

# Modelarea piețelor financiare

## Seminarul 1: Fapte stilizate și randamente — aplicații



Daniel Traian PELE
Antoaneta Amza

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

## Ce putem afirma despre randamente?

- Întrebarea seminarului
  - **care dintre faptele stilizate ale randamentelor rămîn valide atunci cînd inferența ține cont de cozile groase și de gruparea volatilității (*volatility clustering*)?**
- Trei tipuri de exerciții
  - **A**: exerciții de calcul, pe hîrtie
  - **B**: exerciții pe date reale, rezolvate în notebook
  - **C**: întrebări deschise, de exemplu corectarea unui răspuns dat de un instrument de inteligență artificială (AI)
- Fiecare exercițiu este marcat
  - **[Rezolvat]**: rezolvarea completă este prezentată la seminar
  - **[Propus]**: îl rezolvați dumneavoastră; soluția se discută la seminar

## Ce exerciții rezolvăm mai întîi?

- Exercițiile de azi, în ordine
  - A1: agregarea randamentelor în timp și anualizarea
  - A2(a): amplitudinea unui drawdown (scăderea față de un vîrf anterior)
  - B1: statistici descriptive și grosimea cozilor pentru trei piețe
  - A3: testele Jarque–Bera și Ljung–Box, pas cu pas
  - B2: cozi groase și valoarea la risc (VaR) de 1% pe BET-TR (*BET Total Return*, indicele BET cu dividendele reinvestite), indicele BET (*Bucharest Exchange Trading*, indicele de referință al Bursei de Valori București) al Bursei de Valori București cu dividendele reinvestite
  - A4: kurtosisul unei mixturi de distribuții Normale
  - B3: randamente necorelate, dar nu independente (Bitcoin)
  - B4: testul Ljung–Box robust la gruparea volatilității
  - B5: agregarea randamentelor și apropierea de distribuția Normală
  - C1: corectarea unui răspuns AI cu erori
- După cursul 1 (estimarea cozilor și benzi de autocorelație robuste)
  - A5 și B6: estimatorul Hill pentru indicele cozii (*tail index*)
  - A6: lărgirea benzii de autocorelație în prezența grupării volatilității

## Ce păstrăm din fiecare analiză?

- Livrabile
  - răspunsurile la exercițiile A
  - cîte o figură și un tabel pentru fiecare exercițiu B
  - codul corectat din C1
  - o frază de interpretare pentru fiecare exercițiu
- Notebook-ul de azi: capitol1_notebook_seminar.ipynb (Colab); prețuri: data/market
  - fiecare cerință indică secțiunea din notebook
- Seminar: prof. Daniel Traian Pele, danpele@ase.ro; drd. Antoaneta Amza, antoaneta.amza@csie.ase.ro

## Cum citim seriile și indicatorii?

- **JB**: testul de normalitate Jarque–Bera.
- **ACF**: funcția de autocorelație (*AutoCorrelation Function*): corelațiile unei serii cu valorile ei decalate în timp.
- **BVB**: Bursa de Valori București.
- **UTC**: timpul universal coordonat (*Coordinated Universal Time*).
- **BNR**: Banca Națională a României.
- **S&P 500**: Standard & Poor's 500 — indice al acțiunilor a 500 de companii mari din Statele Unite.
- **OHLC**: prețurile de deschidere, maxim, minim și închidere (*Open, High, Low, Close*).
- **SUA**: Statele Unite ale Americii.
- **SPY**: simbolul fondului care urmărește indicele S&P 500.
- **VIX**: indicele volatilității implicite a S&P 500 (*Cboe Volatility Index*).

# Harta exercițiilor

| Exercițiu | Întrebare | Tip | Model |
|---|---|---|---|
| A1 | cum se agregă randamentele în timp? | Rezolvat | – |
| A2(a) | cît de mare este un drawdown? | Propus | A1 |
| B1 | cît de groase sînt cozile și cît de robust este kurtosisul? | Rezolvat | – |
| A3 | de la statistică la valoarea p (*p-value*): JB și Ljung–Box | Rezolvat | – |
| B2 | cît de mare este VaR 1% pentru BET-TR? | Propus | B1 |
| A4 | de ce generează volatilitatea aleatoare cozi groase? | Rezolvat | – |
| B3 | sînt randamentele Bitcoin independente? | Rezolvat | – |
| B4 | ce autocorelații rămîn semnificative la inferența robustă? | Rezolvat | – |
| B5 | se apropie randamentele de distribuția Normală cînd orizontul crește? | Propus | B1, B3 |
| C1 | ce este greșit într-un răspuns AI? | Propus | A3, B3, B4 |
| După curs | A5 [Rezolvat], B6 [Rezolvat], A6 [Propus] (grosimea cozii, lărgimea benzii ACF) | recapitulare; rezolvați A6 | A4, B4 |
| Pachete de extindere | A1(c), A2(b–f), A7–A11, B1(d), B7–B15, C2, C3 (opționale) | Rezolvat sau Propus | indicat la fiecare cerință |

# Datele folosite

| Serie | Tip de preț | Perioadă | Calendar | Randamente |
|---|---|---|---|---|
| BET-TR | închidere zilnică | 23 sept. 2014 – 18 sept. 2026 | zilele de tranzacționare BVB | 2 992 |
| Bitcoin | închidere zilnică (USD) | 17 sept. 2014 – 18 sept. 2026 | în fiecare zi, 24:00 UTC | 4 384 |
| EUR/RON | cursul de referință BNR | 4 iul. 2005 – 18 sept. 2026 | zilele lucrătoare din România | 5 352 |
| S&P 500 | OHLC zilnic | 2 ian. 1990 – 18 sept. 2026 | zilele de tranzacționare din SUA | 9 245 |
| ETF-ul SPY | OHLC, închidere ajustată (pentru dividende și splituri) | 2 ian. 2008 – 18 sept. 2026 (B15) | zilele de tranzacționare din SUA | – |
| VIX, VIX pe 9 zile | închidere zilnică | doar C3 | zilele de tranzacționare din SUA | – |
| Aur | spot, USD pe uncie | 2 ian. 1990 – 18 sept. 2026 | zilele lucrătoare | 9 559 |

- Prețurile zilnice din datele cursului, cîte o serie pe activ (cîmpul folosit: close / adjusted close)
- EUR/RON, numărul de lei (RON) pentru un euro (EUR), este cursul oficial de referință BNR, nu o serie de cotații de piață
- ETF (Exchange-Traded Fund): fond tranzacționat la bursă; randamentele sînt în % pe slide-uri
- Calendare: fiecare serie pe propriul calendar
  - analizele pe mai multe serii aliniază întîi *prețurile* pe zilele comune

## Fișierele pentru SPY și indicii VIX

- Fișiere CSV (*Comma-Separated Values*, fișier text cu valori separate prin virgulă) furnizate în `data/market`; toate se încheie la 18 sept. 2026.
- Numărul de mai jos reprezintă observații de preț, înainte de diferențiere.

| Fișier | Prima dată | Rînduri | Exercițiu |
|---|---|---|---|
| SPY.US.csv | 2008-01-02 | 4 708 | B15, C3 |
| VIX.INDX.csv | 1990-01-02 | 9 284 | C3 |
| VIX9D.INDX.csv | 2011-01-03 | 3 951 | C3 |

- Cîmpuri: `open, high, low, close, adjusted_close, volume`, indexate prin `date`.
- Încărcați cu `read_market("SPY.US")` din `mfm_data`; pentru SPY selectați datele din 2008. Pentru VIX folosiți `"VIX.INDX"`, respectiv `"VIX9D.INDX"`.

## Acțiunile BVB pentru C2

- Fișiere furnizate: `data/market/<ticker>.RO.csv`; ultima dată: 18 sept. 2026.
- C2 folosește `volume` și `close` brut, în lei. Numerele sînt rînduri de preț, înainte de filtrare.

| Simbol | Prima dată | Rînduri | Simbol | Prima dată | Rînduri |
|---|---|---|---|---|---|
| TLV | 1997-10-16 | 7 379 | SNG | 2013-11-12 | 3 284 |
| SNP | 2001-09-04 | 6 464 | H2O | 2023-07-12 | 793 |
| BRD | 2001-01-16 | 6 628 | SNN | 2013-11-04 | 3 290 |
| TGN | 2008-01-25 | 4 796 | EL | 2014-07-07 | 3 071 |
| FP | 2010-09-30 | 4 087 | DIGI | 2017-05-17 | 2 368 |
| TEL | 2006-08-29 | 5 164 | | | |

# Pregătire: rulați și verificați

```
# sectiunea "Date" din notebook defineste load_data(name)
closes = {a: load_data(a)['Close'] for a in
          ['sp500', 'bettr', 'btc', 'eurron']}
rets = {a: np.log(c).diff().dropna()
        for a, c in closes.items()}
r_bet, r_btc = rets['bettr'], rets['btc']
r_eur, r_sp  = rets['eurron'], rets['sp500']
for a, r in rets.items():
    print(a, r.index[0].date(), len(r))
```

- Rezultat așteptat: data primului randament și numărul de randamente

| Serie | primul randament | $T$ |
|---|---|---|
| S&P 500 | 1990-01-03 | 9 245 |
| BET-TR | 2014-09-24 | 2 992 |
| Bitcoin | 2014-09-18 | 4 384 |
| EUR/RON | 2005-07-05 | 5 352 |

- Convenții în cod
  - momente: `scipy.stats.skew` și `kurtosis` (exces de kurtosis, fără corecție de eșantion mic)
  - abateri standard cu `ddof=1`
  - cuantile cu `numpy.quantile` (liniar)
- Dacă un număr diferă, opriți-vă și verificați seria, tipul de preț și data de început

## Prețuri și randamente

- Indicii și prețurile
  - $t = 0, 1, \ldots, T$: indicele zilei de tranzacționare; $t - 1$ este ziua precedentă
  - $P_t$: prețul de închidere din ziua $t$; $T$: numărul de randamente
- **Randamentul simplu**: variația relativă a prețului într-o zi

$$R_t = \frac{P_t}{P_{t-1}} - 1$$

- **Randamentul logaritmic**: logaritmul natural (ln) al raportului dintre prețuri

$$r_t = \ln \frac{P_t}{P_{t-1}} = \ln(1 + R_t)$$

  - $r_t \approx R_t$ pentru variații mici; $r_t < R_t$ ori de cîte ori $R_t \neq 0$
  - un exemplu: $P_0 = 100$, $P_1 = 102$: $R_1 = 2{,}00\%$, $r_1 = \ln 1{,}02 = 1{,}98\%$

## Media, abaterea standard, standardizarea

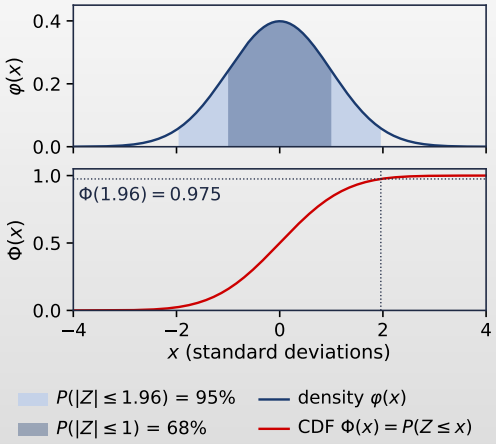
- Convenții de scriere
  - notația ˆ: $\hat{\theta}$ este valoarea estimată din eșantion a unei mărimi teoretice $\theta$
  - $\sum_t$: suma după toate zilele $t = 1, \ldots, T$
- **Media de selecție** $\bar{r} = \frac{1}{T} \sum_t r_t$: randamentul mediu zilnic
- **Abaterea standard de selecție** (volatilitatea zilnică): dispersia tipică a randamentelor în jurul mediei

$$\hat{\sigma} = \sqrt{\frac{1}{T-1} \sum_t (r_t - \bar{r})^2}$$

- Randamentul centrat $x_t = r_t - \bar{r}$: abaterea față de medie
- Randamentul standardizat $z_t = (r_t - \bar{r})/\hat{\sigma}$: numărul de abateri standard față de medie
  - $z_t = -3$: o zi cu trei abateri standard sub medie
- **i.i.d.**: independente și identic distribuite, ipoteza de referință a testelor clasice
  - randamentul de azi nu depinde de cele trecute, iar distribuția lui este aceeași în fiecare zi

# Curba distribuției Normale

- $Z \sim N(0,1)$
  - $\varphi(x)$: înălțimea curbei în $x$
  - $\Phi(x)$: aria de sub curbă la stînga lui $x$
- Ariile colorate din grafic
  - $P(|Z| \leq 1) = 68\%$
  - $P(|Z| \leq 1{,}96) = 95\%$
- $\Phi^{-1}$: inversa lui $\Phi$, adică cuantilele lui $N(0,1)$
  - $\Phi^{-1}(0{,}975) = 1{,}96$, deci $Z$ ia valori în $[-1{,}96;\ 1{,}96]$ cu probabilitatea 95%
- Standardizarea: dacă $X \sim N(\mu, \sigma^2)$, atunci $Z = (X - \mu)/\sigma \sim N(0,1)$



## Agregarea în timp și anualizarea

- **Agregarea în timp**: randamentul pe zilele $1, \dots, n$, de la $P_0$ la $P_n$

$$1 + R_{0\to n} = \frac{P_n}{P_0} = \prod_{t=1}^{n}(1 + R_t), \qquad r_{0\to n} = \ln \frac{P_n}{P_0} = \sum_{t=1}^{n} r_t$$

  - $\prod$: produsul factorilor; randamentele simple se compun (se înmulțesc factorii $1 + R_t$)
  - randamentele logaritmice se adună: $\sum_{t=1}^{n}(\ln P_t - \ln P_{t-1}) = \ln P_n - \ln P_0$, deoarece termenii intermediari se anulează ($+\ln P_1$ cu $-\ln P_1$ etc.)
  - între active, într-un portofoliu, se adună ponderat randamentele simple (A1(c))

- **Anualizarea**: $A$ este numărul de intervale de randament dintr-un an

$$\bar{r}_{\text{an}} = A\,\bar{r}, \qquad \hat{\sigma}_{\text{an}} = \sqrt{A}\,\hat{\sigma}, \qquad \text{deoarece } \operatorname{Var}\Big(\sum_{t=1}^{A} r_t\Big) = A\sigma^2$$

  - $A = 252$ (BET-TR, S&P 500, EUR/RON: zilele de tranzacționare), $A = 365$ (Bitcoin, tranzacționat zilnic)
  - $A\sigma^2$ presupune randamente necorelate, cu variață constantă (A7: cazul autocorelat)
  - exemplu: $\hat{\sigma} = 1\%$ pe zi și $A = 252$: $\hat{\sigma}_{\text{an}} = 15{,}87 \times 1\% = 15{,}9\%$

## Asimetria, kurtosisul, autocorelația

- **Asimetria** $\hat{S} = \frac{1}{T}\sum_t z_t^3$: media cuburilor randamentelor standardizate
  - $\hat{S} = 0$ pentru o distribuție simetrică
  - $\hat{S} < 0$: pierderile extreme sînt mai mari, în valoare absolută, decît cîștigurile extreme
- **Kurtosisul** $\hat{K} = \frac{1}{T}\sum_t z_t^4$: media puterilor a patra, determinată mai ales de zilele extreme
  - $\hat{K} = 3$ pentru distribuția Normală
  - **excesul de kurtosis** $\hat{K} - 3$: 0 pentru distribuția Normală, pozitiv pentru cozi groase
- **Autocorelația** la lagul $k$: corelația dintre randamentele aflate la $k$ zile distanță

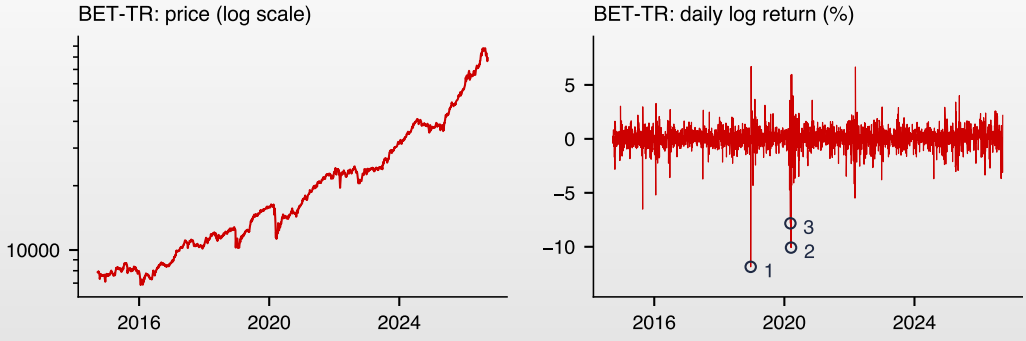
$$\hat{\rho}_k = \frac{\sum_{t=k+1}^{T} x_t x_{t-k}}{\sum_{t=1}^{T} x_t^2}$$

  - $k = 1, 2, \ldots$: lagul, în zile; $\hat{\rho}_k \in [-1, 1]$; $\rho_k$ fără $\hat{}$: valoarea teoretică
  - $\hat{\rho}_k > 0$: o zi peste medie tinde să fie urmată, după $k$ zile, de o zi peste medie
  - $\hat{\rho}_k \approx 0$: nicio legătură liniară între randamentele aflate la $k$ zile distanță

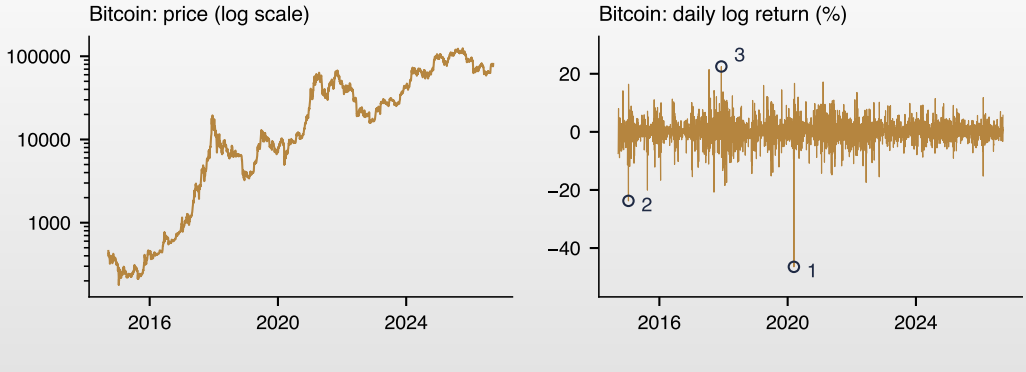
## Testele statistice

- **Distribuția** $\chi^2_m$ (hi-pătrat cu $m$ grade de libertate)
  - distribuția sumei pătratelor a $m$ variabile $N(0,1)$ independente; media ei este $m$
  - valori critice la 5%: 3,84 ($m = 1$), 5,99 ($m = 2$), 11,07 ($m = 5$), 18,3 ($m = 10$)
- **Valoarea critică** la 5%: valoarea depășită cu probabilitatea 5% sub $H_0$
  - respingem $H_0$ la pragul de 5% dacă statistica depășește valoarea critică
- **Valoarea p**: probabilitatea, sub $H_0$, a unei statistici cel puțin la fel de extreme ca cea observată
  - $p < 0{,}05$ este echivalent cu depășirea valorii critice de 5%
  - nu este probabilitatea ca $H_0$ să fie adevărată

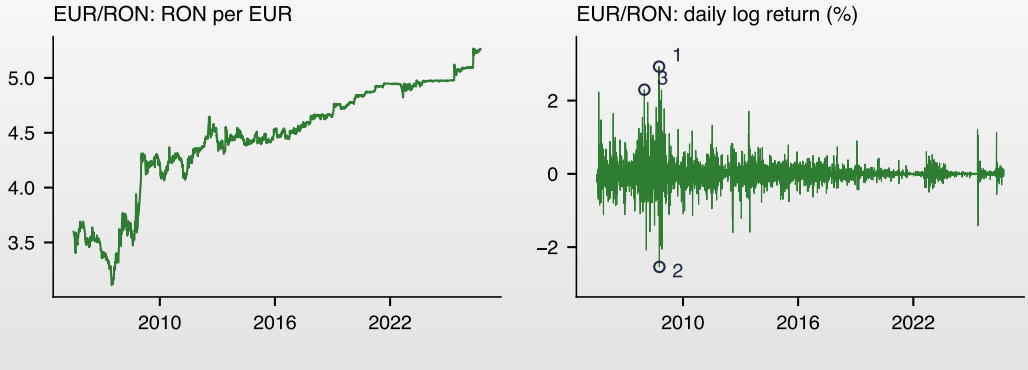
# BET-TR: prețuri, randamente și zile extreme



# Bitcoin: prețuri, randamente și zile extreme



# EUR/RON: prețuri, randamente și zile extreme

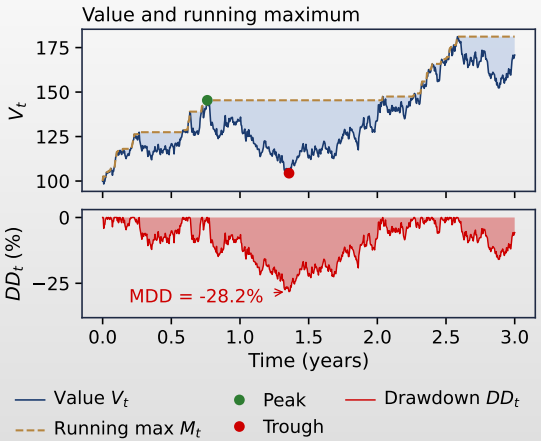


# Fapte stilizate și episoade de piață

- **Fapt stilizat**: o proprietate statistică comună multor active, piețe și perioade Cont (2001)
- Cele patru fapte stilizate de azi
  - **cozi groase**: zilele extreme sînt mai frecvente decît sub distribuția Normală cu aceeași varianță
  - **gruparea volatilității**: variațiile mari sînt urmate de variații mari, de orice semn
  - **lipsa autocorelației liniare**: randamentul de azi nu este o funcție liniară a randamentelor trecute, $\hat{\rho}_k \approx 0$
  - **gaussianitate prin agregare**: randamentele pe orizonturi mai lungi au cozi mai apropiate de cele ale distribuției Normale
- Fiecare zi extremă corespunde unui eveniment identificabil
  - taxa pe activele bancare (19 dec. 2018)
  - crahul din martie 2020, la începutul pandemiei de COVID-19 (boala provocată de coronavirus)
  - deprecierea leului din 2008
- Perioadele calme alternează cu perioade turbulente în toate cele trei serii
- EUR/RON este aproape constant după 2013
- Analiza grafică precede testarea formală

MFM_ch1_seminar_tests

# Drawdown-ul




- **Vîrful curent**: cea mai mare valoare atinsă pînă la momentul $t$

$$M_t = \max_{s \le t} V_s$$

  - $V_t$: valoarea portofoliului sau prețul la momentul $t$
- **Drawdown-ul**: pierderea relativă față de vîrful curent

$$DD_t = \frac{V_t}{M_t} - 1 \le 0$$

  - $DD_t = 0$ la fiecare nou maxim; graficul lui $DD_t$ este curba „sub apă” (*underwater*)
- **Drawdown-ul maxim** (MDD, Maximum DrawDown) $MDD = \min_t DD_t$
  - cea mai mare scădere de la un vîrf la un minim ulterior
  - în grafic: $-28{,}2\%$, de la punctul verde la punctul roșu
- **CAGR** (Compound Annual Growth Rate) $= (V_n/V_0)^{1/Y} - 1$: rata anuală constantă de creștere
  - $Y$: numărul de ani dintre $V_0$ și $V_n$

## Rapoartele Sharpe, Sortino și Calmar

- **Raportul Sharpe**: randamentul anual în exces pe unitatea de volatilitate anuală

$$SR = \frac{A\,\bar{R} - r_f}{\sqrt{A}\,\hat{\sigma}_R}$$

  - $\bar{R}$, $\hat{\sigma}_R$: media și abaterea standard a randamentelor simple; $r_f$: rata anuală a dobînzii fără risc
  - $SR = 0{,}5$: fiecare punct procentual de volatilitate anuală aduce 0,5 puncte procentuale de randament în exces
- **Raportul Sortino** Sortino & van der Meer (1991): penalizează doar zilele cu pierderi

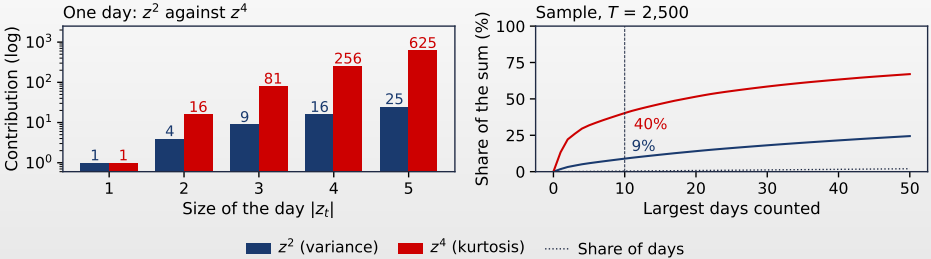
$$\text{Sortino} = \frac{A\,\bar{R} - r_f}{\sqrt{A}\,\hat{\sigma}_D}, \qquad \hat{\sigma}_D = \sqrt{\frac{1}{T}\sum_{t=1}^{T} \min(R_t, 0)^2}$$

  - $\hat{\sigma}_D$: semiabaterea inferioară; zilele cu $R_t \geq 0$ contribuie cu 0
  - Sortino $>$ Sharpe nu dovedește asimetrie: pentru un randament simetric cu media aproape zero, semiabaterea inferioară este aproximativ $\sigma/\sqrt{2}$, deci Sortino este deja circa 1,4 ori Sharpe
- **Raportul Calmar** $= \text{CAGR}/|MDD|$ (slide-ul anterior)
  - creșterea anuală pe unitatea de pierdere maximă de la un vîrf
  - depinde de un singur episod, deci este o măsură descriptivă, fără o eroare standard simplă

## Momentele de ordinul $k$ și kurtosisul

- **Momentele** unei variabile aleatoare $r$, cu media $\mu$ și abaterea standard $\sigma$
  - momentul de ordinul $k$: $\mathbb{E}[r^k]$; momentul centrat de ordinul $k$: $\mathbb{E}[(r-\mu)^k]$
  - momentul standardizat de ordinul $k$: $\mathbb{E}[z^k]$, cu $z = (r-\mu)/\sigma$
  - $\mu = \mathbb{E}[r]$ este momentul de ordinul 1; varianța $\sigma^2 = \mathbb{E}[(r-\mu)^2]$ este momentul centrat de ordinul 2
  - asimetria $S = \mathbb{E}[z^3]$ și kurtosisul $K = \mathbb{E}[z^4]$ sînt momentele standardizate de ordinul 3 și 4
  - **moment de ordinul patru finit**: $\mathbb{E}[r^4] < \infty$; altfel $\hat{K}$ nu se stabilizează cînd $T$ crește
- **Kurtosisul de selecție** $\hat{K} = \frac{1}{T}\sum_t z_t^4$ și excesul $\hat{K} - 3$
- De ce momentul de ordinul patru măsoară cozile
  - contribuția unei zile crește foarte repede cu $|z_t|$: $1^4 = 1$, $2^4 = 16$, $3^4 = 81$, $10^4 = 10\,000$
  - deci $\hat{K}$ este determinat de zilele extreme, nu de zilele obișnuite
  - o zi cu $z_t = 10$ adaugă singură $10^4/T$ la $\hat{K}$: pentru $T = 2\,500$, adică $+4$

# Contribuția zilelor extreme la kurtosis




- Stînga: contribuția unei singure zile la $\sum_t z_t^2$ (varianța) și la $\sum_t z_t^4$ (kurtosisul), pe scală logaritmică
  - o zi cu $|z_t| = 5$ contează de 25 de ori mai mult decît o zi cu $|z_t| = 1$ în varianță, dar de 625 de ori mai mult în kurtosis
- Dreapta: un eșantion simulat cu cozi groase, $T = 2\,500$; ponderea celor mai mari zile în cele două sume
  - cele mai mari 10 zile (0,4% din zile) dau 9% din $\sum_t z_t^2$, dar 40% din $\sum_t z_t^4$
  - de aceea $\hat{K}$ se schimbă mult cînd se elimină cîteva zile (B1)
- **Ce credeți**: ce statistică se schimbă mai mult dacă eliminăm ziua cea mai extremă a eșantionului, abaterea standard sau kurtosisul?
  - **Răspuns**: kurtosisul, pentru că ziua contribuie cu $z_t^4$, nu cu $z_t^2$

## Cuantilele și o măsură robustă a cozilor

- **Cuantila** de ordin $p$, notată $q_p$: valoarea sub care se află fracțiunea $p$ a randamentelor

$$q_p = F^{-1}(p), \qquad F(x) = P(r_t \leq x)$$

  - $F$: funcția de repartiție; empiric, $q_p$ este valoarea de pe poziția $pT$ în randamentele sortate crescător
  - $q_{0,5}$ este mediana; $q_{0,25}$ și $q_{0,75}$ sînt cuartilele
- Alternativa robustă Kim & White (2004): excesul de kurtosis Crow–Siddiqui pe cuantile, notat $K_{CS}$

$$K_{CS} = \frac{q_{0,975} - q_{0,025}}{q_{0,75} - q_{0,25}} - 2{,}91$$

  - numărătorul: lățimea intervalului care conține 95% din zile; numitorul: lățimea celor 50% din mijloc
  - 2,91 este valoarea raportului pentru distribuția Normală, deci măsura este 0 pentru distribuția Normală
  - valoare pozitivă: cozile sînt mai lungi, față de partea centrală, decît la distribuția Normală
  - depinde doar de cuantile, deci cîteva zile extreme o modifică foarte puțin

## Testul Jarque–Bera

- **Jarque–Bera** Jarque & Bera (1980): măsoară cît de departe sînt asimetria și excesul de kurtosis de valorile distribuției Normale, ambele egale cu 0

$$JB = T\Big[\frac{\hat{S}^2}{6} + \frac{(\hat{K}-3)^2}{24}\Big]$$

  - 6 și 24: variațele lui $\sqrt{T}\,\hat{S}$ și $\sqrt{T}(\hat{K}-3)$ sub $H_0$, pentru $T$ mare
  - fiecare termen este, așadar, pătratul unei variabile aproximativ $N(0,1)$
  - $JB \geq 0$, iar $JB = 0$ doar dacă $\hat{S} = 0$ și $\hat{K} = 3$
- $H_0$: randamente i.i.d. cu distribuția Normală
  - sub $H_0$, $JB \sim \chi^2_2$ pentru $T$ mare (doi termeni); valoarea critică la 5% este 5,99
  - un JB mare: asimetria sau cozile sînt incompatibile cu distribuția Normală
- Exemplu: $T = 1\,000$, $\hat{S} = 0$, $\hat{K} = 6$: $JB = 1\,000 \times 9/24 = 375 > 5{,}99$, respingem $H_0$

## Value-at-Risk și distribuția Normală

- **Valoarea la risc (VaR) de 1%** pe o zi: pragul de pierdere depășit cu probabilitatea de 1%

$$P\left(L_t > \mathrm{VaR}_\alpha\right) = \alpha, \qquad L_t = -r_t, \qquad \mathrm{VaR}_\alpha = -q_\alpha$$

  - $L_t$: pierderea zilei $t$; $\alpha$: nivelul VaR, aici $1\% = 0{,}01$
  - $q_\alpha$: cuantila de ordin $\alpha$ a randamentului; semnul minus transformă cuantila, negativă, într-o pierdere pozitivă
  - VaR nu este pierderea maximă: în aproximativ 1% din zile pierderea îl depășește, iar VaR nu spune cît

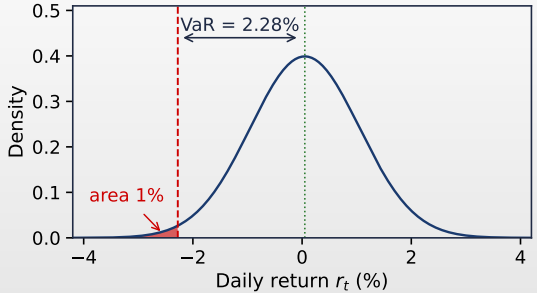
- Dacă randamentul are distribuția Normală $N(\mu, \sigma^2)$

$$q_\alpha = \mu + \sigma\, \Phi^{-1}(\alpha)$$

  - $\Phi^{-1}(0{,}01) = -2{,}326$: 1% din valorile unei variabile $N(0,1)$ sînt sub $-2{,}326$
  - exemplu: $\mu = 0{,}05\%$, $\sigma = 1\%$: $q_{0{,}01} = 0{,}05 - 2{,}326 = -2{,}28\%$, deci VaR $1\% = 2{,}28\%$

- **VaR empiric**: minus cuantila de selecție de 1%
  - pe eșantionul pe care a fost estimat este depășit, prin construcție, în 1% din zile

# Cuantila și VaR pe o densitate




- Density of $r_t \sim N(\mu, \sigma^2)$
- Left tail, probability $\alpha$ = 1%
- Quantile $q_{0.01}$ = -2.28%
- Mean $\mu$ = 0.05%

- **Densitatea** $f$: aria de sub curbă între $a$ și $b$ este $P(a < r_t \leq b)$; aria totală este 1
- **Cuantila** $q_{0,01}$: punctul care lasă la stînga o arie de 1% (zona roșie)

$$\int_{-\infty}^{q_{0,01}} f(x)\,dx = 0{,}01$$

- **VaR 1%** $= -q_{0,01}$: distanța de la 0 la cuantilă, exprimată ca pierdere pozitivă
  - cu cozi mai groase, aria de 1% se află mai departe la stînga, iar VaR crește (B2)

## Distribuția Student-$t$

- **Student-$t$** cu $\nu > 0$ grade de libertate: o distribuție simetrică, cu cozi mai groase decît distribuția Normală
  - $P(|X| > x)$ descrește proporțional cu $x^{-\nu}$ cînd $x$ crește: cu cît $\nu$ este mai mic, cu atît cozile sînt mai groase
  - pentru $\nu \to \infty$ se obține distribuția Normală
- Momentele: $\mathbb{E}|X|^p$ este finit doar pentru $p < \nu$
  - varianța există pentru $\nu > 2$ și este egală cu $\nu/(\nu - 2)$; kurtosisul există pentru $\nu > 4$
- Cuantilele: $t_\nu^{-1}(\alpha)$ este cuantila de ordin $\alpha$ a distribuției Student-$t$ standard
  - $t_3^{-1}(0{,}01) = -4{,}54$
  - standardizarea la varianță unitară: se împarte la abaterea standard a distribuției, $\sqrt{\nu/(\nu-2)}$
  - pentru $\nu = 3$: $\sqrt{3/(3-2)} = \sqrt{3}$, deci cuantila standardizată este $-4{,}54/\sqrt{3} = -2{,}62$
  - ea este mai extremă decît $-2{,}33$, cuantila de 1% a distribuției Normale
- Estimarea: prin verosimilitate maximă, pentru variabila $m + s\,Y$, cu $Y$ Student-$t$ standard
  - $m$: parametrul de locație (centrul); $s > 0$: parametrul de scală (lățimea); `scipy.stats.t.fit`

## Densitatea Student-$t$ și indicele cozii

- **Densitatea Student-$t$ standard** cu $\nu$ grade de libertate

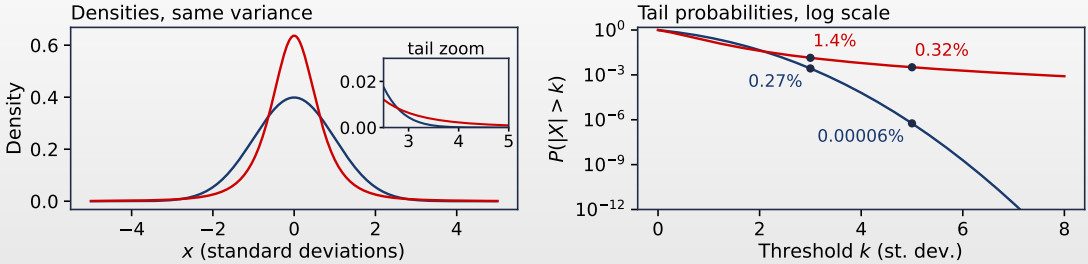
$$f_\nu(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

  - $\Gamma$: funcția gamma, $\Gamma(n) = (n-1)!$ pentru $n$ întreg; fracția din față face ca aria totală să fie 1
  - pentru $|x|$ mare, $f_\nu(x)$ descrește ca $|x|^{-(\nu+1)}$, o putere a lui $|x|$
  - densitatea Normală descrește ca $e^{-x^2/2}$, mult mai repede: de aici diferența dintre cozi
- **Indicele cozii** $\alpha$: viteza cu care scade probabilitatea unei mișcări mari

$$P(|r_t| > x) \propto x^{-\alpha}, \quad x \text{ mare}$$

  - $\propto$: proporțional cu; un $\alpha$ mic înseamnă cozi groase; pentru Student-$t$, $\alpha = \nu$; alt $\alpha$ decît nivelul VaR
  - $\mathbb{E}|r_t|^p$ este finit doar pentru $p < \alpha$: cu $\alpha \leq 4$, kurtosisul nu există
- După curs
  - A5 și B6 estimează $\alpha$ cu estimatorul Hill; A6 derivă varianța lui $\hat{\rho}_k$

# Distribuția Normală și Student-$t$ cu aceeași varianță




- Stînga: $N(0,1)$ și Student-$t$ cu $\nu = 3$, rescalată la varianța 1 (împărțită la $\sqrt{3}$)
  - aceeași varianță, dar Student-$t$ are vîrful mai înalt, flancurile mai subțiri și cozile mai groase (zoom)
- Dreapta: $P(|X| > k)$ pe scală logaritmică
  - $k = 3$: 1,4% pentru Student-$t$, față de 0,27% pentru distribuția Normală, de 5 ori mai probabil
  - $k = 5$: 0,32% față de 0,00006%, de peste 5 000 de ori mai probabil

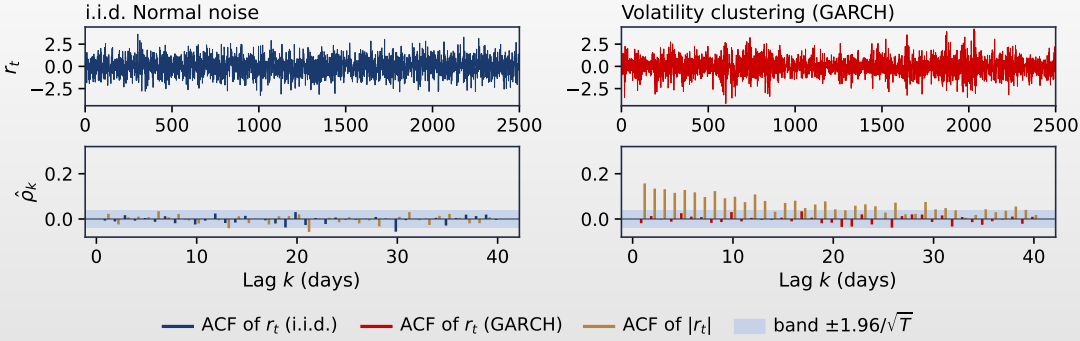
## ACF și testul Ljung–Box

- **ACF** (funcția de autocorelație): șirul $\hat{\rho}_k$, $k = 1, 2, \ldots$, reprezentat în funcție de lag
  - pentru randamente i.i.d. $\hat{\rho}_k \approx N(0, 1/T)$, deci abaterea standard a lui $\hat{\rho}_k$ este $1/\sqrt{T}$
  - banda de 95%: $\pm 1{,}96/\sqrt{T}$ ($T = 2\,500$: $\pm 0{,}039$); un $\hat{\rho}_k$ în afara ei este semnificativ la 5%
- **Gruparea volatilității** în ACF: $|r_t|$ și $r_t^2$ sînt autocorelate, deși $r_t$ nu este
- **Ljung–Box** Ljung & Box (1978): testează împreună primele $m$ autocorelații

$$Q(m) = T(T+2) \sum_{k=1}^{m} \frac{\hat{\rho}_k^2}{T-k}$$

  - $m$: numărul de laguri testate; $H_0$: $\rho_1 = \cdots = \rho_m = 0$
  - pentru randamente i.i.d., $Q(m) \sim \chi^2_m$; un $Q$ peste valoarea critică: cel puțin o autocorelație nenulă
  - aplicat pe $r_t$, testează predictibilitatea liniară; aplicat pe $r_t^2$, testează gruparea volatilității

# ACF pentru zgomot i.i.d. și pentru gruparea volatilității




- Seria din dreapta, simulată dintr-un model GARCH (heteroscedasticitate condiționată autoregresivă generalizată, *Generalised AutoRegressive Conditional Heteroskedasticity*): $r_t = \sigma_t \varepsilon_t$, $\sigma_t^2 = 0{,}05 + 0{,}10\, r_{t-1}^2 + 0{,}85\, \sigma_{t-1}^2$, $\varepsilon_t$ i.i.d. $N(0,1)$
  - o zi cu $|r_{t-1}|$ mare crește volatilitatea de azi: alternează perioade calme și turbulente
- ACF a lui $r_t$ rămîne în banda $\pm 1{,}96/\sqrt{T} = \pm 0{,}039$ ($T = 2\,500$) pentru ambele serii
- ACF a lui $|r_t|$: în bandă pentru zgomotul i.i.d., mult peste bandă pentru seria din dreapta (0,16 la lagul 1)

## Testul ARCH-LM

- **ARCH-LM** Engle (1982): testul multiplicatorului Lagrange (LM) pentru efecte ARCH (*AutoRegressive Conditional Heteroskedasticity*, heteroscedasticitate condiționată autoregresivă), adică dacă pătratul randamentului de azi depinde de pătratele trecute (gruparea volatilității)
- Regresia auxiliară, estimată prin metoda celor mai mici pătrate (OLS, *Ordinary Least Squares*)

$$x_t^2 = a_0 + a_1 x_{t-1}^2 + \cdots + a_q x_{t-q}^2 + u_t$$

  - $q$: numărul de laguri; $a_0, \ldots, a_q$: coeficienții; $u_t$: eroarea
  - $H_0$: $a_1 = \cdots = a_q = 0$, adică fără gruparea volatilității
- Statistica $LM = (T - q)R^2$, comparată cu $\chi_q^2$
  - $R^2 \in [0, 1]$: proporția din variația lui $x_t^2$ explicată de pătratele trecute
  - $T - q$: numărul de observații folosite în regresie
  - exemplu: $T = 1\,000$, $q = 5$, $R^2 = 0{,}02$: $LM = 995 \times 0{,}02 = 19{,}9 > 11{,}07$, valoarea critică la 5% a lui $\chi_5^2$: volatilitatea este grupată

## Mixtura de distribuții Normale

- **Mixtură de distribuții Normale** Clark (1973): randamentul este un șoc Normal scalat de o volatilitate aleatoare

$$r_t = \sigma_t \varepsilon_t$$

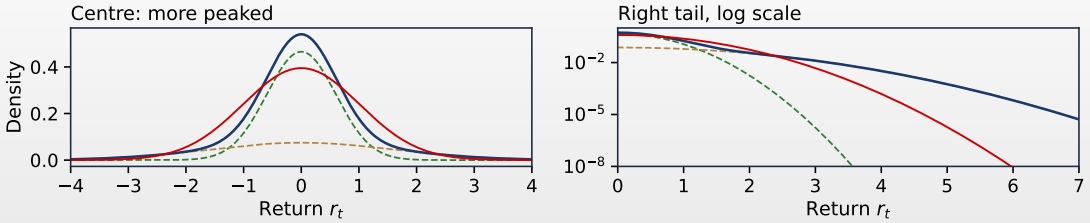
  - $\varepsilon_t$: șocul, i.i.d. $N(0,1)$ și independent de $\sigma_t$
  - $\sigma_t > 0$: volatilitatea zilei $t$, ea însăși aleatoare (de exemplu, după ritmul sosirii informației)

- Kurtosisul teoretic al lui $r_t$

$$K = \frac{\mathbb{E}r^4}{(\mathbb{E}r^2)^2} = \frac{3\,\mathbb{E}\sigma^4}{(\mathbb{E}\sigma^2)^2} \geq 3$$

  - 3 este $\mathbb{E}\varepsilon^4$, kurtosisul distribuției $N(0,1)$
  - $K = 3$ doar dacă $\sigma_t$ este constantă; cu cît $\sigma_t$ variază mai mult, cu atît $K$ este mai mare
  - cozi groase generate de șocuri cu distribuția Normală (A4)

# O mixtură de două distribuții Normale




- Zile calme, cu probabilitatea 0,7: $\sigma_t = 0{,}6$; zile turbulente, cu probabilitatea 0,3: $\sigma_t = 1{,}6$

$$f(x) = 0{,}7\,\varphi_{\mathbf{0},\mathbf{6}}(x) + 0{,}3\,\varphi_{\mathbf{1},\mathbf{6}}(x), \qquad \varphi_s(x) = \frac{1}{s}\,\varphi\Big(\frac{x}{s}\Big)$$

- $\varphi_s$: densitatea lui $N(0, s^2)$; varianța mixturii $0{,}7 \times 0{,}6^2 + 0{,}3 \times 1{,}6^2 = 1{,}02$; $K = 3\,\mathbb{E}\sigma^4/(\mathbb{E}\sigma^2)^2 = 5{,}9$
- mai multe zile foarte mici și mai multe zile foarte mari decît la distribuția Normală cu aceeași varianță
- $P(|r_t| > 4)$: 0,37% în mixtură, față de 0,0075% pentru distribuția Normală

## Inferența în prezența grupării volatilității

- Cînd volatilitatea este grupată, $\mathrm{Var}(\hat{\rho}_k) > 1/T$
  - banda $\pm 1{,}96/\sqrt{T}$ este prea îngustă, iar $Q(m)$ respinge $H_0$ prea des
- Varianța lui $\hat{\rho}_k$ se estimează direct din date Lobato, Nankervis & Savin (2001)

$$\hat{\tau}_k = \frac{\sum_{t>k} x_t^2 x_{t-k}^2}{\left(\sum_t x_t^2\right)^2}$$

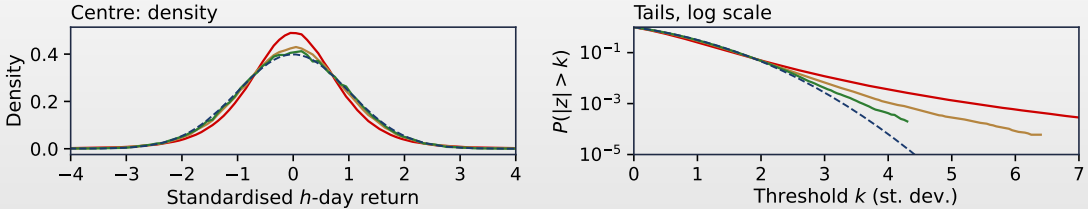
  - pentru randamente i.i.d. $\hat{\tau}_k \approx 1/T$; $T\hat{\tau}_k$ arată de cîte ori varianța lui $\hat{\rho}_k$ depășește $1/T$
  - banda robustă $\pm 1{,}96\sqrt{\hat{\tau}_k}$
- Ljung–Box robust (B4): fiecare $\hat{\rho}_k^2$ este împărțit la propria varianță estimată

$$\tilde{Q}(m) = \sum_{k=1}^{m} \frac{\hat{\rho}_k^2}{\hat{\tau}_k}$$

  - același $H_0$ ca $Q(m)$, comparat tot cu $\chi_m^2$
  - exemplu: $T\hat{\tau}_1 = 4$ implică o bandă robustă de $\sqrt{4} = 2$ ori mai largă decît $\pm 1{,}96/\sqrt{T}$

## Gaussianitatea prin agregare

- **Randamentul logaritmic pe $h$ zile**: suma a $h$ randamente zilnice consecutive, $r_t^{(h)} = r_t + r_{t-\mathbf{1}} + \cdots + r_{t-h+\mathbf{1}}$
  - $h$: orizontul, în zile; $\kappa$: excesul de kurtosis al randamentului zilnic
  - pentru randamente i.i.d., excesul de kurtosis pe $h$ zile este $\kappa/h \to 0$: suma se apropie de distribuția Normală



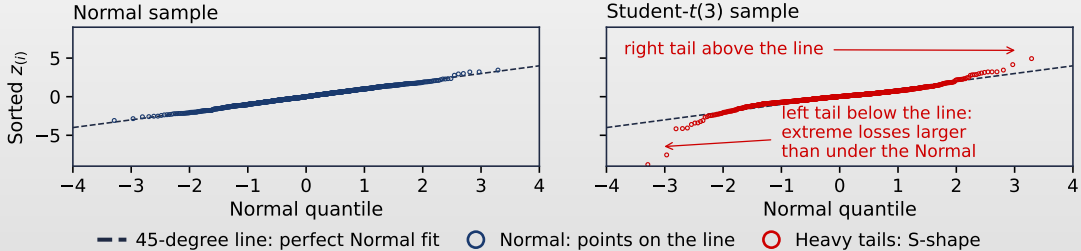

- Randamente zilnice simulate i.i.d. cu cozi groase ($\kappa = 6$), însumate pe $h = 1, 5, 20$ zile și standardizate
  - excesul de kurtosis scade de la 6 la 1,2 și la 0,3; coada (dreapta) se apropie de cea a distribuției Normale
  - cu gruparea volatilității, apropierea este mai lentă decît $\kappa/h$ (B5)

# Graficul QQ

- **Graficul QQ** (cuantilă–cuantilă, *QQ plot*): cuantilele datelor față de cuantilele distribuției $N(0,1)$, cîte un punct pentru fiecare zi

$$\left(\Phi^{-1}\left(\frac{i-0{,}5}{T}\right),\ z_{(i)}\right), \qquad i = 1, \ldots, T$$

  - $z_{(1)} \leq \cdots \leq z_{(T)}$: randamentele standardizate sortate crescător; $(i-0{,}5)/T$: proporția corespunzătoare




- Pe dreapta de 45 de grade: distribuția Normală (stînga); formă de S: cozi groase (dreapta, Student-$t$ cu $\nu = 3$)
  - coada stîngă sub dreaptă, coada dreaptă deasupra ei: cuantilele extreme ale datelor depășesc cuantilele Normale

# A1 [Rezolvat]: randamente și anualizare — cerința

- **Întrebări**
  - ce randament se adună în timp?
  - cum se anualizează mărimile zilnice?
- **Date de intrare**
  - o acțiune închide la 100, 110, 99, 104,5 în zilele 0–3 și nu plătește dividend
  - un indice are un randament logaritmic mediu zilnic de 0,04% și o volatilitate zilnică de 1,2%
  - Bitcoin are o volatilitate zilnică de 3,5%
  - randamentele sînt necorelate serial, cu varianță constantă
- **Cerințe**
  1. Calculați cele trei randamente simple $R_t$ și cele trei randamente logaritmice $r_t$, toate în %.
  2. Calculați randamentele pe 3 zile $P_3/P_0 - 1$ și $\ln(P_3/P_0)$ și comparați-le cu $\sum_t R_t$ și $\sum_t r_t$.
  3. Anualizați randamentul logaritmic mediu și volatilitatea indicelui cu $A = 252$.
  4. Anualizați volatilitatea Bitcoin și justificați alegerea lui $A$.
- **De raportat**: un tabel cu coloanele zi, preț, $R_t$, $r_t$ și cele trei valori anualizate
- Notebook: secțiunea A1

◂ Înapoi ▸ Lista exercițiilor

# A1 [Rezolvat]: rezolvare (1/2)



# A1 [Rezolvat]: rezolvare (2/2)

| $t$ | $P_t$ | $R_t$ | $r_t$ |
|---|---|---|---|
| 0 | 100 | – | – |
| 1 | 110 | +10,00% | +9,53% |
| 2 | 99 | −10,00% | −10,54% |
| 3 | 104,5 | +5,56% | +5,41% |
| sumă | | 5,56% | 4,40% |
| $0 \to 3$ | | 4,50% | 4,40% |

- Suma randamentelor logaritmice este $\ln(104{,}5/100)$: termenii intermediari $\ln P_t$ se anulează
- Randamentele simple se compun: $1{,}1 \times 0{,}9 \times 1{,}0556 - 1 = 4{,}5\%$

- Indice
  - randamentul logaritmic mediu anualizat $252 \times 0{,}04\% = 10{,}08\%$
  - volatilitatea $\sqrt{252} \times 1{,}2\% = 15{,}87 \times 1{,}2\% = 19{,}05\%$
- Bitcoin se tranzacționează zilnic: $\sqrt{365} \times 3{,}5\% = 66{,}9\%$
  - cu $\sqrt{252}$ s-ar obține 55,6%, cu 17% mai puțin

MFM_ch1_seminar_tests

# A2(a) [Propus]: amplitudinea unui drawdown — cerința

- **Model**: A1 [Rezolvat]; secțiunea A2 din notebook
- **Întrebări**
  - care este cea mai mare pierdere față de un vîrf anterior?
  - cum se raportează aceasta la rata anuală de creștere?
- **Date de intrare**
  - valorile portofoliului la sfîrșit de an $V_0, \ldots, V_5 = 100, 120, 90, 110, 130, 117$
  - CAGR pe 5 ani $(V_5/V_0)^{1/5} - 1 = 3{,}19\%$: rata anuală constantă care duce $V_0$ la $V_5$
- **Cerințe**: completați tabelul
  1. Calculați vîrful $M_t = \max_{s \le t} V_s$, cel mai mare $V_s$ din anii $s = 0, \ldots, t$.
  2. Calculați drawdown-ul $DD_t = V_t/M_t - 1 \le 0$, pierderea relativă față de vîrf.
  3. Determinați drawdown-ul maxim $MDD = \min_t DD_t$ și anul în care apare.
  4. Calculați raportul Calmar CAGR$/|MDD|$, creșterea anuală pe unitatea de pierdere maximă.

| Anul | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $V_t$ | 100 | 120 | 90 | 110 | 130 | 117 |
| $M_t$ | 100 | 120 | ... | ... | ... | ... |
| $DD_t$ | 0% | 0% | ... | ... | ... | ... |

- **Interpretare**: de ce scăderea la 117 din anul 5 nu este drawdown-ul maxim?

◂ Înapoi ▸ Lista exercițiilor

# B1 [Rezolvat]: statistici descriptive — cerința

- **Întrebări**
  - cît de groase sînt cozile BET-TR, Bitcoin și EUR/RON?
  - în ce măsură depinde kurtosisul de cîteva zile extreme?
- **Date**: `r_bet`, `r_btc`, `r_eur` din secțiunea de pregătire
  - $T = 2\,992,\ 4\,384,\ 5\,352$; $A = 252,\ 365,\ 252$
- **Cerințe**
  1. Calculați media, abaterea standard, volatilitatea anualizată, asimetria, excesul de kurtosis, minimul și maximul fiecărei serii.
  2. Calculați statistica JB (definiție) și cei doi termeni ai săi.
  3. Comparați JB cu 5,99, valoarea critică la 5% a lui $\chi^2_2$.
  4. Recalculați fiecare exces de kurtosis după eliminarea zilei cu cel mai mare $|r_t|$ și după eliminarea celor mai mari cinci.
  5. Recalculați excesul de kurtosis Crow–Siddiqui pe cuantile, $K_{CS}$ (definiție), pe aceleași eșantioane reduse.
- **De raportat**: un tabel pentru pașii 1–3 și un grafic al kurtosisului în funcție de numărul de zile eliminate
- **Interpretare**: ce măsură a grosimii cozilor puteți raporta cu încredere?

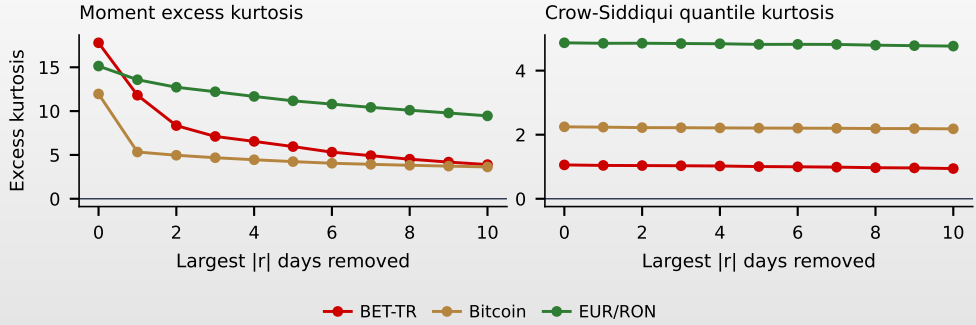
◂ Înapoi ▸ Lista exercițiilor

# B1 [Rezolvat]: BET-TR pas cu pas

- Pasul 1: $\bar{r} = 0{,}077\%$, $\hat{\sigma} = 0{,}978\%$, volatilitatea $\sqrt{252} \times 0{,}978\% = 15{,}5\%$
  - $\hat{S} = -1{,}40$, $\hat{K} - 3 = 17{,}79$
  - cea mai mare scădere zilnică, $-11{,}86\%$, pe 19 dec. 2018 (taxa pe activele bancare)
  - cea mai mare creștere zilnică, $+6{,}71\%$, pe 24 dec. 2018
- Pasul 2: $JB = 40\,433$
  - termenul de asimetrie $2\,992 \times 1{,}40^2/6 = 977$
  - termenul de kurtosis $2\,992 \times 17{,}79^2/24 = 39\,455$, 98% din statistică
- Pasul 3: $40\,433 > 5{,}99$, $p < 10^{-10}$
- Pasul 4: exces de kurtosis 17,8 cu toate zilele
  - 11,8 fără 19 dec. 2018
  - 5,95 fără cele mai mari cinci zile: 19 dec. 2018, 16 mar. 2020, 9 mar. 2020, 24 dec. 2018, 9 mar. 2022
- Pasul 5: Crow–Siddiqui 1,06 cu toate zilele, 1,01 fără cele mai mari cinci zile
- O singură zi generează o treime din kurtosisul calculat pe momente
  - măsura bazată pe cuantile rămîne practic neschimbată

| | BET-TR | Bitcoin | EUR/RON |
|---|---|---|---|
| Medie (% pe zi) | 0,077 | 0,118 | 0,007 |
| Abatere std. (% pe zi) | 0,978 | 3,485 | 0,279 |
| Volatilitate (% pe an) | 15,5 | 66,6 | 4,4 |
| Asimetrie | −1,40 | −0,70 | 0,73 |
| Exces de kurtosis | 17,8 | 12,0 | 15,1 |
| Min (%) | −11,9 | −46,5 | −2,5 |
| Max (%) | 6,7 | 22,5 | 2,9 |
| JB | 40 433 | 26 482 | 51 575 |
| Fără cele 5 zile extreme | 5,95 | 4,24 | 11,17 |

# B1 [Rezolvat]: sensibilitatea kurtosisului




- Toate statisticile JB depășesc de peste 4 000 de ori valoarea critică 5,99
- Bitcoin fără ziua de 12 mar. 2020: kurtosis 12,0 $\rightarrow$ 5,3
- Interpretare: cozile sînt groase conform ambelor măsuri
  - raportați excesul de kurtosis pe cuantile, care nu depinde de cîteva zile extreme

MFM_ch1_seminar_tests

# A3 [Rezolvat]: Jarque–Bera și Ljung–Box, pas cu pas — cerința

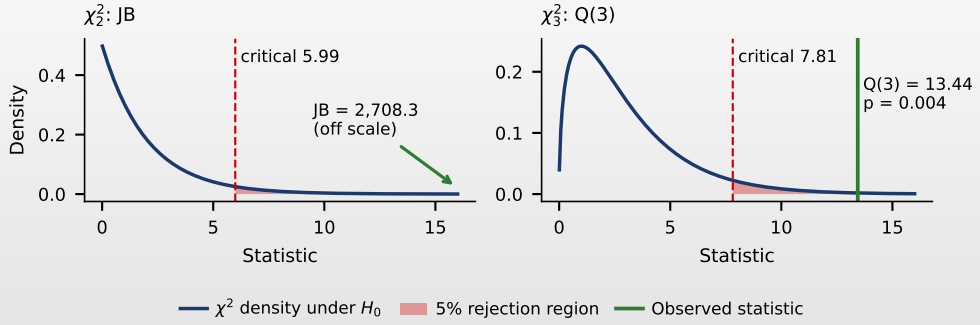
- **Întrebări**
  - cum se trece de la o statistică de test la o decizie?
  - pe ce ipoteze se sprijină decizia?
- **Date de intrare**: două eșantioane ipotetice de randamente zilnice; teste la pragul de 5%
  - eșantionul 1: $T = 2\,500$, asimetria $\hat{S} = -0{,}5$, kurtosisul $\hat{K} = 8$ (nu excesul)
  - eșantionul 2: $T = 1\,000$, $\hat{\rho}_1 = 0{,}10$, $\hat{\rho}_2 = -0{,}05$, $\hat{\rho}_3 = 0{,}03$
- **Cerințe**
  1. Calculați cei doi termeni ai statisticii JB Jarque & Bera (1980) (definiție), statistica și $p = P(\chi^2_2 \geq JB)$.
  2. Calculați cei trei termeni ai statisticii Ljung–Box $Q(3)$ Ljung & Box (1978) (definiție, cu $m = 3$), statistica și $p = P(\chi^2_3 \geq Q)$.
  3. Formulați $H_0$ pentru fiecare test.
  4. Identificați, pentru fiecare distribuție de referință $\chi^2$, ipoteza care nu mai este satisfăcută în prezența grupării volatilității.
- **De raportat**: termenii, cele două statistici, valorile p și deciziile
- Notebook: secțiunea A3

◂ Înapoi ▸ Lista exercițiilor

## A3 [Rezolvat]: pas cu pas

1. $\hat{S}^2/6 = 0{,}25/6 = 0{,}042$; $(\hat{K}-3)^2/24 = 25/24 = 1{,}042$
   - $JB = 2\,500 \times 1{,}084 = 2\,708{,}3$, $p < 10^{-10}$: respingem $H_0$
   - 96% din JB este termenul de kurtosis
2. $0{,}01/999 = 1{,}001 \times 10^{-5}$, $0{,}0025/998 = 0{,}251 \times 10^{-5}$, $0{,}0009/997 = 0{,}090 \times 10^{-5}$
   - $Q(3) = 1\,000 \times 1\,002 \times 1{,}342 \times 10^{-5} = 13{,}44$, $p = 0{,}004$: respingem $H_0$
3. $H_0$ pentru fiecare test
   - JB: randamente i.i.d. cu distribuția Normală
   - Ljung–Box: absența autocorelației
4. Ipotezele din spatele distribuțiilor de referință
   - Ljung–Box: $\chi^2_3$ presupune randamente i.i.d.; cu volatilitatea grupată, $\mathrm{Var}(\hat{\rho}_k) > 1/T$, iar $Q$ respinge prea des (A6, B4)
   - JB: $\chi^2_2$ presupune și independența; cu volatilitatea grupată, respingerea privește ipoteza comună „i.i.d. și Normală”

# A3 [Rezolvat]: distribuțiile de referință




- Valoarea p este aria de sub densitatea $\chi^2$ la dreapta statisticii observate
- Interpretare: pentru un eșantion i.i.d. ambele respingeri sînt valide
  - pentru randamente cu volatilitate grupată, doar respingerea prin JB a ipotezei "i.i.d. și distribuție Normală" este validă
  - în acest caz, valoarea p a lui $Q(3)$ este subestimată

MFM_ch1_seminar_tests

# B2 [Propus]: cozi groase și VaR 1% pe BET-TR — cerința (1/2)

- **Model**: B1 [Rezolvat]; secțiunea B2 din notebook
- **Întrebare**: cît de mare este VaR 1% pe o zi, adică pragul de pierdere depășit cu probabilitatea de 1%, sub trei ipoteze de distribuție?
- **Date**: `r_bet`, $T = 2\,992$
  - modelele sînt estimate și verificate pe **aceleași** eșantion: comparația este descriptivă
  - nu este un backtesting (verificarea VaR pe date ulterioare estimării)
- **Cerințe**
  1. Standardizați $z_t = (r_t - \bar{r})/\hat{\sigma}$ și numărați zilele cu $|z_t| > 3$.
  2. Comparați ponderea lor cu ponderea teoretică sub distribuția Normală, $P(|Z| > 3) = 0{,}27\%$ pentru $Z \sim N(0,1)$.
  3. Estimați o distribuție Student-$t$ pentru $z_t$ prin verosimilitate maximă (`scipy.stats.t.fit`: grade de libertate $\nu$, locație $m$, scală $s$).

