








Statistica piețelor financiare

Capitolul 1: Date, randamente și indicatori



Daniel Traian PELE

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

Acronime folosite în acest material (1/2)

- **ADF**: Augmented Dickey–Fuller (test) — testul Dickey–Fuller augmentat
- **AI**: Artificial Intelligence (inteligență artificială)
- **BCE**: Banca Centrală Europeană
- **BET**: Bucharest Exchange Trading (index) — indicele de referință al Bursei de Valori București
- **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)
- **COVID-19**: COronaVIrus Disease 2019 (boala provocată de noul coronavirus, identificată în 2019)
- **CRSP**: Center for Research in Security Prices (University of Chicago) — centrul de cercetare a prețurilor valorilor mobiliare, Universitatea din Chicago
- **EOD**: End Of Day (daily closing data) — date de sfîrșit de zi, adică prețuri de închidere
- **EODHD**: EOD Historical Data (market-data provider) — furnizor de date de piață
- **ETF**: Exchange-Traded Fund (fond tranzacționat la bursă)

Acronime folosite în acest material (2/2)

- **FRED**: Federal Reserve Economic Data (baza de date economice a Rezervei Federale din St. Louis)
- **GARCH**: Generalised AutoRegressive Conditional Heteroskedasticity (heteroscedasticitate condiționată autoregresivă generalizată)
- **LSEG**: London Stock Exchange Group (data vendor, formerly Refinitiv) — grupul London Stock Exchange, furnizor de date, fost Refinitiv
- **MDD**: Maximum DrawDown (drawdown-ul maxim)
- **NYSE**: New York Stock Exchange (Bursa din New York)
- **OHLCV**: Open, High, Low, Close, Volume (deschidere, maxim, minim, închidere, volum)
- **OMV**: Österreichische Mineralölverwaltung (Administrația austriacă a uleiurilor minerale, grupul OMV)
- **OUG**: Ordonanță de Urgență a Guvernului
- **PIB**: Produsul intern brut
- **SE**: Standard Error (eroare standard)
- **SFE**: Statistics of Financial Markets, Exercises (Quantlet collection of the textbook) — colecția de Quantlet-uri a manualului Statistics of Financial Markets
- **SR**: Sharpe Ratio (raportul Sharpe)
- **SUA**: Statele Unite ale Americii

Întrebarea de azi, traseul și rezultatele învățării

- **Întrebarea**: cum transformăm datele brute de piață în cifre care permit compararea investițiilor?
  - un nivel de preț spune puțin; un randament, o măsură a riscului și un drawdown sînt mult mai informative
- **Traseul și rezultatele învățării**: la finalul capitolului veți putea
  - alege o sursă de date de încredere și verifica o serie de prețuri înainte de utilizare
  - calcula randamente simple și logaritmice și le agrega în timp și între active
  - anualiza mediile și volatilitățile cu frecvența reală a datelor
  - descrie o serie de randamente prin momente, cuantile și valori extreme
  - calcula și interpreta CAGR, Sharpe, Sortino, drawdown-ul maxim, Calmar și volatility drag
  - compara diferența dintre doi indicatori cu eroarea lor de estimare
- Date reale în tot capitolul: BET, S&P 500, DAX, Bitcoin și cinci acțiuni de la BVB

Bibliografie și instrumente

- Manual: Franke, Härdle & