Lista exercițiilor

B6 [Propus]: Backtesting VaR și ES pentru Bitcoin, 2019–2026 — Cerință

Rezolvarea se discută la seminar

- **Întrebare**: Se păstrează clasamentul din B3 pentru Bitcoin?
- **Date și ipoteze**
  - Pierderile zilnice Bitcoin (închiderea în USD, datele cursului), evaluare 1 ian. 2019 – 18 sept. 2026, $T = 2\,818$ de zile, fiecare zi calendaristică; fiecare prognoză folosește doar datele pînă la $t - 1$.
  - Cele patru modele, ferestrele și testele din B3; dispersia Newey–West cu 8 laguri; nivelul 5% (10% pentru mulțimea de încredere a modelelor).
- **Cerințe**
  - (a) Numărați depășirile VaR 1% și calculați rata, p_{uc} și p_{ind} pentru fiecare model.
  - (b) Aplicați testul DQ cu indicatorii depășirilor din cele 4 zile anterioare și VaR_t.
  - (c) Calculați scorurile FZ0, testele Diebold–Mariano și mulțimea de încredere a modelelor la 10%.
  - (d) Aplicați verificarea comparativă a ES față de HS.
  - (e) Reprezentați pierderile zilnice și VaR 1% HS și FHS din 1 ianuarie 2022.
- **Raportați**
  - Cele două tabele din B3 pentru Bitcoin și graficul de la (e).
- **Interpretare**
  - Este clasamentul modelelor același ca la BET?
  - Pot testele să discrimineze între modele?
- **Model**: B3

înapoi Lista exercițiilor

A10 [Propus]: Praguri pentru o sută de factori — Cerință

Rezolvarea se discută la seminar

- **Întrebare**: Cît de înalt trebuie să fie pragul pentru $|t|$ cînd sînt testați o sută de factori?
- **Date și ipoteze**
  - 100 de factori candidați, fiecare testat cu o statistică t bilaterală, aproximativ $N(0,1)$ sub H_0; nivelul țintă 5%.
  - Bonferroni și BH ca la A2; la dependență arbitrară, procedura Benjamini–Yekutieli înlocuiește $0{,}05k/m$ cu $0{,}05k/[m\,c(m)]$, $c(m) = \sum_{i=1}^{m} 1/i$ Benjamini & Yekutieli (2001).
- **Cerințe**
  - (a) Calculați nivelul Bonferroni pe test la FWER 5% și pragul $|t| > z_{1-0{,}05/(2m)}$.
  - (b) Calculați pragul BH pentru $|t|$ cînd sînt declarați exact 10 factori și cînd este declarat doar 1.
  - (c) Calculați $c(100)$ și pragul Benjamini–Yekutieli cînd sînt declarați 10 factori.
  - (d) Comparați pragurile cu $|t| > 3$ propus de Harvey, Liu & Zhu (2016).
- **Raportați**
  - Nivelurile pe test și cele patru praguri pentru $|t|$.
- **Interpretare**
  - De ce depinde pragul BH de numărul factorilor declarați?
  - Mai este valid BH cînd factorii sînt corelați pozitiv?
- **Model**: A2

înapoi Lista exercițiilor

A11 [Rezolvat]: Kurtosis-ul de selecție cînd indicele cozii este sub 4 — Cerință

- **Întrebare**: Ce estimează kurtosis-ul de selecție cînd indicele cozii este sub 4?
- **Date și ipoteze**
  - $|X|$ are o coadă de tip Pareto: $P(|X| > x) = Cx^{-\alpha}$ pentru $x \geq x_0 > 0$, $C > 0$; α este indicele cozii.
  - Configurația simulării: Student-t cu 3 grade de libertate (indicele cozii 3, dispersie finită), mărimi de eșantion $T = 10^3$, 10^4, 10^5, cîte 200 de eșantioane pentru fiecare mărime.
- **Cerințe**
  - (a) Folosiți $\mathbb{E}|X|^4 = \int_0^\infty 4x^3 P(|X| > x)\, dx$ pentru a arăta că $\mathbb{E}|X|^4 = \infty$ cînd $\alpha \leq 4$.
  - (b) Simulați eșantioanele, calculați excesul de kurtosis de selecție al fiecăruia și raportați mediana și cuantila de 90% pentru fiecare T.
  - (c) Explicați ce estimează în acest caz un interval bootstrap pentru kurtosis Athreya (1987).
- **Raportați**
  - Demonstrația de la (a) și un tabel cu mediana și cuantila de 90% pentru cele trei mărimi de eșantion.
- **Interpretare**
  - Ce ar trebui raportat în locul kurtosis-ului cînd indicele cozii este sub 4?