Înapoi Lista exercițiilor

# B2 [Propus]: cozi groase și VaR 1% pe BET-TR — cerința (2/2)

- **Cerințe** (continuare)
  4. Calculați $\text{VaR}_{1\%} = -q_r$, cu $q_r = \bar{r} + \hat{\sigma} q_z$, pentru trei variante ale cuantilei de 1%.
     - $q_z$: cuantila de 1% a lui $z_t$; $q_r$: aceeași cuantilă exprimată în randamente (inversa standardizării)
     - Student-$t$: $q_z = m + s\, t_\nu^{-1}(0{,}01)$, cu $t_\nu^{-1}$ cuantila Student-$t$ standard
     - distribuția Normală: $q_z = \Phi^{-1}(0{,}01) = -2{,}326$
     - empiric: $q_r$ = cuantila de 1% a lui $r_t$
  5. Numărați depășirile (zilele cu $r_t < q_r$, adică cu o pierdere mai mare decît VaR) și comparați-le cu numărul așteptat $0{,}01\, T = 29{,}9$.
- **De raportat**: un tabel metodă / VaR / depășiri și un grafic al cozii stîngi cu cele trei cuantile
- **Interpretare**: de ce are VaR-ul empiric, prin construcție, circa 30 de depășiri?

# A4 [Rezolvat]: kurtosisul unei mixturi de volatilitate — cerința

- **Întrebare**: poate volatilitatea aleatoare să genereze, singură, cozi groase?
- **Ipoteze**: $r_t = \sigma_t \varepsilon_t$, $\varepsilon_t$ i.i.d. $N(0,1)$ și independent de $\sigma_t$
  - $\sigma_t$: volatilitatea aleatoare a zilei $t$; $\varepsilon_t$: șocul standardizat
  - $\sigma_t = 1\%$ cu probabilitatea 0,8 și 2% cu probabilitatea 0,2
  - $\sigma_t$ este extras **independent în fiecare zi**
- **Cerințe**
  1. Folosind independența și $\mathbb{E}\varepsilon^4 = 3$, arătați că kurtosisul teoretic (nu excesul) este $K = \mathbb{E}r^4/(\mathbb{E}r^2)^2 = 3\,\mathbb{E}\sigma^4/(\mathbb{E}\sigma^2)^2$.
  2. Completați un tabel cu cele două stări: probabilitatea, $\sigma^2$ (în %$^2$) și $\sigma^4$ (în %$^4$).
  3. Calculați $\mathbb{E}\sigma^2$, $\mathbb{E}\sigma^4$ și $K$.
  4. Folosiți inegalitatea lui Jensen, $\mathbb{E}[X^2] \geq (\mathbb{E}X)^2$ cu $X = \sigma^2$, pentru a arăta că $K \geq 3$.
  5. Stabiliți dacă acest model produce gruparea volatilității.
- **De raportat**: tabelul, $K$ și excesul de kurtosis
- **Interpretare**: puneți rezultatul în legătură cu ipoteza mixturii de distribuții a lui Clark (1973)
- Notebook: secțiunea A4

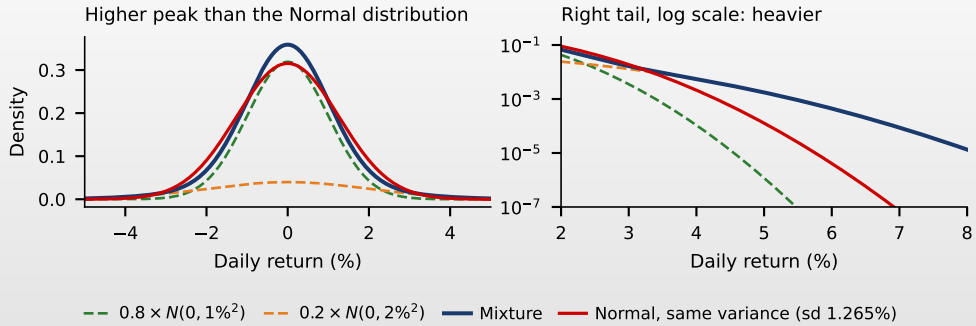
Înapoi Lista exercițiilor

# A4 [Rezolvat]: momente

| Stare | Prob. | $\sigma^2$ | $\sigma^4$ |
|---|---|---|---|
| calmă | 0,8 | 1 | 1 |
| turbulentă | 0,2 | 4 | 16 |
| medie | | 1,6 | 4,0 |

1. $\mathbb{E}r^4 = \mathbb{E}\sigma^4\,\mathbb{E}\varepsilon^4 = 3\,\mathbb{E}\sigma^4$ și $\mathbb{E}r^2 = \mathbb{E}\sigma^2$
2. Tabelul din stînga
3. $\mathbb{E}\sigma^2 = 1{,}6$, $\mathbb{E}\sigma^4 = 4{,}0$
   - $K = 3 \times 4{,}0/1{,}6^2 = 12/2{,}56 = 4{,}69$; excesul de kurtosis 1,69
4. $\mathbb{E}[(\sigma^2)^2] \geq (\mathbb{E}\sigma^2)^2$, cu egalitate doar pentru $\sigma$ constant: $K \geq 3$
5. Fără gruparea volatilității: stările sînt independente de la o zi la alta
   - deci $|r_t|$ este necorelat
   - în A9, stările devin persistente

# A4 [Rezolvat]: densități




- $P(|r_t| > 4\%)$: 0,92% în mixtură, 0,16% pentru distribuția Normală cu aceeași variață
- Interpretare: timpul operațional aleator (ritmul sosirii informației) al lui Clark (1973) este o astfel de mixtură
  - variabilitatea volatilității generează cozi groase

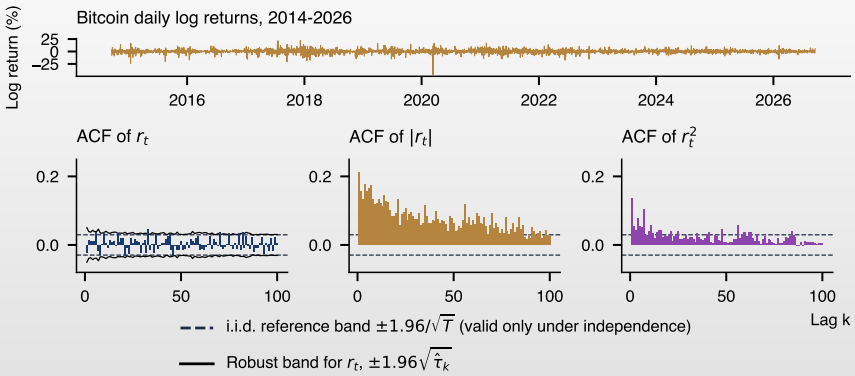
MFM_ch1_seminar_tests

# B3 [Rezolvat]: necorelate, dar nu independente (Bitcoin) — cerința

- **Întrebare**: conțin randamentele trecute ale Bitcoin informație despre direcția sau doar despre mărimea randamentului curent?
- **Date**: `r_btc`, 18 sept. 2014 – 18 sept. 2026, $T = 4\,384$
- **Cerințe**
  1. Reprezentați grafic randamentele.
  2. Reprezentați grafic ACF pentru $r_t$, $|r_t|$ și $r_t^2$ la lagurile 1–100, cu banda de referință i.i.d. $\pm 1{,}96/\sqrt{T} = \pm 0{,}030$.
  3. Calculați Ljung–Box $Q(10)$ pe $r_t$ și pe $r_t^2$ (`acorr_ljungbox`), cu valorile p din $\chi^2_{10}$.
  4. Aplicați testul ARCH-LM cu $q = 5$ laguri Engle (1982) (definiție) și comparați $(T-5)R^2$ cu $\chi^2_5$.
- **De raportat**: graficul, autocorelațiile la lagul 1 și cele trei statistici
  - fiecare însoțită de seria analizată și de ipoteza $H_0$
- **Interpretare**: $Q(10)$ pe $r_t$ are $p \approx 3\%$
  - ținînd cont de pasul 4, este validă distribuția lui de referință i.i.d.?
- Notebook: secțiunea B3

◂ Înapoi ▸ Lista exercițiilor

# B3 [Rezolvat]: randamente și ACF




- Direcția: ACF a lui $r_t$ nu prezintă nicio structură
- Mărimea: $|r_t|$ și $r_t^2$ sînt autocorelate pe zeci de laguri

MFM_ch1_seminar_tests

## B3 [Rezolvat]: teste

2. $\hat{\rho}_1$: $-0{,}022$ pentru $r_t$ (în interiorul benzii), 0,213 pentru $|r_t|$ și 0,137 pentru $r_t^2$
3. Rezultatele $Q(10)$
   - pe $r_t$: 20,2, $p = 0{,}028$
   - pe $r_t^2$: 212,7, $p \approx 4 \times 10^{-40}$
4. ARCH-LM(5) $= 114{,}7$, $p \approx 4 \times 10^{-23}$
   - distribuția sa de referință $\chi^2$ presupune $\mathbb{E}r^4 < \infty$, deci valoarea p este doar orientativă

- Valorile absolute ale randamentelor sînt dependente în timp: randamentele nu sînt i.i.d.
  - valoarea p de 2,8% a lui $Q(10)$ pe $r_t$ se bazează pe o ipoteză încălcată
- Banda i.i.d. este doar un reper, nu un test valid pentru $|r_t|$ și $r_t^2$
  - A6 și B4 corectează inferența pentru $r_t$

## B4 [Rezolvat]: Ljung–Box robust la heteroscedasticitate — cerința (1/2)

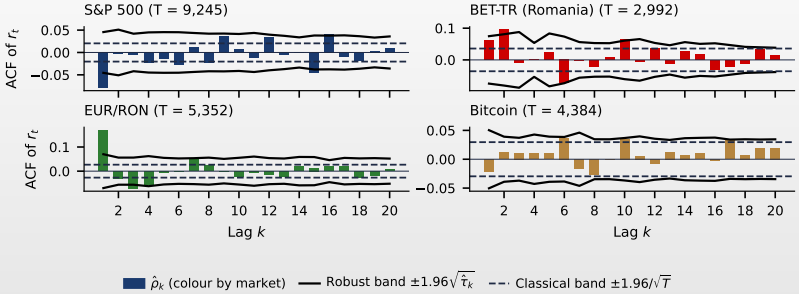
- **Întrebare**: ce autocorelații ale randamentelor rămîn semnificative atunci cînd varianța lui $\hat{\rho}_k$ este corectată pentru gruparea volatilității?
- **Date**: `r_sp`, `r_bet`, `r_eur`, `r_btc`; $x_t = r_t - \bar{r}$
- **Cerințe**
  1. Pentru $k = 1, \ldots, 20$ calculați $\hat{\rho}_k$ și $\hat{\tau}_k$ Lobato, Nankervis & Savin (2001), varianța estimată a lui $\hat{\rho}_k$ (definiție).
  2. Reprezentați $\hat{\rho}_k$ cu banda clasică $\pm 1{,}96/\sqrt{T}$ și banda robustă $\pm 1{,}96\sqrt{\hat{\tau}_k}$.
  3. Calculați $Q(10)$ clasic și $\tilde{Q}(10)$ robust.
  4. Comparați-le cu $\chi^2_{10}$, a cărei valoare critică la 5% este 18,3.
  5. Numărați lagurile dintre 1 și 20 aflate în afara fiecărei benzi.
     - sînt verificări punctuale, diferite de testul conjunct pe 10 laguri de la pașii 3–4

Înapoi Lista exercițiilor

# B4 [Rezolvat]: Ljung–Box robust la heteroscedasticitate — cerința (2/2)

- **Ipotezele lui $\tilde{Q}$**
  - un moment de ordinul patru finit, $\mathbb{E}r^4 < \infty$, necesar pentru ca $\hat{\tau}_k$ să estimeze o varianță finită
  - termeni încrucișați nuli, $\mathbb{E}[x_t^2 x_{t-k} x_{t-l}] = 0$ pentru $k \neq l$
  - adică $x_t^2$ nu este corelat cu produsele $x_{t-k}x_{t-l}$ ale randamentelor trecute de la laguri diferite
- **Interpretare**
  - pentru ce serii rămîne semnificativă autocorelația?
  - este credibilă ipoteza momentului de ordinul patru?
- Notebook: secțiunea B4

# B4 [Rezolvat]: ACF cu ambele benzi




- S&P 500, lagul 1: $\hat{\rho}_1 = -0{,}079$, $T\hat{\tau}_1 = 5{,}00$
  - varianța lui $\hat{\rho}_1$ este de 5 ori mai mare decît valoarea i.i.d. $1/T$
  - banda robustă este $1{,}96\sqrt{5{,}00/9\,245} = 0{,}046$, nu $0{,}020$
- La lagul 1, banda robustă este de $\sqrt{T\hat{\tau}_1}$ ori mai largă decît cea clasică: de la 1,7 ori (Bitcoin) pînă la 2,6 ori (EUR/RON)

MFM_ch1_seminar_tests

# B4 [Rezolvat]: teste conjuncte și interpretare

| | S&P 500 | BET-TR | EUR/RON | Bitcoin |
|---|---|---|---|---|
| $Q(10)$ (p) | 91,6 ($< 10^{-10}$) | 70,0 ($< 10^{-10}$) | 221,7 ($< 10^{-10}$) | 20,2 (0,028) |
| $\tilde{Q}(10)$ (p) | 19,0 (0,041) | 16,9 (0,076) | 37,4 ($<$ 0,001) | 11,9 (0,29) |
| Laguri în afara benzii, clasic / robust | 8 / 3 | 5 / 2 | 6 / 2 | 3 / 0 |
| Laguri în afara benzii robuste | 1, 15, 16 | 2, 10 | 1, 3 | niciunul |

- La 5%, testul clasic respinge $H_0$ pentru toate cele patru serii
  - testul robust respinge doar pentru S&P 500 (la limită) și EUR/RON
- Bitcoin (B3): valoarea p de 2,8% devine 29%
  - nicio dovadă de autocorelație liniară la lagurile 1–10
- Pentru EUR/RON rămîne semnificativă autocorelația $\hat{\rho}_1 = 0{,}17$
  - în regimul de flotare administrată (*managed float*), cursul de referință BNR se ajustează prin variații mici, distribuite pe mai multe zile

MFM_ch1_seminar_tests

## B4 [Rezolvat]: ipotezele de verificat

- Termeni încrucișați nuli
  - testul lui Dalla, Giraitis & Phillips (2022) păstrează doar termenii încrucișați semnificativi
  - nu găsește niciunul pentru S&P 500, BET-TR și Bitcoin
- Moment de ordinul patru finit
  - dacă $P(|r_t| > x)$ descrește proporțional cu $x^{-\alpha}$ și indicele cozii $\alpha \leq 4$, $\mathbb{E}r^4$ este infinit
  - atunci valorile p robuste nu sînt valide
  - A5 și B6 estimează indicele cozii după curs

# B5 [Propus]: gaussianitate prin agregare (Bitcoin) — cerința

- **Model**: B1 și B3 [Rezolvat]; secțiunea B5 din notebook
- **Întrebare**: se apropie randamentele Bitcoin pe $h$ zile de distribuția Normală cînd $h$ crește?
- **Date**: `r_btc`, $T = 4\,384$, în blocuri care nu se suprapun, începînd cu prima zi
  - primul bloc de 5 zile este 18–22 sept. 2014, al doilea 23–27 sept. 2014
  - ultimul bloc incomplet (4 zile pentru $h = 5$ și $h = 20$) se elimină
- **Cerințe**
  1. Construiți randamentele logaritmice pe $h$ zile (sume de $h$ randamente zilnice consecutive) pentru $h = 1, 5, 20$.
  2. Reprezentați un grafic QQ al fiecărei serii standardizate față de distribuția Normală, cu $N$, numărul de blocuri, în titlu.
  3. Calculați excesul de kurtosis și eroarea standard (SE, *standard error*) de referință pentru date i.i.d. Normale, $\sqrt{24/N}$.
     - $\sqrt{24/N}$: eroarea standard a excesului de kurtosis estimat din $N$ observații i.i.d. Normale (24, ca în JB)
- **De raportat**: cele trei grafice QQ și un tabel $h$, $N$, exces de kurtosis, $\sqrt{24/N}$
- **Interpretare**
  - de ce subestimează $\sqrt{24/N}$ incertitudinea în acest caz?
  - de ce blocurile disjuncte, deși nu au zile comune, nu elimină dependența volatilității?

◂ Înapoi ▸ Lista exercițiilor

# C1 [Propus]: critica unui răspuns AI — răspunsul

- **Promptul** trimis unui asistent AI: "Sînt randamentele logaritmice zilnice ale BET-TR distribuite Normal? Sînt autocorelate? Dă cod Python și o concluzie"
- **Răspunsul**: presupuneți NumPy importat ca `np` și o coloană `close` în `bet`; evaluați erorile care schimbă o valoare raportată sau o concluzie

```
from scipy import stats
from statsmodels.stats.diagnostic import acorr_ljungbox
r = np.log(bet["close"]).diff().dropna()      # daily log returns
k = stats.kurtosis(r)                          # sample kurtosis
print("excess kurtosis:", k - 3)               # the Normal has kurtosis 3
jb, p = stats.jarque_bera(r)
lb = acorr_ljungbox(r**2, lags=[10])           # Ljung-Box, 10 lags
```

- **Concluzia asistentului**
  - "Exces de kurtosis 14,8: cozi groase"
  - "valoarea p JB 0,000: probabilitatea ca randamentele BET-TR să urmeze distribuția Normală este 0%"
  - "Ljung–Box $Q(10) = 713{,}5$, $p < 0{,}001$: randamentele sînt puternic autocorelate, deci BET-TR este predictibil și nu este un mers aleator"

# C1 [Propus]: critica unui răspuns AI — cerința (1/2)

- **Model**: A3, B3 și B4 [Rezolvat]; secțiunea C1 din notebook
- **Date**: `r_bet`, cele 2 992 de randamente din pregătire
- **Cerințe**
  1. Identificați toate erorile de fond din cod și din cele trei concluzii.
  2. Pentru fiecare eroare, precizați cum poate fi detectată.
     - un test pe date simulate sau o valoare din slide-urile introductive
     - un exercițiu rezolvat sau documentația bibliotecii
  3. Scrieți codul corectat, în formă executabilă.