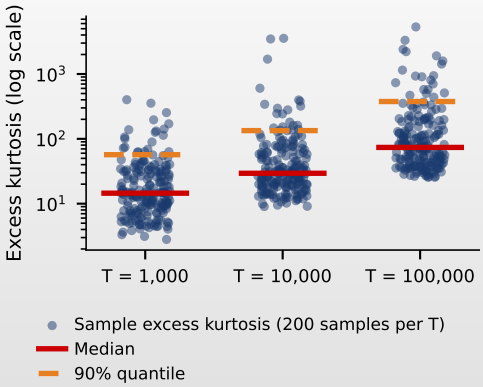
Înapoi Lista exercițiilor

A11 [Rezolvat]: Rezolvare (1/2)

- (a) Pasul 1: pentru $x \geq x_0$ integrandul este $4x^3 \cdot Cx^{-\alpha} = 4Cx^{3-\alpha}$ și este nenegativ peste tot.
- Pasul 2: $\mathbb{E}|X|^4 \geq 4C \int_{x_0}^{\infty} x^{3-\alpha}\, dx$, iar această integrală diverge cînd $3 - \alpha \geq -1$, adică $\alpha \leq 4$.
- Pasul 3: pentru $2 < \alpha \leq 4$ dispersia este finită, dar $\mathbb{E}X^4 = \infty$, deci kurtosis-ul populației $\mathbb{E}X^4/(\mathbb{E}X^2)^2$ este infinit, iar kurtosis-ul de selecție nu are o limită finită.
- (b) Pași: extragem 200 de eșantioane din fiecare mărime, calculăm excesul de kurtosis de selecție, luăm mediana și cuantila de 90% din cele 200 de valori.

T	10^3	10^4	10^5
Excesul de kurtosis median	14,5	29,4	73,7
Cuantila de 90%	56,9	134,0	376,4

A11 [Rezolvat]: Rezolvare (2/2)



- Observație: mediana crește cu T, iar cuantila de 90% crește mai repede: statistica nu se stabilizează, ci crește fără limită.
- (c) Nu există o valoare a populației pe care un interval să o acopere, iar bootstrap-ul unor astfel de momente este inconsistent Athreya (1987): un interval bootstrap pentru kurtosis nu are nicio garanție de acoperire.
- Interpretare: cînd indicele cozii poate fi sub 4, raportați indicele cozii estimat, cu intervalul lui (Hill cu bootstrap pe blocuri, B1), nu kurtosis-ul; un kurtosis de selecție mare spune doar că cozile sînt groase, nu cît de groase.

MFM_ch19_seminar

B7 [Propus]: Faptele stilizate ale Bitcoin — Cerință

Rezolvarea se discută la seminar

- **Întrebare**: Se păstrează faptele stilizate din B1 pentru Bitcoin?
- **Date și ipoteze**
  - Bitcoin în USD (închiderea zilnică, datele cursului), în fiecare zi calendaristică, 18 sept. 2014 – 18 sept. 2026: $T = 4\,384$ de randamente log în %.
  - Aceleași setări ca la B1: $k = 5\%$ din zilele cu pierdere, 999 de reeșantionări bootstrap cu blocuri de 20 de zile, 10 laguri, nivelul 5%; rezultatele BET din B1.
- **Cerințe**
  - (a) Calculați abaterea standard, asimetria și excesul de kurtosis.
  - (b) Estimați indicele cozii Hill cu CI de 95% asimptotic și prin bootstrap pe blocuri.
  - (c) Calculați $Q(10)$ pe r_t și pe r_t^2 și $Q^*(10)$ robust și decideți la 5%.
  - (d) Numărați zilele dincolo de ± 4 abateri standard și comparați cu distribuția Normală.
  - (e) Testați H_0: $\alpha_{\mathrm{BTC}} = \alpha_{\mathrm{BET}}$ cu $z = (\hat{\alpha}_{\mathrm{BTC}} - \hat{\alpha}_{\mathrm{BET}}) / \sqrt{\mathrm{se}^2_{\mathrm{BTC}} + \mathrm{se}^2_{\mathrm{BET}}}$, folosind erorile standard bootstrap pe blocuri și presupunînd estimări independente (testul este detaliat în A12).
- **Raportați**
  - Un tabel Bitcoin față de BET cu (a)–(d); testul de la (e) cu z, p și decizia.
- **Interpretare**
  - Are Bitcoin cozi mai groase decît BET sau este doar mai volatil?
  - Rămîne autocorelația liniară a Bitcoin semnificativă în testul robust?
- **Model**: B1

◂ Înapoi ▸ Lista exercițiilor

A12 [Propus]: Compararea a doi indici ai cozii — Cerință

Rezolvarea se discută la seminar

- **Întrebare**: Are Bitcoin cozi mai groase decît BET?
- **Date și ipoteze**
  - Estimatorul Hill $\hat{\alpha}$ din slide-ul „Indicele cozii și estimatorul Hill", $k = 5\%$ din zilele cu pierdere; $\mathrm{se}(\hat{\alpha}) \approx \hat{\alpha}/\sqrt{k}$ pentru date i.i.d. Hill (1975).
  - Estimări date din B1 și B7 (nu trebuie rulate în prealabil): BET $\hat{\alpha} = 2{,}39$ ($k = 153$), Bitcoin $\hat{\alpha} = 2{,}74$ ($k = 104$); erori standard bootstrap pe blocuri (999 de reeșantionări, blocuri de 20 de zile): BET 0,233, Bitcoin 0,227; ipoteză simplificatoare: cele două estimări sînt independente (eșantioanele se suprapun în timp, deci șocurile comune ar adăuga $-2\,\mathrm{Cov}(\hat{\alpha}_{\mathrm{BTC}}, \hat{\alpha}_{\mathrm{BET}})$ la dispersia diferenței); nivelul 5%.
- **Cerințe**
  - (a) Calculați erorile standard i.i.d. ale ambelor estimări.
  - (b) Testați H_0: $\alpha_{\mathrm{BTC}} = \alpha_{\mathrm{BET}}$ cu $z = (\hat{\alpha}_{\mathrm{BTC}} - \hat{\alpha}_{\mathrm{BET}})/\sqrt{\mathrm{se}^2_{\mathrm{BTC}} + \mathrm{se}^2_{\mathrm{BET}}}$, bilateral, cu erorile standard i.i.d.
  - (c) Repetați testul cu erorile standard bootstrap pe blocuri.
- **Raportați**
  - Toate erorile standard, ambele statistici z, valorile p și deciziile.
- **Interpretare**
  - De ce poate formula i.i.d. să evalueze greșit incertitudinea pentru randamente?
  - Depinde concluzia de ea?
- **Model**: A11, B1

◂ Înapoi ▸ Lista exercițiilor

B8 [Propus]: Alinierea calendarelor de tranzacționare pentru Bitcoin și SPY — Cerință

Rezolvarea se discută la seminar

- **Întrebare**: Ce efect are alinierea randamentelor în locul prețurilor asupra statisticilor Bitcoin–SPY?
- **Date și ipoteze**
  - Bitcoin (închiderea zilnică în USD, în fiecare zi calendaristică) și SPY, ETF-ul pe S&P 500 (închiderea ajustată pentru dividende și divizări, doar zilele lucrătoare), din datele cursului; randamente log zilnice în %.
  - Două perioade: 1 ianuarie 2015 – 31 decembrie 2019 (1 258 de zile comune) și 1 ianuarie 2020 – 18 septembrie 2026 (1 687 de zile comune).
- **Cerințe**
  - (a) Aliniere greșită: calculați randamentele log ale fiecărui activ pe propriul calendar, apoi păstrați zilele comune.
  - (b) Aliniere corectă: păstrați zilele comune ale prețurilor, apoi calculați randamentele log (un randament de luni acoperă atunci intervalul vineri–luni pentru ambele active).
  - (c) Pentru fiecare perioadă și fiecare aliniere calculați corelația Bitcoin–SPY, abaterea standard a randamentelor Bitcoin în toate zilele și în zilele de luni și randamentul cumulat Bitcoin în %, $100\,[\exp(\sum_t r_t/100) - 1]$ (cu r_t în %).
  - (d) Reprezentați corelația pe fereastră mobilă de 250 de zile, din 2016, pentru ambele alinieri.
- **Raportați**
  - Un tabel (statistici × perioade × alinieri) și graficul de la (d).
- **Interpretare**
  - Ce statistică este cea mai afectată de alinierea greșită?
  - VaR-ul unui portofoliu Bitcoin–SPY doar cu poziții lungi, dintr-un model varianță–covarianță construit pe el, ar fi prea mic sau prea mare?
- **Model**: B4