Înapoi Lista exercițiilor

# C1 [Propus]: critica unui răspuns AI — cerința (2/2)

- **Cerințe** (continuare)
  4. Raportați excesul de kurtosis, JB cu valoarea p, $Q(10)$ pe $r_t$ și pe $r_t^2$ și $\tilde{Q}(10)$ robust din B4.
  5. Formulați ipotezele din spatele valorii p robuste.
  6. Scrieți concluzia corectată în trei fraze.
- **De raportat**
  - codul corectat
  - un tabel valoare AI / valoare corectă
  - un grafic al ACF pentru $r_t$ și $r_t^2$
- Codul generat de AI se execută de obicei fără mesaje de eroare
  - erorile sînt de natură statistică, nu sintactică
  - consemnați erorile identificate în `AI_ERRORS.md`

# A5 [Rezolvat]: estimatorul Hill, pas cu pas — cerința (1/2)

- **Întrebări**
  - cum se estimează indicele cozii din cele mai mari observații?
  - cît de precisă este estimarea?
- **Date de intrare**: cele mai mari 11 randamente zilnice absolute (%) ale unui eșantion ipotetic: 12,0; 9,5; 8,2; 7,4; 6,9; 6,1; 5,8; 5,5; 5,2; 5,0; 4,8
- **Notații**
  - $X_{(1)} \geq X_{(2)} \geq \ldots$: valorile absolute ordonate descrescător (statisticile de ordine)
  - $k$: numărul de observații din coadă; $X_{(k+1)}$: pragul de la care începe coada
- **Ipoteză**: peste pragul $X_{(k+1)}$, coada urmează o distribuție Pareto, $P(X > x) \propto x^{-\alpha}$
  - $\alpha$: indicele cozii; un $\alpha$ mic înseamnă o coadă groasă; $\mathbb{E}|X|^p$ este finit doar pentru $p < \alpha$
  - observațiile sînt independente

Înapoi Lista exercițiilor

# A5 [Rezolvat]: estimatorul Hill, pas cu pas — cerința (2/2)

- **Estimatorul Hill** Hill (1975): inversul mediei log-rapoartelor dintre cele mai mari $k$ valori și prag

$$\hat{\alpha} = \Big(\frac{1}{k}\sum_{i=1}^{k} \ln \frac{X_{(i)}}{X_{(k+1)}}\Big)^{-1}$$

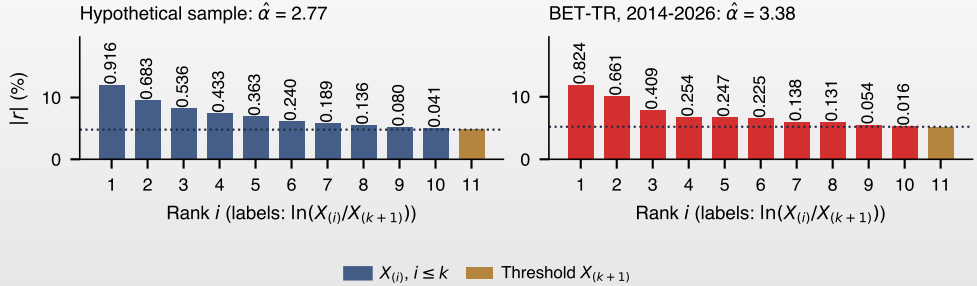
  - o coadă groasă are log-rapoarte mari, deci un $\hat{\alpha}$ mic

- **Cerințe**
  1. Pentru $k = 10$ calculați log-rapoartele $\ln(X_{(i)}/X_{(11)})$, media lor și $\hat{\alpha} = 1/\text{medie}$.
  2. Calculați eroarea standard $\text{SE} = \hat{\alpha}/\sqrt{k}$ (valabilă pentru observații i.i.d.) și intervalul de 95% $\hat{\alpha} \pm 1{,}96\,\text{SE}$.
  3. Precizați ce momente $\mathbb{E}|X|^p$ ($p < \alpha$) există dacă $\alpha$ este egal cu estimarea punctuală.
  4. Precizați care dintre aceste concluzii rămîn valabile la nivelul de încredere de 95%.
  5. Repetați pașii 1–2 pe cele mai mari 11 valori $|r_t|$ ale BET-TR, folosind tabelul afișat de notebook.

- **De raportat**: $\hat{\alpha}$, SE și intervalul pentru ambele eșantioane

- Notebook: secțiunea A5

# A5 [Rezolvat]: eșantionul ipotetic




1. Etichetele din grafic sînt log-rapoartele față de pragul 4,8: suma 3,616, media 0,3616, $\hat{\alpha} = 1/0{,}3616 = 2{,}77$
2. SE $= 2{,}77/\sqrt{10} = 0{,}87$; intervalul [1,05; 4,48]
   - o aproximare grosieră prin distribuția Normală, valabilă pentru cozi Pareto i.i.d.
3. Pentru $\alpha = 2{,}77$, varianța este finită, iar momentul de ordinul patru este infinit
4. Intervalul conține valori sub 2 și peste 4: nicio concluzie nu este susținută la 95%

MFM_ch1_seminar_tests

# A5 [Rezolvat]: tabelul BET-TR

| $i$ | Data | $r_t$ (%) | $X_{(i)}$ (%) | $\ln(X_{(i)}/X_{(\mathbf{11})})$ |
|---|---|---|---|---|
| 1 | 19 dec. 2018 | −11,857 | 11,857 | 0,824 |
| 2 | 16 mar. 2020 | −10,075 | 10,075 | 0,661 |
| 3 | 9 mar. 2020 | −7,827 | 7,827 | 0,409 |
| 4 | 24 dec. 2018 | +6,707 | 6,707 | 0,254 |
| 5 | 9 mar. 2022 | +6,662 | 6,662 | 0,247 |
| 6 | 24 aug. 2015 | −6,512 | 6,512 | 0,225 |
| 7 | 24 mar. 2020 | +5,972 | 5,972 | 0,138 |
| 8 | 17 mar. 2020 | +5,928 | 5,928 | 0,131 |
| 9 | 4 mar. 2022 | −5,490 | 5,490 | 0,054 |
| 10 | 12 mar. 2020 | −5,285 | 5,285 | 0,016 |
| 11 | 18 ian. 2016 | −5,202 | 5,202 | prag |

- Calculul pentru primul rînd:
  $\ln(11{,}857/5{,}202) = \ln 2{,}279 = 0{,}824$
- Suma celor zece log-rapoarte este 2,958;
  $\hat{\alpha}_{\mathbf{10}} = 10/2{,}958 = 3{,}38$
- SE $= 3{,}38/\sqrt{10} = 1{,}07$; intervalul [1,28; 5,48]
- Șapte dintre cele zece zile se situează în decembrie 2018 sau în martie 2020
  - valorile extreme sînt grupate în timp, deci SE calculată sub ipoteza i.i.d. subestimează incertitudinea
- Interpretare: zece observații nu sînt suficiente pentru a stabili dacă momentul de ordinul patru există
  - B6 folosește $k = 2\%$ din $n$

# B6 [Rezolvat]: estimatorul Hill pe ambele cozi, BET-TR și Bitcoin — cerința (1/2)

- **Întrebare**: sînt pierderile mari mai frecvente decît cîștigurile mari de aceeași mărime?
- **Date**: `r_bet` ($n = 2\,992$), `r_btc` ($n = 4\,384$)
  - pierderi `loss = -r[r < 0]`, cîștiguri `gain = r[r > 0]`
  - fiecare serie ordonată descrescător
- **Notații**
  - $\hat{\alpha}(k)$: estimatorul Hill din A5, calculat cu cele mai mari $k$ valori ale unei cozi
  - $\hat{\alpha}_-$: indicele cozii pierderilor; $\hat{\alpha}_+$: indicele cozii cîștigurilor; $\hat{\alpha}_- < \hat{\alpha}_+$ înseamnă că pierderile au coada mai groasă
  - $\lfloor \cdot \rfloor$: partea întreagă
- **Cerințe**
  1. Reprezentați grafic estimarea Hill $\hat{\alpha}(k)$ pentru fiecare coadă, pentru $k$ între 1% și 10% din $n$ (volumul întregului eșantion), cu banda $\hat{\alpha}(1 \pm 1{,}96/\sqrt{k})$.
  2. Calculați $\hat{\alpha}_-$ și $\hat{\alpha}_+$ la $k = \lfloor 0{,}02\, n \rfloor$ (BET-TR 59, Bitcoin 87), cu intervale de 95%.
     - cursul folosește $k = 2{,}5\%$ din $n$ pentru $|r_t|$, care combină ambele cozi

◂ Înapoi ▸ Lista exercițiilor

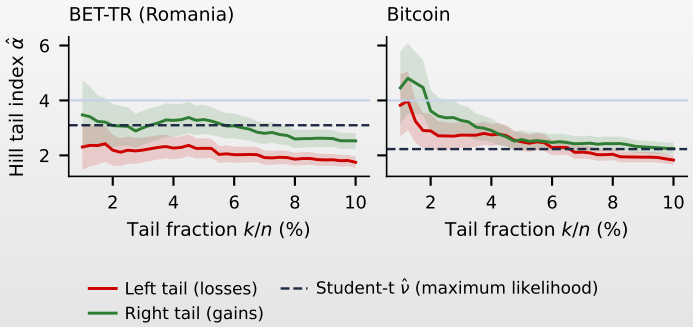
# B6 [Rezolvat]: estimatorul Hill pe ambele cozi, BET-TR și Bitcoin — cerința (2/2)

- **Cerințe** (continuare)
  3. Calculați statistica aproximativă pentru cozi independente și valoarea p bilaterală.

$$z = \frac{\hat{\alpha}_- - \hat{\alpha}_+}{\sqrt{\hat{\alpha}_-^2/k + \hat{\alpha}_+^2/k}}$$

     - numitorul: eroarea standard a diferenței, din $\mathrm{SE} = \hat{\alpha}/\sqrt{k}$ pentru fiecare coadă
     - sub $H_0$: $\alpha_- = \alpha_+$, $z \approx N(0,1)$; $|z| > 1{,}96$: respingem $H_0$ la 5%
  4. Comparați rezultatele cu $\hat{\nu}$ al distribuției Student-$t$ estimate pe randamentele standardizate: 3,10 pentru BET-TR (B2), 2,23 pentru Bitcoin.
- **Interpretare**
  - este asimetria dintre cîștiguri și pierderi semnificativă statistic?
  - cum schimbă dependența dintre cei doi estimatori ai cozilor SE a diferenței?
- Notebook: secțiunea B6

# B6 [Rezolvat]: graficele Hill și estimările




- BET-TR: $\hat{\alpha}_- = 2{,}19$ [1,63; 2,75], $\hat{\alpha}_+ = 3{,}08$ [2,30; 3,87]
- Bitcoin: $\hat{\alpha}_- = 2{,}89$ [2,28; 3,50], $\hat{\alpha}_+ = 3{,}62$ [2,86; 4,38]
- De exemplu, pierderile BET-TR: $2{,}19 \times (1 \pm 1{,}96/\sqrt{59}) = 2{,}19 \times (1 \pm 0{,}255)$

MFM_ch1_seminar_tests

## B6 [Rezolvat]: testul și interpretarea

- BET-TR: $\sqrt{2{,}19^2/59 + 3{,}08^2/59} = \sqrt{0{,}0812 + 0{,}1612} = 0{,}492$; $z = (2{,}19 - 3{,}08)/0{,}492 = -1{,}82$, $p = 0{,}069$
- Bitcoin: $\sqrt{2{,}89^2/87 + 3{,}62^2/87} = 0{,}497$; $z = (2{,}89 - 3{,}62)/0{,}497 = -1{,}47$, $p = 0{,}14$
- Decizie: coada pierderilor este mai groasă pe ambele piețe, dar $H_0 : \alpha_- = \alpha_+$ nu este respinsă la 5%
  - statistica tratează cele două cozi ca eșantioane independente
  - dependența schimbă incertitudinea: $\mathrm{Var}(\hat{\alpha}_- - \hat{\alpha}_+) = \mathrm{Var}(\hat{\alpha}_-) + \mathrm{Var}(\hat{\alpha}_+) - 2\,\mathrm{Cov}(\hat{\alpha}_-, \hat{\alpha}_+)$
  - gruparea în timp a valorilor extreme poate mări fiecare dispersie, iar termenul de covarianță o poate compensa: direcția nu se cunoaște dinainte
  - un bootstrap pe blocuri aplicat simultan ambelor cozi, pe randamentele originale, îi estimează direcția și mărimea (extindere)
- $\hat{\nu}$ este estimat pe întreaga distribuție, pe cînd $\hat{\alpha}$ folosește doar coada
  - valoarea 2,23 pentru Bitcoin este sub ambii indici ai cozilor
  - Student-$t$ nu poate descrie simultan partea centrală și cozile

MFM_ch1_seminar_tests

# A6 [Propus]: lărgirea benzii autocorelației în prezența grupării volatilității — cerința (1/2)

- **Model**: A4 [Rezolvat]; secțiunea A6 din notebook
- **Cadru**: $x_t = \sigma_t \varepsilon_t$, cu $\varepsilon_t$ i.i.d. $N(0,1)$, independent de procesul staționar al volatilității
  - în acest caz, $x_t$ este necorelat serial
  - $\mathrm{Var}(\hat{\rho}_1) \approx v_1 / T$, cu $v_1 = \mathbb{E}[x_t^2 x_{t-1}^2] / (\mathbb{E}x_t^2)^2$ Lobato, Nankervis & Savin (2001)
  - $v_1$: factorul cu care gruparea volatilității crește varianța lui $\hat{\rho}_1$ față de valoarea i.i.d. $1/T$
- **Date de intrare**
  - A4: $\mathbb{E}\sigma^2 = 1{,}6$, $\mathrm{Var}(\sigma^2) = 4{,}0 - 1{,}6^2 = 1{,}44$ (în $\%^4$)
  - regimuri persistente (A9): $\mathrm{Cov}(\sigma_t^2, \sigma_{t-1}^2) = \lambda \, \mathrm{Var}(\sigma^2)$ cu $\lambda = 0{,}875$
  - $\lambda \in [0,1)$: persistența regimurilor; $\lambda = 0$: stări independente de la o zi la alta
  - S&P 500: $T = 9\,245$, $\hat{\rho}_1 = -0{,}079$, $T\hat{\tau}_1 = 5{,}00$

Înapoi Lista exercițiilor

# A6 [Propus]: lărgirea benzii autocorelației în prezența grupării volatilității — cerința (2/2)

- **Cerințe**
  1. Arătați că $v_1 = 1$ pentru mixtura din A4, ale cărei stări sînt independente de la o zi la alta.
  2. Arătați că $v_1 = 1 + \lambda\,\mathrm{Var}(\sigma^2)/(\mathbb{E}\sigma^2)^2$ pentru regimuri persistente.
  3. Calculați $v_1$ și factorul de lărgire a benzii $\sqrt{v_1}$.
  4. Comparați cu B4: calculați statistica $t = \hat{\rho}_1/\mathrm{SE}$ pentru autocorelația de lag 1 a S&P 500 cu SE i.i.d. $1/\sqrt{T}$ și cu SE robustă $\sqrt{\hat{\tau}_1}$ ($|t| > 1{,}96$: semnificativă la 5%).
  5. Calculați cele două semilățimi ale benzii de 95%.
- **Interpretare**: ce factor lărgește banda: cozile groase sau volatilitatea persistentă?
- **De raportat**: $v_1$ și $\sqrt{v_1}$; un tabel i.i.d./robust cu statisticile $t$ și semilățimile benzilor

## După curs: cerințele rămase

- Aceste trei cerințe folosesc instrumente introduse în Capitolul 1
  - A5 [Rezolvat] și B6 [Rezolvat]: estimatorul Hill al indicelui cozii, pas cu pas și pe ambele cozi
  - A6 [Propus]: care mecanism modifică banda autocorelației din B4
- Parcurgeți A5 și B6 după curs
- Rezolvați A6 înaintea seminarului următor; rezolvarea se discută atunci

# Pachete de extindere: conținut și cunoștințe necesare

| Pachet | Exerciții | Cunoștințe necesare |
|---|---|---|
| 1. Aritmetica randamentelor și drawdown | A1(c) [Rezolvat], A2(b–f), A7 | A1, A2(a) |
| 2. Mixturi, regimuri și cozi | A8 [Rezolvat], A9, A10, A11 | A4, A5, A6 |
| 3. Momente robuste și verificarea modelelor | B1(d) [Rezolvat], B7, B8 | B1, B4, B6 |
| 4. Dependență și asimetrie | B9, B10, B11, B12 | B3, B4 |
| 5. Randament ajustat la risc și estimatori de volatilitate bazați pe amplitudinea zilnică | B13, B14, B15 | A1, A8, B1 |
| 6. Proiecte | C2, C3 | B4, studiul de caz din curs |

- Secțiunile din notebook poartă aceleași etichete ca exercițiile
- Rezolvările exercițiilor [Propus] se discută la seminarul următor

## Mersul aleator și mișcarea browniană

- **Mersul aleator**: logaritmul prețului, $X_t = \ln P_t$, adună șocuri independente

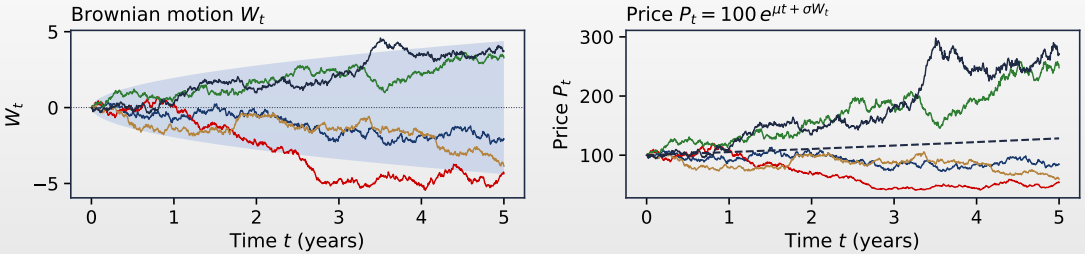
$$X_t = X_{t-1} + r_t, \qquad r_t \text{ i.i.d. } N(\mu\Delta,\ \sigma^2\Delta)$$

  - $\Delta$: lungimea unui pas, în ani (o zi de tranzacționare: $\Delta = 1/252$)
  - $\mu$: drift-ul anual al lui $X_t$; $\sigma$: volatilitatea anuală
  - reperul față de care se definesc faptele stilizate: randamente i.i.d. cu distribuția Normală
- **Mișcarea browniană standard** $W_t$, $t \geq 0$: limita mersului aleator cînd $\Delta \to 0$
  - $W_0 = 0$; creșterea $W_t - W_s$, $s < t$, este independentă de trecut și are distribuția $N(0,\ t-s)$
  - deci $W_t \sim N(0, t)$: abaterea standard crește ca $\sqrt{t}$, iar $|W_t| \leq 1{,}96\sqrt{t}$ cu probabilitatea 95%
- **Prețul** ca mișcare browniană geometrică (A2, A8)

$$X_t = X_0 + \mu t + \sigma W_t, \qquad P_t = P_0\, e^{\mu t + \sigma W_t}$$

  - $P_t > 0$; randamentul logaritmic pe un interval de lungime $\Delta$ are distribuția $N(\mu\Delta,\ \sigma^2\Delta)$

# Traiectorii simulate ale mișcării browniene




- Stînga: cinci traiectorii simulate ale lui $W_t$ pe 5 ani și banda de 95% $\pm 1{,}96\sqrt{t}$, care se lărgește ca $\sqrt{t}$
- Dreapta: aceleași traiectorii transformate în prețuri, $P_t = 100\,e^{\mu t + \sigma W_t}$, cu $\mu = 0{,}05$ și $\sigma = 0{,}20$
  - după 5 ani, prețurile simulate sînt între 54 și 270: incertitudinea crește cu orizontul
  - linia întreruptă $100\,e^{\mu t}$ este mediana prețului, deoarece mediana lui $W_t$ este 0