◂ Înapoi ▸ Lista exercițiilor

C3: reproducerea convenției erorilor standard

- Date: cîmpurile close din fișierele BET.csv și STOXX50E.INDX.csv (STOXX50E: simbolul indicelui Euro Stoxx 50), din datele de piață ale cursului, inițial pentru 2002–2018
- Articolul folosește erori standard convenționale simple; reproduceți întîi această convenție
- Tendința: înmulțiți eroarea standard (*Standard Error*, SE) obișnuită a pantei OLS cu $100/\bar{m}$ ($\bar{m}$: media măsurii anuale m_t din C3), tratînd numitorul ca fix
- Variațiile logaritmice și relative: abaterea standard de selecție a variațiilor anuale, împărțită la rădăcina numărului lor
- Păstrați numai variațiile între ani calendaristici consecutivi observați
- Aceste SE ignoră dependența indusă de diferențiere și incertitudinea numitorului tendinței
  - sînt convenția articolului, nu o afirmație de validitate universală a inferenței
- Analiză de sensibilitate separată: erori HAC sau bootstrap pe blocuri; identificați deciziile care se schimbă
- Păstrați acest tabel separat de cel care reproduce convenția articolului.

C3 [Propus]: Erori nestandard: mini-multivers — Cerință (1/2)

Rezolvarea se discută la seminar

- **Întrebare**: Cît de mult modifică deciziile metodologice, singure, tendința estimată a eficienței pieței?
- **Date și ipoteze**
  - Context: în Menkveld et al. (2024) 164 de echipe au testat aceleași ipoteze pe aceleași date; intervalul intercuartilic (IQR) al estimărilor tendinței anuale a eficienței pieței (eroarea nestandard) a fost de 6,7 puncte procentuale, față de un SE median de 2,5% (Tabelul I, Panoul C).
  - Date: valorile zilnice de închidere ale indicelui Euro Stoxx 50 și ale BET (datele cursului), 2002–2018, apoi pînă la ultimul an complet; H_0: „eficiența pieței nu s-a schimbat în timp" (Sec. I.A).
- **Cerințe**
  - (a) Înainte de orice rulare, fixați în scris, printr-un commit datat, un plan cu opțiunile metodologice din Tabelul V disponibile pentru date zilnice: valorile extreme (nimic; m_t anual winsorizat sau cu extremele eliminate la percentilele 2,5 și 97,5, anii eliminați rămîn lipsă); măsura, în fiecare an (raportul variațelor, *variance ratio*, VR: $|VR(q) - 1|$, $VR(q) = \mathrm{Var}(\sum_{i=0}^{q-1} r_{t-i})/(q\mathrm{Var}(r_t))$, sume suprapuse cu corecțiile Lo–MacKinlay din Capitolul 2, $q = 5, 21$; R^2 al unui model autoregresiv AR(1) cu termen liber); modelul (tendință liniară; diferență logaritmică; diferență relativă).
  - (b) Pentru fiecare indice și fiecare dintre cele 27 de specificații (3 măsuri $\times$ 3 reguli pentru valorile extreme $\times$ 3 modele), calculați măsura anuală m_t, variația ei anuală medie în % și eroarea standard a estimării: $100\, b/\bar{m}$ din regresia măsurii pe o tendință liniară, media lui $100\,\Delta \ln m_t$ sau media lui $100\,(m_t/m_{t-1} - 1)$. Ipoteze nule: pentru m_t staționar fără tendință, panta și variația logaritmică medie sînt centrate în zero, variația relativă medie nu.
  - (c) Raportați mediana, IQR-ul (eroarea nestandard a multiversului dumneavoastră) și intervalul dintre percentilele 10 și 90 ale celor 27 de estimări; comparați IQR-ul cu 1,35$\times$ mediana erorilor standard.