# A1(c) [Rezolvat]: randamentul unui portofoliu — cerința și rezolvarea

- **Întrebare**: ce randament se agregă liniar între active?
- **Date de intrare**: un portofoliu investit 60% în activul A și 40% în activul B, pe o zi; $R_A = +5\%$, $R_B = -2\%$
- **Cerințe**
  1. Calculați randamentul simplu al portofoliului $R_p = 0{,}6R_A + 0{,}4R_B$.
  2. Calculați randamentul logaritmic ponderat $0{,}6\, r_A + 0{,}4\, r_B$, cu $r_A = \ln(1 + R_A)$, $r_B = \ln(1 + R_B)$.
  3. Calculați randamentul logaritmic adevărat al portofoliului $\ln(1 + R_p)$.
- **Rezolvare**
  1. $R_p = 0{,}6 \times 5\% + 0{,}4 \times (-2\%) = 2{,}20\%$
  2. $0{,}6 \times \ln 1{,}05 + 0{,}4 \times \ln 0{,}98 = 0{,}6 \times 4{,}879\% - 0{,}4 \times 2{,}020\% = 2{,}119\%$
  3. $\ln 1{,}022 = 2{,}176\%$
- Interpretare: randamentele simple se agregă liniar între active, cele logaritmice în timp (A1)
  - pentru variații zilnice, diferența dintre cele două reguli este mică
  - pentru variații mari, diferența devine importantă

Înapoi Lista exercițiilor

# A2(b–f) [Propus]: un reper pentru drawdown — cerința (1/2)

- **Model**: A2(a); secțiunea A2 din notebook
- **Ipoteză**: logaritmul prețului este $X_t = \mu t + \sigma W_t$
  - $W_t$ este o mișcare browniană standard: creșteri independente, $W_t \sim N(0, t)$
  - $\mu$: drift-ul anual al logaritmului prețului; $\sigma$: volatilitatea anuală; $T$: orizontul, în ani (aici nu numărul de randamente)
  - drawdown-ul maxim logaritmic $D = \max_{s \le t \le T}(X_s - X_t) = -\ln(1 + MDD)$: cea mai mare scădere a lui $X$ de la un vîrf $s$ la un moment ulterior $t$
  - pentru $\mu = 0$: $\mathbb{E}[D] = \sqrt{\pi/2}\,\sigma\sqrt{T}$ Magdon-Ismail et al. (2004), cu $\sqrt{\pi/2} = 1{,}2533$
- **Date de intrare**: aur spot, 1990–2026
  - $\sigma = 16{,}3\%$, $T = 36{,}7$ ani, randament logaritmic mediu 6,5% pe an
  - $MDD$ observat $= -44{,}7\%$, deci $D = 0{,}592$

înapoi Lista exercițiilor

# A2(b–f) [Propus]: un reper pentru drawdown — cerința (2/2)

- **Cerințe**
  1. Folosiți $W(cT) \overset{d}{=} \sqrt{c}\, W(T)$ pentru a explica de ce $\mathbb{E}[D]$ este proporțional cu $\sqrt{T}$ când $\mu = 0$ ($c > 0$: factorul de scalare a timpului; $\overset{d}{=}$: aceeași distribuție).
  2. Calculați raportul valorilor $\mathbb{E}[D]$ pentru 40 și 10 ani ai aceluiași proces.
  3. Calculați reperul fără drift pentru aur și comparați-l cu $D$ observat.
  4. Simulați 2 000 de traiectorii zilnice (`numpy.random.default_rng(42)`), din log-preț zero, cu 252 de pași pe an și volatilitatea aurului de mai sus cu $\mu = 0$ și cu $\mu = 6{,}5\%$.
  5. Simulați o singură dată traiectoriile de 40 de ani; evaluați primele `round(252*T)` observații (36,7 ani: 9 248 pași). Reprezentați și tabelați media lui $D$ în funcție de $T$, pentru ambele drift-uri, la fiecare an 1–40 și la 36,7 ani.
- **De raportat**: raportul de la pasul 2, reperul $\mathbb{E}[D]$ pentru aur, tabelul și graficul mediei lui $D$ în funcție de $T$
- **Interpretare**: ce implică pașii 1–5 pentru rapoartele Calmar calculate pe eșantioane de lungimi diferite?

# A7 [Propus]: regula $\sqrt{h}$ în prezența autocorelației — cerința (1/2)

- **Model**: A1 [Rezolvat]; secțiunea A7 din notebook
- **Ipoteză**: randamentele urmează un proces autoregresiv de ordinul 1, AR(1), staționar

$$r_t = \rho\, r_{t-1} + e_t, \qquad |\rho| < 1$$

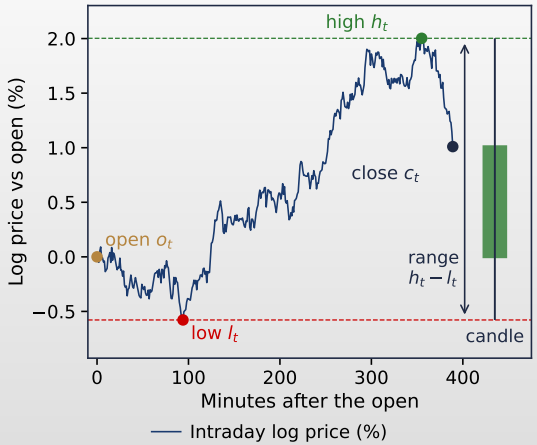
  - $e_t$: șoc necorelat în timp, cu media 0
  - $\rho$: coeficientul autoregresiv, egal cu autocorelația de lag 1; $\rho > 0$: randamentele de același semn tind să se succeadă
  - $\mathrm{Var}(r_t) = \sigma^2$; autocovarianța la lagul $j$ este $\gamma_j = \mathrm{Cov}(r_t, r_{t-j}) = \sigma^2 \rho^j$, iar $\gamma_0 = \sigma^2$
- **Regula $\sqrt{h}$**: varianța randamentului pe $h$ zile este $h\sigma^2$, adevărat doar pentru randamente necorelate
- **Raportul variațelor** (*variance ratio*, VR): $VR(h) = \mathrm{Var}\big(\sum_{j=1}^{h} r_{t+j}\big)/(h\sigma^2)$
  - $VR(h) = 1$ pentru randamente necorelate; $VR(h) > 1$ pentru $\rho > 0$, deci regula $\sqrt{h}$ subestimează riscul
- **Punct de plecare**, $h = 3$: $\mathrm{Var}(r_1 + r_2 + r_3) = 3\gamma_0 + 2(2\gamma_1 + \gamma_2)$: două perechi la lagul 1, o pereche la lagul 2

◂ Înapoi ▸ Lista exercițiilor

# A7 [Propus]: regula $\sqrt{h}$ în prezența autocorelației — cerința (2/2)

- **Cerințe**
  1. Generalizați numărul de perechi: arătați că $\mathrm{Var}\left(\sum_{j=1}^{h} r_{t+j}\right) = h\sigma^2\left[1 + 2\sum_{j=1}^{h-1}(1 - j/h)\rho^j\right]$.
  2. Folosind $\sum_{j=1}^{h-1}\rho^j = \rho(1-\rho^{h-1})/(1-\rho)$ și $\sum_{j\geq 1} j|\rho|^j = |\rho|/(1-|\rho|)^2$, arătați că $VR(h) \to (1+\rho)/(1-\rho)$ cînd $h \to \infty$.
  3. Calcul ilustrativ: calculați $\sqrt{VR(252)}$ și volatilitățile anuale corectate, tratînd autocorelațiile de selecție la lagul 1 drept parametri AR(1).
     - EUR/RON: $\rho = 0{,}170$, volatilitate 4,44%
     - S&P 500: $\rho = -0{,}079$, volatilitate 17,99%
  4. Explicați de ce gruparea volatilității, singură, nu modifică rezultatul de la pasul 1.
- **Interpretare**: ce verificare a ACF trebuie făcută înaintea pasului 3?
- **De raportat**: derivarea variației, două volatilități anuale corectate și verificarea ACF.

# Amplitudinea zilnică și volatilitatea Parkinson




- Pentru pachetele de extindere (A8, B15)
- $O_t, H_t, L_t, C_t$: deschiderea, maximul, minimul și închiderea zilei $t$
  - cu litere mici, logaritmii lor: $o_t = \ln O_t$, $h_t = \ln H_t$, $l_t = \ln L_t$, $c_t = \ln C_t$
- **Amplitudinea logaritmică** (*range*) $h_t - l_t = \ln(H_t/L_t)$
  - folosește maximul și minimul traiectoriei, nu doar prețul de închidere
  - în grafic: amplitudinea este 2,58%, iar $c_t - o_t$ doar 1,01%
- **Estimatorul Parkinson** Parkinson (1980)

$$\hat{\sigma}_P^2 = \frac{1}{4 \ln 2} \cdot \frac{1}{n} \sum_{t=1}^{n} (h_t - l_t)^2$$

  - $n$: numărul de zile; $\hat{\sigma}_P^2$: varianța zilnică estimată
  - $4 \ln 2 = 2{,}77$: pentru o mișcare browniană fără drift, $\mathbb{E}[(h_t - l_t)^2] = 4 \ln 2 \, \sigma^2$

# A8 [Rezolvat]: volatilitatea Parkinson și eficiența ei — cerința (1/2)

- **Întrebare**: cîtă informație despre volatilitate conține amplitudinea zilnică?
- **Notații**
  - $O_t, H_t, L_t, C_t$: deschiderea, maximul, minimul și închiderea zilei $t$
  - logaritmii lor: $o_t = \ln O_t$, $h_t = \ln H_t$, $l_t = \ln L_t$, $c_t = \ln C_t$
  - amplitudinea logaritmică: $h_t - l_t = \ln(H_t/L_t)$
- **Date de intrare**: $H/L$ pentru cinci zile: 1,020, 1,015, 1,030, 1,010, 1,025; $A = 252$
- **Ipoteză**: pe parcursul zilei, logaritmul prețului este o mișcare browniană fără drift, cu volatilitatea $\sigma$
  - fără salturi peste noapte
- **Estimatorul Parkinson** Parkinson (1980): varianța zilnică estimată din amplitudine

$$\hat{\sigma}_P^2 = \frac{\overline{(h-l)^2}}{4 \ln 2}$$

  - $\overline{(h-l)^2}$: media pe zile a pătratelor amplitudinii
  - $4 \ln 2 = 2{,}7726$: sub ipoteză, $\mathbb{E}[(h-l)^2] = 4 \ln 2 \, \sigma^2$, deci împărțirea face estimatorul nedeplasat

Înapoi Lista exercițiilor

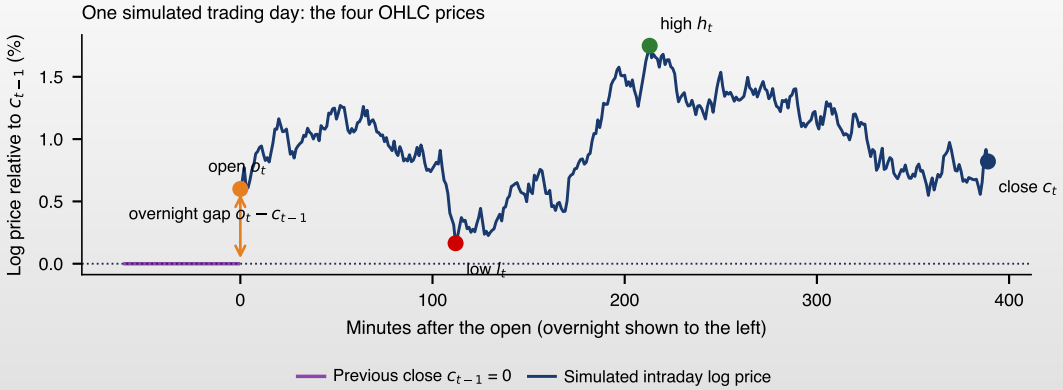
# A8 [Rezolvat]: volatilitatea Parkinson și eficiența ei — cerința (2/2)

**Cerințe**

1. Completați un tabel cu cinci rînduri: $\ln(H/L)$, pătratul său și media pătratelor.
2. Calculați $\hat{\sigma}_P^2$, volatilitatea zilnică și cea anuală.
3. Precizați ce informație nu este folosită de estimator.
4. Explicați ce efect are un drift puternic asupra estimării.
5. Folosind $\mathbb{E}[(h-l)^2] = 4\ln 2\,\sigma^2$ și $\mathbb{E}[(h-l)^4] = 9\zeta(3)\,\sigma^4$, comparați varianța zilnică a lui $(h-l)^2/(4\ln 2)$ cu $\mathrm{Var}(r_t^2) = 2\sigma^4$.
   - $\zeta(3) = \sum_{n\geq 1} n^{-3} = 1{,}202$: funcția zeta a lui Riemann în punctul 3
   - $r_t^2$: estimatorul de la închidere la închidere (*close-to-close*) al variației zilnice, dintr-un randament cu distribuția $N(0, \sigma^2)$
   - estimatorul cu varianța mai mică este mai eficient: are nevoie de mai puține zile pentru aceeași precizie

Notebook: secțiunea A8

# A8 [Rezolvat]: o zi de tranzacționare simulată



# A8 [Rezolvat]: estimarea Parkinson din cinci intervale

| $H/L$ | $\ln(H/L)$ | pătrat ($\times 10^{-4}$) |
|---|---|---|
| 1,020 | 0,0198 | 3,92 |
| 1,015 | 0,0149 | 2,22 |
| 1,030 | 0,0296 | 8,74 |
| 1,010 | 0,0100 | 0,99 |
| 1,025 | 0,0247 | 6,10 |
| medie | | 4,39 |

2. $\hat{\sigma}_P^2 = 0{,}000439/2{,}7726 = 0{,}000158$; $\hat{\sigma}_P = 1{,}26\%$ pe zi, $1{,}26\% \times 15{,}87 = 20{,}0\%$ pe an
3. Informație nefolosită: deschiderea și saltul de peste noapte Yang & Zhang (2000)
4. Un drift lărgește amplitudinea și introduce o deplasare pozitivă în $\hat{\sigma}_P$ Rogers & Satchell (1991)

- O zi simulată: amplitudinea $h_t - l_t$ folosește întreaga traiectorie, randamentul close-to-close doar valorile de la capete

MFM_ch1_seminar_tests

## A8 [Rezolvat]: eficiența

5. $\mathrm{Var}\big((h-l)^2\big) = \big(9\zeta(3) - 16(\ln 2)^2\big)\sigma^4 = (10{,}819 - 7{,}687)\,\sigma^4 = 3{,}131\,\sigma^4$
   - după împărțirea la $(4\ln 2)^2 = 7{,}687$: $0{,}407\,\sigma^4$ pe zi pentru termenul Parkinson
   - close-to-close: $\mathrm{Var}(r_t^2) = 3\sigma^4 - \sigma^4 = 2\sigma^4$ pe zi

- Pentru o medie pe $n$ zile independente, ambele variațe se împart la $n$
  - raportul rămîne $2/0{,}407 = 4{,}9$
- Amplitudinea unei zile conține aproximativ informația a cinci zile de randamente close-to-close
- Interpretare: cîștigul de eficiență presupune tranzacționare continuă, fără drift și fără salturi peste noapte
  - o deschidere neactualizată (*stale open*: calculată din prețuri ale componentelor încă netranzacționate) afectează estimatorii care folosesc deschiderea
  - adică Garman–Klass, Rogers–Satchell și Yang–Zhang, nu Parkinson (B15)

# A9 [Propus]: originea autocorelației lui $r_t^2$ — cerința (1/2)

- **Model**: A4 [Rezolvat] și A6; secțiunea A9 din notebook
- **Ipoteză**: ca în A4, $\sigma_t^2 \in \{1, 4\}$ (în $\%^2$), cu probabilitățile staționare (0,8; 0,2)
  - dar starea urmează un lanț Markov cu două stări Hamilton (1989): starea de mîine depinde doar de starea de azi
  - matricea de tranziție $\begin{pmatrix} 0{,}975 & 0{,}025 \\ 0{,}1 & 0{,}9 \end{pmatrix}$ (rîndurile: starea curentă, calmă $L$ sau turbulentă $H$)
  - $p_{LL} = 0{,}975$ și $p_{HH} = 0{,}9$ sînt probabilitățile de a rămîne în aceeași stare
  - $\pi_L = 0{,}8$, $\pi_H = 0{,}2$: probabilitățile staționare, adică ponderile pe termen lung ale celor două stări
- **Cerințe**
  1. Verificați condiția de echilibru $\pi_L(1 - p_{LL}) = \pi_H(1 - p_{HH})$, adică $0{,}8 \times 0{,}025 = 0{,}2 \times 0{,}1$: fluxul din $L$ în $H$ este egal cu fluxul din $H$ în $L$.
  2. Calculați $\lambda = p_{LL} + p_{HH} - 1$, persistența lanțului: $\lambda = 0$ pentru stări independente, $\lambda \to 1$ pentru regimuri foarte persistente.

Înapoi Lista exercițiilor

# A9 [Propus]: originea autocorelației lui $r_t^2$ — cerința (2/2)

- **Cerințe** (continuare)
  3. Cu indicatorul $I_t = 1$ în starea turbulentă, $\sigma_t^2 = 1 + 3I_t$; arătați că $\mathrm{Cov}(I_t, I_{t-k}) = \pi_H(1-\pi_H)\lambda^k$.
  4. Deduceți că $\mathrm{Cov}(r_t^2, r_{t-k}^2) = \mathrm{Var}(\sigma^2)\lambda^k$ pentru $k \geq 1$.
  5. Folosind A4, calculați $\mathrm{Var}(\sigma^2)$, $\mathrm{Var}(r^2) = 3\mathbb{E}\sigma^4 - (\mathbb{E}\sigma^2)^2$ și $\mathrm{corr}(r_t^2, r_{t-k}^2)$ la lagurile 1 și 10.
  6. Reprezentați grafic ACF implicată de model, comparativ cu ACF empirică a lui $r_t^2$ pentru S&P 500 (0,31 la lagul 1, 0,04 la lagul 100).
- **Interpretare**: care sînt cele două caracteristici ale datelor pe care acest model nu le poate reproduce?

# A10 [Propus]: Cozi GARCH cu inovații Normale — cerința (1/2)

- **Model**: A4 [Rezolvat]; secțiunea A10 din notebook
  - o introducere anticipată în Capitolul 5
- **Ecuațiile modelului** GARCH(1,1) Bollerslev (1986): varianța de azi depinde de șocul de ieri și de varianța de ieri

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

  - $\mu$: randamentul mediu; $\varepsilon_t$: șocul zilei $t$; $z_t$: inovația standardizată, i.i.d. $N(0,1)$
  - $\sigma_t^2$: varianța condiționată, cunoscută la sfîrșitul zilei $t-1$
  - $\omega > 0$: nivelul de bază al variației; $\alpha \geq 0$: reacția la șocul de ieri; $\beta \geq 0$: persistența variației de ieri
  - $\alpha + \beta < 1$: varianța revine spre media ei pe termen lung, $\omega/(1-\alpha-\beta)$
  - aici $\alpha$ este un coeficient GARCH, nu indicele cozii

înapoi Lista exercițiilor

# A10 [Propus]: Cozi GARCH cu inovații Normale — cerința (2/2)

- **Rezultat**: dacă $1-(\alpha+\beta)^2-2\alpha^2>0$, kurtosisul necondiționat al lui $r_t$ este

$$K=\frac{3[1-(\alpha+\beta)^2]}{1-(\alpha+\beta)^2-2\alpha^2}>3$$

- **Cerințe**
  1. Calculați $K$ pentru $(\alpha,\beta)=(0{,}10;\ 0{,}85)$ și $(0{,}10;\ 0{,}88)$.
  2. Formulați condiția pentru un moment de ordinul patru finit.
  3. Comparați-o cu $\alpha+\beta<1$.
  4. Cu $\alpha+\beta=0{,}99$, determinați $\alpha$ pentru care modelul reproduce kurtosisul S&P 500, $K=13{,}9$.
  5. Estimați modelul pe `r_sp` în %.
     - `arch_model(100*r_sp, mean='Constant', vol='GARCH', p=1, q=1, dist='normal')`
  6. Calculați $K$ implicat.
  7. Reprezentați graficul QQ al lui $\hat{z}_t=\hat{\varepsilon}_t/\hat{\sigma}_t$.
- **Interpretare**: ce indică graficul QQ despre ipoteza distribuției Normale pentru $z_t$?
- **De raportat**: condiții pentru momente, parametri estimați, kurtosis implicat și graficul QQ.

# A11 [Propus]: estimatorul Hill ca estimator de verosimilitate maximă — cerința

- **Model**: A5 [Rezolvat]; secțiunea A11 din notebook
- **Ipoteză**: fiind dat pragul $u = X_{(k+1)}$, cele $k$ depășiri sînt statisticile de ordine ale unui eșantion Pareto i.i.d.
  - densitatea Pareto $f(x) = \alpha u^{\alpha} x^{-\alpha-1}$, $x > u$, cu indicele cozii $\alpha$
- **Noțiuni**
  - log-verosimilitatea $\ell(\alpha) = \sum_{i=1}^{k} \ln f(X_{(i)})$: cît de plauzibile sînt datele pentru o valoare dată a lui $\alpha$
  - estimatorul de verosimilitate maximă: valoarea lui $\alpha$ care maximizează $\ell(\alpha)$
  - $\ell'$, $\ell''$: derivatele de ordinul 1 și 2 în raport cu $\alpha$
  - informația Fisher $-\mathbb{E}\,\ell''(\alpha)$; $\mathrm{SE}(\hat{\alpha}) \approx 1/\sqrt{\text{informația}}$
- **Cerințe**
  1. Scrieți log-verosimilitatea $\ell(\alpha)$ a celor $k$ depășiri (ordonarea adaugă doar constanta $\ln k!$).
  2. Rezolvați $\ell'(\alpha) = 0$ și arătați că maximul este estimatorul Hill $\hat{\alpha} = k \big/ \sum_{i=1}^{k} \ln(X_{(i)}/u)$ Hill (1975).
  3. Calculați $\ell''(\alpha)$, informația $k/\alpha^2$ și $\mathrm{SE}(\hat{\alpha}) \approx \hat{\alpha}/\sqrt{k}$.
  4. Verificați SE pe datele din A5.
  5. Simulați $n = 2\,000\,000$ de valori $|X|$, $X$ Student-$t$ cu $\nu = 3$ (`numpy.random.default_rng(1)`).
  6. Reprezentați $\hat{\alpha}_k$ (estimarea Hill cu $k$ valori) în funcție de $k/n$ și raportați $k = 200$, $2\,000$, $20\,000$, $200\,000$.
- **Interpretare**: de ce se îndepărtează $\hat{\alpha}_k$ de 3 cînd $k$ este mare?

◂ Înapoi ▸ Lista exercițiilor

# B1(d) [Rezolvat]: momente pe cuantile, cu intervale bootstrap (1/2)

- **Întrebare**: se confirmă asimetria negativă și excesul de kurtosis cu măsuri care nu presupun existența momentelor?
- **Măsuri** Kim & White (2004), ambele egale cu 0 pentru distribuția Normală ($q_p$: cuantila de ordin $p$)
  - asimetria Hinkley $\frac{q_{0,95}+q_{0,05}-2q_{0,5}}{q_{0,95}-q_{0,05}}$: poziția mediei între $q_{0,05}$ și $q_{0,95}$
  - ia valori în $[-1, 1]$; negativă: coada stîngă este mai lungă decît cea dreaptă
  - kurtosisul Crow–Siddiqui $K_{CS}$ (definiție)
- **Bootstrap pe blocuri**: 1 000 de reeșantionări (`default_rng(42)`)
  - extrageți uniform $\lceil T/20 \rceil$ zile de început ($\lceil \cdot \rceil$: rotunjire în sus) și preluați cîte 20 de randamente consecutive, începînd cu fiecare
  - blocurile păstrează gruparea volatilității din interiorul lor
  - trunchiați seria la $T$ observații și recalculați ambele măsuri
  - intervalul percentil de 95%: cuantilele de 2,5% și 97,5% ale celor 1 000 de valori recalculate

Înapoi Lista exercițiilor

# B1(d) [Rezolvat]: momente pe cuantile, cu intervale bootstrap (2/2)

| | Asimetrie (momente) | Hinkley [int. 95%] | Exces kurt. | Crow–Siddiqui [int. 95%] |
|---|---|---|---|---|
| BET-TR | −1,40 | −0,036 [−0,085; 0,025] | 17,8 | 1,06 [0,73; 1,43] |
| Bitcoin | −0,70 | −0,018 [−0,069; 0,032] | 12,0 | 2,25 [1,88; 2,64] |
| EUR/RON | 0,73 | 0,041 [−0,022; 0,093] | 15,1 | 4,87 [3,74; 6,13] |

- Între paranteze drepte: intervalele bootstrap de încredere de 95%
- Toate intervalele pentru kurtosis se situează integral peste 0: cozile groase sînt confirmate fără a folosi momente
- Asimetria pe momente și asimetria Hinkley măsoară caracteristici diferite ale distribuției
  - toate intervalele Hinkley conțin 0
  - aceasta nu dovedește simetria

MFM_ch1_inference

# B7 [Propus]: un kurtosis care nu converge — cerința

- **Model**: B1 [Rezolvat]; secțiunea B7 din notebook
- **Date**: `r_sp`; fereastra $n$ cuprinde randamentele anterioare datei de 3 ianuarie a anului $1990 + n$, $n = 1, \ldots, 36$
  - 253 de randamente pentru $n = 1$, 2 527 pentru $n = 10$, 9 067 pentru $n = 36$
- **Simulare**: 500 de traiectorii Student-$t$ de lungime 9 245 pentru $\nu = 3$ și pentru $\nu = 6$ (`default_rng(42)`)
  - trunchiate la aceleași lungimi ale ferestrelor
- **Cerințe**
  1. Calculați excesul de kurtosis al S&P 500 pe fiecare fereastră.
  2. Identificați cele mai mari trei salturi și zilele care le generează.
  3. Pentru fiecare $\nu$ calculați cuantilele de 5%, 50% și 95% ale kurtosisului simulat pentru fiecare $n$.
  4. Reprezentați curba empirică suprapusă peste ambele benzi, pe o scală logaritmică simetrică.
- **Interpretare**
  - de ce se stabilizează lent banda pentru $\nu = 6$, deși kurtosisul lui $t_6$ este finit?
  - ce informație transmite, în acest caz, valoarea "exces de kurtosis = 10,9"?
- **De raportat**: kurtosis pe ferestre cumulative, benzi de cuantile simulate și cele mai mari trei salturi.

◂ Înapoi ▸ Lista exercițiilor

# B8 [Propus]: faptele stilizate reproduse de GARCH-$t$ — cerința (1/2)

- **Model**: B3 și B6 [Rezolvat], A10; secțiunea B8 din notebook
- **Date**: `r_sp` exprimat în % ($T = 9\,245$)
  - medie constantă
  - $\omega$ în $\%^2$ pe zi

```
from arch import arch_model
from arch.univariate import StudentsT
am  = arch_model(100 * r_sp, mean='Constant', vol='GARCH', p=1, q=1, dist='t')
res = am.fit(disp='off')
am.distribution = StudentsT(seed=np.random.default_rng(42))  # reproducible paths
sim = am.simulate(res.params, nobs=len(r_sp), burn=1000)['data']
```

# B8 [Propus]: faptele stilizate reproduse de GARCH-$t$ — cerința (2/2)

**Cerințe**

1. Estimați modelul și raportați toți cei cinci parametri, cu unitățile lor.
2. Verificare: simulați o traiectorie și calculați-i excesul de kurtosis.
3. Apoi simulați 200 de traiectorii.
4. Pentru fiecare traiectorie, calculați următoarele statistici.
   - excesul de kurtosis
   - ACF a lui $|r_t|$ la lagurile 1, 20, 100, 250
   - $\hat{\alpha}$ Hill pentru $|r_t|$ la $k = 2\%$ din $T$
   - $L(1) = \mathrm{corr}(r_t, |r_{t+\mathbf{1}}|)$: corelația dintre randamentul de azi și mărimea celui de mîine (efectul de levier, *leverage effect*, B9)
5. Construiți intervalele 5–95%, separat pentru fiecare statistică.
6. Precizați, pentru fiecare valoare observată, dacă se află în interiorul intervalului.

**Interpretare**

- ce fapte stilizate nu sînt reproduse?
- ce componentă a modelului ar trebui modificată pentru fiecare dintre ele?

**De raportat**: statistici observate lîngă intervalele simulate și modificări propuse modelului.

Înapoi Lista exercițiilor

# B9 [Propus]: efectul de levier, BET-TR față de Bitcoin — cerința

- **Model**: B3 [Rezolvat]; secțiunea B9 din notebook
- **Convenția lagului**: $L(k) = \text{corr}(r_t, |r_{t+k}|)$ Christie (1982)
  - $L(k) < 0$: după scăderi, randamentele ulterioare sînt mai mari în valoare absolută (efectul de levier)
  - $k$ este exprimat în zile de observație ale fiecărei serii
  - pentru $k > 0$, randamentul curent este asociat cu o valoare absolută ulterioară
  - exemplu, $k = 1$: randamentul BET-TR de luni, 16 mar. 2020, este asociat cu $|r|$ de marți, 17 mar. 2020
- **Date**: `r_bet` și `r_btc`, fiecare pe propriul calendar
- **Cerințe**
  1. Calculați $L(k)$ pentru $k = -20, \ldots, 20$.
  2. Reprezentați ambele serii cu benzile de referință i.i.d. $\pm 1{,}96/\sqrt{T}$ (0,036 și 0,030).
- **Cerințe** (continuare)
  3. Raportați $L(1)$ și $L(5)$ pentru ambele piețe.
  4. Precizați dacă se află în afara benzilor de referință.
  5. Propuneți două explicații alternative pentru un efect mai slab la Bitcoin.
- **Interpretare**
  - de ce o bandă de referință nu este un interval de încredere în prezența grupării volatilității?
  - ce informație suplimentară oferă B10?

înapoi Lista exercițiilor

# B10 [Propus]: efectul de levier ca regresie — cerința (1/2)

- **Model**: B4 [Rezolvat], B9; secțiunea B10 din notebook
- **Regresia**: $|r_{t+1}| = c + \beta^+ r_t^+ + \beta^- r_t^- + \sum_{j=1}^{5} \delta_j |r_{t-j}| + u_{t+1}$
  - $r_t^+ = \max(r_t, 0) \geq 0$, $r_t^- = \min(r_t, 0) \leq 0$: partea pozitivă și partea negativă a randamentului
  - $c$: constanta; $\delta_j$: efectul lui $|r_{t-j}|$, variabile de control pentru gruparea volatilității; $u_{t+1}$: eroarea
  - un cîștig $+a$ contribuie cu $\beta^+ a$, o pierdere $-a$ contribuie cu $-\beta^- a$
  - un răspuns simetric înseamnă $\beta^+ = -\beta^-$
  - rîndul din matricea de regresori pentru ziua $t$: $(1,\ r_t^+,\ r_t^-,\ |r_{t-1}|, \ldots, |r_{t-5}|)$
  - variabila dependentă: $|r_{t+1}|$
- **Date**: `r_sp`, `r_bet`, `r_btc`, `r_eur`

◂ Înapoi ▸ Lista exercițiilor

## B10 [Propus]: efectul de levier ca regresie — cerința (2/2)

- **Cerințe**
  1. Estimați prin OLS cu erori standard Newey–West Newey & West (1987), robuste la heteroscedasticitate și autocorelație, cu lungimea de trunchiere $\lfloor 4(T/100)^{2/9} \rfloor$ (numărul de laguri ale erorilor luate în calcul).
  2. Testați $H_0 : \beta^+ + \beta^- = 0$ (răspuns simetric) la 5%, față de $\chi^2_1$, cu statistica Wald
     - $W = (\hat{\beta}^+ + \hat{\beta}^-)^2 / [\mathrm{Var}(\hat{\beta}^+) + \mathrm{Var}(\hat{\beta}^-) + 2\,\mathrm{Cov}(\hat{\beta}^+, \hat{\beta}^-)]$
     - numitorul: varianța lui $\hat{\beta}^+ + \hat{\beta}^-$, din matricea de covarianță Newey–West; $W$ este pătratul unei statistici $t$
  3. Reprezentați valoarea estimată a lui $|r_{t+1}|$ după un cîștig $+a$ și după o pierdere $-a$, cu variabilele de control fixate la mediile lor.
  4. Arătați de ce $|r_t|$ nu poate fi adăugat printre variabilele de control.
- **De raportat**: tabelul coeficienților, test Wald, graficul răspunsului și explicația coliniarității.

# B11 [Propus]: EUR/RON pe regimuri — cerința (1/3)

- **Model**: B1, B4 [Rezolvat]; secțiunea B11 din notebook
- **Date**: `r_eur` (cursul de referință BNR), seria împărțită în trei perioade
  - iul. 2005 – 2011, 2012–2019, 2020–2026: 1 656, 2 013, 1 683 de randamente
  - lagurile se construiesc în interiorul fiecărei perioade
- **Cerințe**
  1. Pentru fiecare perioadă, estimați prin OLS regresia $r_t = a + \rho\, r_{t-1} + e_t$, cu SE robuste la heteroscedasticitate și autocorelație (HAC, *heteroskedasticity and autocorrelation consistent*), ca în B10.
     - $a$: constanta; $\rho$: autocorelația de lag 1; $e_t$: eroarea
     - lungimea de trunchiere: $\lfloor 4(T/100)^{2/9} \rfloor$
  2. Raportați $\hat{\rho}$, SE, volatilitatea și excesul de kurtosis pentru fiecare perioadă.

Înapoi Lista exercițiilor

# B11 [Propus]: EUR/RON pe regimuri — cerința (2/3)

**Cerințe** (continuare)

3. Testați egalitatea lui $\rho$ între perioade cu statistica de mai jos.

$$W = \sum_j w_j (\hat{\rho}_j - \bar{\rho})^2 \sim \chi^2_2$$

Reperul presupune estimări asimptotic independente între cele trei perioade.

- $j = 1, 2, 3$: perioada; ponderi $w_j = 1/\mathrm{SE}_j^2$, deci estimările precise contează mai mult
- $\bar{\rho} = \sum_j w_j \hat{\rho}_j / \sum_j w_j$: media ponderată a celor trei estimări
- $W$ măsoară dispersia estimărilor în jurul lui $\bar{\rho}$; 2 grade de libertate = 3 perioade − 1

# B11 [Propus]: EUR/RON pe regimuri — cerința (3/3)

- **Cerințe** (continuare)
  4. Separat, identificați o ruptură în varianță cu suma cumulată a pătratelor (CUSUM) Inclán & Tiao (1994).
     - $C_k = \sum_{t \le k} x_t^2$: suma pătratelor pînă în ziua $k$; $x_t$ este randamentul minus media întregii serii
     - $D_k = C_k / C_T - k/T$: fracțiunea din varianța totală acumulată pînă în ziua $k$, minus fracțiunea de timp scursă
     - $D_k \approx 0$ pentru orice $k$ dacă varianța este constantă
     - ruptura este estimată la $\arg\max_k |D_k|$, ziua $k$ în care $|D_k|$ este maxim
- **Interpretare**
  - de ce este atît de mare SE pentru 2020–2026?
  - de ce data rupturii identificată prin CUSUM are caracter descriptiv și nu constituie un test?

# B12 [Propus]: memorie lungă sau rupturi structurale — cerința (1/3)

- **Model**: B4, B6 [Rezolvat]; secțiunea B12 din notebook (funcția `local_whittle(x, m)`)
- **Parametrul de memorie** $d$: viteza cu care descrește ACF a lui $x_t = |r_t|$
  - $d = 0$: memorie scurtă, ACF descrește geometric
  - $0 < d < 0{,}5$: memorie lungă staționară, ACF descrește lent, ca $k^{2d-1}$
  - $d \geq 0{,}5$: seria nu mai este staționară
- **Date**: $|r_t|$ pentru S&P 500, BET-TR și Bitcoin

Înapoi Lista exercițiilor

## B12 [Propus]: memorie lungă sau rupturi structurale — cerința (2/3)

- **Metoda local Whittle** Robinson (1995): centrați $x_t = |r_t|$
  - frecvențele Fourier $\lambda_j = 2\pi j/T$, $j = 1, \ldots, m$; frecvențele mici corespund ciclurilor lungi
  - periodograma $I(\lambda) = \frac{1}{2\pi T}\left|\sum_t x_t e^{-\mathrm{i}\lambda t}\right|^2$: intensitatea ciclurilor de frecvență $\lambda$ din serie
  - memoria lungă face ca $I(\lambda_j)$ să crească proporțional cu $\lambda_j^{-2d}$ la frecvențe mici
- $\hat{d}$ minimizează pe $d \in [-0{,}49;\ 0{,}99]$

$$\ln\left(\frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2d} I(\lambda_j)\right) - \frac{2d}{m}\sum_{j=1}^{m}\ln\lambda_j$$

  - $m = \lfloor T^p \rfloor$: numărul de frecvențe mici folosite; $p \in (0,1)$ stabilește lățimea de bandă
  - $\mathrm{SE} = 1/(2\sqrt{m})$

# B12 [Propus]: memorie lungă sau rupturi structurale — cerința (3/3)

**Cerințe**

1. Estimați $d$ cu $p = 0{,}65$ și raportați intervalul de încredere de 95%.
2. Reprezentați $\hat{d}$ în funcție de $p$ de la 0,45 la 0,80, cu dreptele $d = 0$ și $d = 0{,}5$.
3. Identificați o ruptură în varianța lui $r_t$ (B11).
4. Reestimați $\hat{d}$ de fiecare parte a rupturii.
5. Reprezentați ACF a lui $|r_t|$ înainte și după ruptură.

**Interpretare**

- este $\hat{d} > 0$ robust la $p$ și la ruptură?
- este rezultatul compatibil și cu un model cu regimuri Diebold & Inoue (2001)?
- ce implică $\hat{d} \geq 0{,}5$?

**De raportat**: estimări ale parametrului de memorie și intervale, graficul lui $\hat{d}$ în funcție de $p$, data rupturii și ACF înainte/după.

# B13 [Propus]: BET-TR față de S&P 500 pe aceeași perioadă — cerința (1/2)

- **Model**: A1, B1 [Rezolvat]; secțiunea B13 din notebook
- **Date**: BET-TR (randament total, lei) și S&P 500 (indice de preț, USD)
  - aliniați *prețurile* pe zilele comune (23 sept. 2014 – 18 sept. 2026)
  - apoi calculați cele 2 903 randamente
  - lungimea intervalelor: 2 231 de o zi, 538 de două sau trei zile, 134 mai lungi (cel mult 6 zile)

◂ Înapoi ▸ Lista exercițiilor

# B13 [Propus]: BET-TR față de S&P 500 pe aceeași perioadă — cerința (2/2)

- **Cerințe**
  1. Reprezentați valoarea investiției, rebazată (scală logaritmică), și cele două curbe de drawdown („sub apă"), $DD_t$ din A2(a).
  2. Calculați următoarele măsuri pentru ambii indici.
     - CAGR $= (P_{\text{final}}/P_{\text{inițial}})^{1/Y} - 1$, cu $Y$ numărul de ani calendaristici; volatilitatea randamentelor simple ($A = 252$, prin convenție; pe zilele comune sînt circa 242 de randamente pe an)
     - raportul Sharpe $252\bar{R}/(\sqrt{252}\,\hat{\sigma}_R)$ cu $r_f = 0$: randamentul anual pe unitatea de volatilitate
     - $\bar{R}$, $\hat{\sigma}_R$: media și abaterea standard a randamentelor simple zilnice; $r_f$: rata dobînzii fără risc
     - raportul Sortino Sortino & van der Meer (1991) $252\bar{R}/(\sqrt{252}\,\hat{\sigma}_D)$, cu semiabaterea inferioară $\hat{\sigma}_D = \sqrt{\frac{1}{T}\sum_t \min(R_t, 0)^2}$, calculată pe toate zilele
     - Sortino penalizează doar zilele cu pierderi
     - MDD și raportul Calmar (A2(a))
- **De raportat**: graficul și un tabel cu două coloane
- **Interpretare**: indicați trei motive pentru care comparația este doar descriptivă

# B14 [Propus]: semnificația diferenței dintre rapoartele Sharpe — cerința (1/2)

- **Model**: B1(d), B13; secțiunea B14 din notebook
- **Referință**: eroarea standard a raportului Sharpe, Lo (2002)
- **Date**: cele 2 903 perechi de randamente din B13
  - randamente simple pentru rapoartele Sharpe (calculate în B13) și corelația din test
  - randamente logaritmice pentru comparațiile de calendar de la cerința 5
- **Cerințe**
  1. Convertiți rapoartele Sharpe în valori zilnice ($\widehat{SR}/\sqrt{252}$).
  2. Calculați testul Jobson–Korkie Jobson & Korkie (1981) cu corecția lui Memmel (2003).
     - $z = (\widehat{SR}_1 - \widehat{SR}_2)/\sqrt{\hat{V}}$
     - $\hat{V} = \frac{1}{T}\left[2(1-\hat{\rho}) + \frac{1}{2}(\widehat{SR}_1^2 + \widehat{SR}_2^2 - 2\widehat{SR}_1\widehat{SR}_2\hat{\rho}^2)\right]$
     - $\widehat{SR}_1, \widehat{SR}_2$: rapoartele Sharpe zilnice ale BET-TR și S&P 500; $\hat{\rho}$: corelația randamentelor lor
     - $\hat{V}$: varianța estimată a diferenței; sub $H_0$ (rapoarte Sharpe egale) $z \approx N(0,1)$

Înapoi Lista exercițiilor

# B14 [Propus]: semnificația diferenței dintre rapoartele Sharpe — cerința (2/2)

- **Cerințe** (continuare)
  3. Aplicați un bootstrap circular pe blocuri, pe perechi de randamente, cu 2 000 de reeșantionări (`default_rng(42)`).
     - extrageți $\lceil T/21 \rceil$ zile de început, preluați cîte 21 de zile consecutive (circular, reluînd seria de la început), trunchiați la $T$
     - aplicați **aceiași** indici ambelor serii
  4. Raportați intervalul percentil de 95% pentru diferența anualizată.
  5. Calculați patru corelații.
     - în aceeași zi
     - BET-TR cu S&P 500 din ziua precedentă și din ziua următoare
     - pe randamente săptămînale (vineri)
- **Interpretare**
  - este diferența semnificativă?
  - ce presupun cele două metode?
  - ce corelație trebuie folosită la construcția unui portofoliu?