◂ Înapoi ▸ Lista exercițiilor

C3 [Propus]: Erori nestandard: mini-multivers — Cerință (2/2)

Rezolvarea se discută la seminar

- **Cerințe (continuare)**
  - (a, continuare) Planul fixează și: regula pentru m_t nul sau lipsă; familia Holm (cele 18 specificații cu tendință liniară sau diferență logaritmică, pentru fiecare indice și perioadă); ultimul an complet disponibil la data-limită.
  - (d) Ierarhizați deciziile metodologice folosind statistica Anderson–Darling pentru k eșantioane (`scipy.stats.anderson_ksamp`) a estimărilor grupate după opțiunea de la fiecare decizie (Sec. II.C).
  - (e) Testați la 5% dacă tendința este nulă pentru cele 18 specificații centrate în zero sub H_0, cu corecția Holm pentru aceste 18 teste și erori standard care țin cont de dependență (HAC, bootstrap pe blocuri); raportați separat respingerile asociate tendințelor pozitive și negative (tendință pozitivă = eficiență în scădere).
  - (e, continuare) Cele 9 variații relative medii nu sînt centrate în zero sub H_0: le raportați descriptiv, fără test față de zero.
- **Raportați**
  - Un tabel cu cele 27 de estimări și erorile lor standard pe indice; mediana, IQR-ul și intervalul 10–90; ierarhia Anderson–Darling a deciziilor; respingerile Holm după semnul tendinței; un boxplot al estimărilor grupate după decizia cea mai influentă.
  - Model: A2 (Holm și BH pe o familie de teste); studiul de caz din curs
- **Interpretare**
  - Estimează IQR-ul celor 27 de specificații eroarea nestandard observată între cele 164 de echipe? Folosiți în răspuns nota 32 a articolului.
  - Dacă semnul tendinței depinde de modelul ales, ce raportați?
  - Ce regulă de alegere a modelului ați fi fixat în plan?

Bibliografie (1/3)

- Andrews, D.W.K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61(4), 821–856.
- Andrews, D.W.K. (1999). Estimation when a parameter is on a boundary. *Econometrica*, 67(6), 1341–1383.
- Athreya, K.B. (1987). Bootstrap of the mean in the infinite variance case. *Annals of Statistics*, 15(2), 724–731.
- Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B*, 57(1), 289–300.
- Benjamini, Y., Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 29(4), 1165–1188.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
- Bollerslev, T., Wooldridge, J.M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews*, 11(2), 143–172.
- Christoffersen, P.F. (1998). Evaluating interval forecasts. *International Economic Review*, 39(4), 841–862.
- Davis, R.A., Mikosch, T. (1998). The sample autocorrelations of heavy-tailed processes with applications to ARCH. *Annals of Statistics*, 26(5), 2049–2080.
- 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.

Bibliografie (2/3)

- Franke, J., Härdle, W.K., Hafner, C.M. (2019). Statistics of Financial Markets: An Introduction, 5th ed. Springer.
- Hansen, P.R., Lunde, A., Nason, J.M. (2011). The model confidence set. *Econometrica*, 79(2), 453–497.
- Harvey, C.R., Liu, Y., Zhu, H. (2016). . . . and the cross-section of expected returns. *Review of Financial Studies*, 29(1), 5–68.
- Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, 3(5), 1163–1174.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70.
- Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *Annals of Statistics*, 17(3), 1217–1241.
- Kupiec, P.H. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, 3(2), 73–84.
- Lobato, I., Nankervis, J.C., Savin, N.E. (2001). Testing for autocorrelation using a modified Box–Pierce Q test. *International Economic Review*, 42(1), 187–205.
- 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.
- Menkveld, A.J., Dreber, A., Holzmeister, F., Huber, J., Johannesson, M., Kirchler, M., et al. (2024). Nonstandard errors. *Journal of Finance*, 79(3), 2339–2390.

Bibliografie (3/3)

- Nelson, D.B. (1990). Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory*, 6(3), 318–334.
- Nolde, N., Ziegel, J.F. (2017). Elicitability and backtesting: perspectives for banking regulation. *Annals of Applied Statistics*, 11(4), 1833–1874.
- 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.
- Pele, D.T., Wesselhöfft, N., Härdle, W.K., Kolossiatis, M., Yatracos, Y.G. (2023). Are cryptos becoming alternative assets?. *European Journal of Finance*, 29(10), 1064–1105.