## B15 [Propus]: estimatori ai volatilității din amplitudinea zilnică, S&P 500 — cerința (1/2)

- **Model**: A8 [Rezolvat]; secțiunea B15 din notebook
  - notebook-ul conține funcțiile pentru toți cei cinci estimatori
  - formulele se găsesc în Capitolul 1
- **Date**: OHLC pentru indicele S&P 500, 2 ian. 2008 – 18 sept. 2026
- ETF-ul SPY, cu toate prețurile ajustate pe aceeași bază
  - $C^*$ = închiderea ajustată (pentru dividende și splituri)
  - $O^*, H^*, L^* = (C^*/C) \times O, H, L$, cu $C$ închiderea neajustată și $C^*/C$ factorul de ajustare al zilei
  - exemplu, 16 mar. 2020: factorul $218{,}08/239{,}85 = 0{,}9092$; $O^* = 241{,}18 \times 0{,}9092 = 219{,}29$, $H^* = 233{,}58$, $L^* = 215{,}82$

## B15 [Propus]: estimatori ai volatilității din amplitudinea zilnică, S&P 500 — cerința (2/2)

- **Cerințe**
  1. Calculați volatilitățile pe ferestre mobile (*rolling*) de 21 de zile: close-to-close, Parkinson, Garman–Klass Garman & Klass (1980), Rogers–Satchell Rogers & Satchell (1991) și Yang–Zhang Yang & Zhang (2000) ($A = 252$).
  2. Calculați mediile lor în timp.
  3. Pentru indice și pentru SPY, comparați mediile estimatorilor close-to-close și Yang–Zhang pe 21 de zile cu două volatilități calculate pe întregul eșantion.
     - volatilitatea de peste noapte: $\mathrm{sd}(o_t - c_{t-\mathbf{1}})\sqrt{252}$, cu $o_t - c_{t-\mathbf{1}}$ randamentul logaritmic de peste noapte (notațiile din A8)
     - volatilitatea de la deschidere la închidere: $\mathrm{sd}(c_t - o_t)\sqrt{252}$
     - sd: abaterea standard de selecție
- **Interpretare**: de ce rămîne Yang–Zhang sub close-to-close pe indice, dar nu și pe SPY?
- **De raportat**: tabel cu volatilitățile medii și explicația contribuției de peste noapte.

# C2 [Propus]: autocorelația BET-TR după reclasificarea FTSE (1/3)

- **Model**: B4 [Rezolvat], B11; o propunere de proiect
- **Context**: pe piețele mai puțin lichide, randamentele indicilor au adesea o autocorelație pozitivă la lagul 1, din cauza prețurilor neactualizate ale componentelor (*stale prices*)
- **Ipoteza de testat**: după reclasificare, lichiditatea a crescut, iar autocorelația de lag 1 a BET-TR a scăzut
  - FTSE (*Financial Times Stock Exchange*, furnizorul de indici FTSE Russell) Russell, furnizor internațional de indici bursieri, a reclasificat România ca piață emergentă secundară, începînd cu 21 sept. 2020
- **Date**: `r_bet` și o măsură a lichidității
  - lichiditatea: valoarea zilnică tranzacționată (volum $\times$ preț de închidere brut, în lei) a 11 acțiuni de primă mînă (*blue chips*) de la BVB
  - acțiunile, după simbolul bursier: TLV (Banca Transilvania), SNP (OMV Petrom), BRD (BRD–Groupe Société Générale), TGN (Transgaz), FP, SNG (Romgaz), H2O (Hidroelectrica), SNN (Nuclearelectrica), EL (Electrica), DIGI, TEL (Transelectrica); fișierele `<simbol>.RO.csv`, cîmpurile `volume` și `close`
  - însumată pe acțiunile listate în ziua respectivă
  - raportați numărul de acțiuni listate în fiecare an

înapoi Lista exercițiilor

# C2 [Propus]: autocorelația BET-TR după reclasificarea FTSE (2/3)

**Cerințe**

1. Calculați $\hat{\rho}_1$ pe ferestre mobile de 250 de zile (pas de 5 zile).
2. Adăugați banda de test de 5% $\pm 1{,}96\,\mathrm{SE}$ în jurul lui zero, cu SE robustă a fiecărei ferestre calculată ca în B4 sub $\rho_1 = 0$; ea testează $H_0 : \rho_1 = 0$ și nu este un interval de încredere în jurul lui $\hat{\rho}_1$.
3. Estimați $\rho_1$ înainte și după 21 sept. 2020 ca pantă AR(1), cu SE HAC (B11).
4. Raportați statistica pentru diferența celor două estimări.

$$z = \frac{\hat{\rho}_{\text{după}} - \hat{\rho}_{\text{înainte}}}{\sqrt{\mathrm{SE}_1^2 + \mathrm{SE}_2^2}}$$

- $\mathrm{SE}_1, \mathrm{SE}_2$: erorile standard HAC înainte și după
- sub $H_0$ (nicio schimbare) $z \approx N(0,1)$, dacă cele două perioade sînt independente

# C2 [Propus]: autocorelația BET-TR după reclasificarea FTSE (3/3)

- **Cerințe** (continuare)
  5. Raportați interacțiunea $e$ din regresia de mai jos (HAC, 10 laguri).

$$r_t = a + b\, r_{t-1} + c\, \text{post}_t + e\, r_{t-1}\text{post}_t$$

     - $\text{post}_t = 1$ după 21 sept. 2020 și 0 înainte
     - $b$: autocorelația înainte; $e$: modificarea ei după reclasificare; $a$, $c$: constanta și modificarea ei
  6. Calculați corelația dintre $\hat{\rho}_1$ pe ferestre mobile și logaritmul valorii medii tranzacționate pe 250 de zile.
- **Livrabil**: un folder Quantlet și un slide
  - o concluzie descriptivă și principala ei limită
  - nu o afirmație cauzală

# C3 [Propus]: dependența VIX de traiectoria prețurilor — cerința (1/5)

- **Model**: studiul de caz din curs
- **Întrebare**: pot randamentele trecute ale S&P 500 să explice, singure, VIX, așa cum arată Guyon & Lekeufack (2023)?
  - 94,6% din varianța sa în 2000–2018, 85,5% în 2019 – mai 2022
- **Date**: închiderile zilnice ale S&P 500, VIX și VIX pe 9 zile
  - $S_{t_i}$: nivelul S&P 500 la închiderea zilei $t_i$; randamente simple $r_{t_i} = S_{t_i}/S_{t_{i-1}} - 1$
  - indicii de volatilitate împărțiți la 100
- **Modelul** (Guyon & Lekeufack (2023), ec. (3.7)): volatilitatea de azi este o funcție liniară de două rezumate ale randamentelor trecute

$$\sigma_t = \beta_0 + \beta_1 R_{1,t} + \beta_2 \Sigma_t$$

  - $\sigma_t$: VIX (sau VIX pe 9 zile) la închiderea zilei $t$, împărțit la 100
  - $\beta_0, \beta_1, \beta_2$: coeficienți estimați din date

Înapoi Lista exercițiilor

## C3 [Propus]: dependența VIX de traiectoria prețurilor — cerința (2/5)

- **Componenta de trend**: o sumă ponderată a randamentelor din ultimele 1 000 de zile

$$R_{1,t} = \sum_{i=0}^{999} K_1(i\Delta t)\, r_{t-i}$$

  - $i$: vechimea randamentului, în zile; $\Delta t = 1/252$: o zi, exprimată în ani; $K_1(i\Delta t)$: ponderea lui
  - $R_{1,t} < 0$ după o perioadă de scăderi, $R_{1,t} > 0$ după o perioadă de creșteri

- **Componenta de volatilitate**: rădăcina pătrată a unei sume ponderate a randamentelor la pătrat

$$\Sigma_t = \Big( \sum_{i=0}^{999} K_2(i\Delta t)\, r_{t-i}^2 \Big)^{1/2}$$

  - o volatilitate istorică ponderată, cu ponderile $K_2$

- **Interpretarea coeficienților**
  - $\beta_1 < 0$: după scăderi ($R_{1,t} < 0$) volatilitatea crește, adică efectul de levier
  - $\beta_2 > 0$: o volatilitate recentă mare ridică volatilitatea de azi, adică gruparea volatilității

## C3 [Propus]: dependența VIX de traiectoria prețurilor — cerința (3/5)

- **Nuclee**, ec. (3.9) din lucrare: o lege de putere decalată în timp

$$K_j(\tau) = \frac{(\tau + \delta_j)^{-\alpha_j}}{Z_j}, \qquad j = 1, 2$$

  - $\tau$: vechimea, în ani; $Z_j$ este ales astfel încît $\sum_{i=0}^{999} K_j(i\Delta t)\Delta t = 1$
  - $\alpha_j > 1$: viteza de scădere a ponderilor (nu indicele cozii); un $\alpha_j$ mic înseamnă ponderi care scad lent în fereastra de 1 000 de zile
  - $\delta_j > 0$: lagul; un $\delta_j$ mic concentrează ponderea pe ultimele cîteva zile
- **Parametri și unități**
  - $\beta_0$ în unități de volatilitate (0,03 = 3 puncte VIX)
  - $\beta_1, \beta_2$ adimensionali
  - $\delta_j$ în ani
- **Valori inițiale** (Tabelul 3 din lucrare, rîndul VIX)
  - $\beta_0 = 0{,}057$, $\beta_1 = -0{,}095$, $\beta_2 = 0{,}82$
  - $\alpha_1 = 1{,}06$, $\delta_1 = 0{,}020$, $\alpha_2 = 1{,}60$, $\delta_2 = 0{,}052$
  - limitele $\alpha_j \in [1{,}0001;\ 20]$, $\delta_j \in [10^{-4};\ 5]$

# C3 [Propus]: dependența VIX de traiectoria prețurilor — cerința (4/5)

- **Cerințe**, pe etape
  1. Nuclee fixe: calculați $R_{1,t}$ și $\Sigma_t$ cu nucleele din Tabelul 3.
  2. Estimați $\beta_0, \beta_1, \beta_2$ prin OLS pe 2000–2018.
  3. Modelul complet: estimați toți cei șapte parametri prin cele mai mici pătrate neliniare pe 1 ian. 2000 – 31 dec. 2018, minimizînd $\sum_t (\sigma_t - \hat{\sigma}_t)^2$, cu $\hat{\sigma}_t$ valoarea dată de model.
     - folosiți `scipy.optimize.least_squares`
     - recalculați variabilele la fiecare evaluare a parametrilor
  4. Raportați $R^2$ și rădăcina erorii pătratice medii (RMSE, definite pe slide-ul următor) pe eșantionul de antrenare și pe perioada 1 ian. 2019 – 15 mai 2022.
     - pentru modelul complet și pentru restricțiile $\beta_1 = 0$ și $\beta_2 = 0$; reestimați parametrii rămași prin cele mai mici pătrate neliniare; nucleul componentei eliminate nu mai intervine
  5. Repetați modelul complet pentru VIX pe 9 zile (antrenare din 2011).
  6. Cu parametrii menținuți ficși, evaluați modelul pe perioada 16 mai 2022 – 18 sept. 2026.

## C3 [Propus]: dependența VIX de traiectoria prețurilor — cerința (5/5)

- **Criteriile de evaluare**, calculate separat pe fiecare eșantion de $T$ zile
  - RMSE $= \sqrt{\frac{1}{T}\sum_t(\sigma_t - \hat{\sigma}_t)^2}$: eroarea tipică, în unități de volatilitate
  - $R^2 = 1 - \sum_t(\sigma_t - \hat{\sigma}_t)^2 / \sum_t(\sigma_t - \bar{\sigma})^2$: proporția din variația lui $\sigma_t$ explicată de model
  - $\bar{\sigma}$: media lui $\sigma_t$; $R^2 = 1$: potrivire perfectă; $R^2 \leq 0$: nu mai bun decît media
- **De raportat**
  - tabelul parametrilor și scorurile, comparate cu Tabelul 2 din lucrare
  - graficul VIX observat și estimat și raportul dintre acestea
  - nucleele estimate
- **Interpretare**: modelul explică VIX din aceeași zi
  - nu este o prognoză pentru ziua următoare
  - ce variabilă explică mai mult, $R_{1,t}$ sau $\Sigma_t$?

## C3: ce schimbă normalizarea ponderilor?

- Ponderile care însumează unu sînt $w_{j,i} = K_j(i\Delta t)\Delta t$, cu $\Delta t = 1/252$.
- Formulele de la cerința 1 folosesc $K_j$, fără factorul $\Delta t$: componenta de trend este de 252 de ori media calculată cu ponderile $w_{1,i}$.
- Componenta de volatilitate este de $\sqrt{252}$ ori rădăcina mediei ponderate a randamentelor la pătrat, calculată cu $w_{2,i}$.
- Exemplu: dacă toate randamentele sînt 0,1%, media ponderată este 0,001, dar componenta de trend este 0,252.
- Coeficienții modelului depind de această scară: folosiți aceeași normalizare la estimare și pe perioada de test.

# Bibliografie (1/3)


- **Bollerslev, T. (1986).** Generalized autoregressive conditional heteroskedasticity. ***Journal of Econometrics***, **31(3), 307–327.**
- **Christie, A.A. (1982).** The stochastic behavior of common stock variances: value, leverage and interest rate effects. ***Journal of Financial Economics***, **10(4), 407–432.**
- **Clark, P.K. (1973).** A subordinated stochastic process model with finite variance for speculative prices. ***Econometrica***, **41(1), 135–155.**
- **Cont, R. (2001).** Empirical properties of asset returns: stylized facts and statistical issues. ***Quantitative Finance***, **1(2), 223–236.**
- **Dalla, V., Giraitis, L., Phillips, P.C.B. (2022).** Robust tests for white noise and cross-correlation. ***Econometric Theory***, **38(5), 913–941.**
- **Diebold, F.X., Inoue, A. (2001).** Long memory and regime switching. ***Journal of Econometrics***, **105(1), 131–159.**
- **Dimson, E. (1979).** Risk measurement when shares are subject to infrequent trading. ***Journal of Financial Economics***, **7(2), 197–226.**
- **Engle, R.F. (1982).** Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. ***Econometrica***, **50(4), 987–1007.**
- **Garman, M.B., Klass, M.J. (1980).** On the estimation of security price volatilities from historical data. ***Journal of Business***, **53(1), 67–78.**
- **Guyon, J., Lekeufack, J. (2023).** Volatility is (mostly) path-dependent. ***Quantitative Finance***, **23(9), 1221–1258.**

# Bibliografie (2/3)


- **Hamilton, J.D. (1989).** A new approach to the economic analysis of nonstationary time series and the business cycle. ***Econometrica*, 57(2), 357–384.**
- **Hill, B.M. (1975).** A simple general approach to inference about the tail of a distribution. ***Annals of Statistics*, 3(5), 1163–1174.**
- **Inclán, C., Tiao, G.C. (1994).** Use of cumulative sums of squares for retrospective detection of changes of variance. ***Journal of the American Statistical Association*, 89(427), 913–923.**
- **Jarque, C.M., Bera, A.K. (1980).** Efficient tests for normality, homoscedasticity and serial independence of regression residuals. ***Economics Letters*, 6(3), 255–259.**
- **Jobson, J.D., Korkie, B.M. (1981).** Performance hypothesis testing with the Sharpe and Treynor measures. ***Journal of Finance*, 36(4), 889–908.**
- **Kim, T.-H., White, H. (2004).** On more robust estimation of skewness and kurtosis. ***Finance Research Letters*, 1(1), 56–73.**
- **Ledoit, O., Wolf, M. (2008).** Robust performance hypothesis testing with the Sharpe ratio. ***Journal of Empirical Finance*, 15(5), 850–859.**
- **Ljung, G.M., Box, G.E.P. (1978).** On a measure of lack of fit in time series models. ***Biometrika*, 65(2), 297–303.**
- **Lo, A.W. (2002).** The statistics of Sharpe ratios. ***Financial Analysts Journal*, 58(4), 36–52.**
- **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.**

# Bibliografie (3/3)


- Magdon-Ismail, M., Atiya, A.F., Pratap, A., Abu-Mostafa, Y.S. (2004). On the maximum drawdown of a Brownian motion. *Journal of Applied Probability*, 41(1), 147–161.
- Memmel, C. (2003). Performance hypothesis testing with the Sharpe ratio. *Finance Letters*, 1(1), 21–23.
- Newey, W.K., West, K.D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703–708.
- Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 53(1), 61–65.
- Robinson, P.M. (1995). Gaussian semiparametric estimation of long range dependence. *Annals of Statistics*, 23(5), 1630–1661.
- Rogers, L.C.G., Satchell, S.E. (1991). Estimating variance from high, low and closing prices. *Annals of Applied Probability*, 1(4), 504–512.
- Sortino, F.A., van der Meer, R. (1991). Downside risk. *Journal of Portfolio Management*, 17(4), 27–31.
- Yang, D., Zhang, Q. (2000). Drift-independent volatility estimation based on high, low, open, and close prices. *Journal of Business*, 73(3), 477–492.

# Acronime folosite în acest material (1/2)

- **ACF**: AutoCorrelation Function — *funcția de autocorelație*
- **AI**: Artificial Intelligence — *inteligență artificială*
- **AR**: AutoRegressive model — model autoregresiv
- **ARCH-LM**: ARCH Lagrange Multiplier (test) — *testul multiplicatorului Lagrange pentru efecte ARCH*
- **BET-TR**: BET Total Return (index) — *indicele BET cu dividendele reinvestite*
- **BNR**: Banca Națională a României
- **BVB**: Bursa de Valori București
- **CAGR**: Compound Annual Growth Rate — *rata anuală compusă de creștere*
- **CI**: Confidence Interval — *interval de încredere*
- **COVID-19**: COronaVIrus Disease 2019 — *boala provocată de coronavirus, 2019*
- **CUSUM**: CUmulative SUM (of squares) — *suma cumulată (a pătratelor)*
- **EGARCH**: Exponential GARCH — *GARCH exponențial*
- **ETF**: Exchange-Traded Fund — *fond tranzacționat la bursă*
- **FIGARCH**: Fractionally Integrated GARCH — *GARCH integrat fracționar*
- **FTSE**: Financial Times Stock Exchange (FTSE Russell, index provider) — *furnizorul de indici FTSE Russell*
- **GARCH**: Generalised AutoRegressive Conditional Heteroskedasticity — *heteroscedasticitate condiționată autoregresivă generalizată*

# Acronime folosite în acest material (2/2)

- **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*
- **HAR**: Heterogeneous AutoRegressive (model) — *model autoregresiv heterogen*
- **JB**: Jarque–Bera (test) — *testul de normalitate Jarque–Bera*
- **MDD**: Maximum DrawDown — *drawdown-ul maxim*
- **OHLC**: Open, High, Low, Close — *deschidere, maxim, minim, închidere*
- **OLS**: Ordinary Least Squares — *metoda celor mai mici pătrate*
- **QQ**: Quantile–Quantile (plot) — *graficul cuantilă–cuantilă*
- **RMSE**: Root Mean Squared Error — *rădăcina erorii pătratice medii*
- **SE**: Standard Error — *eroare standard*
- **SUA**: Statele Unite ale Americii
- **UTC**: Coordinated Universal Time (Temps Universel Coordonné) — *timpul universal coordonat*
- **VIX**: Cboe Volatility Index — *indicele volatilității implicite a S&P 500*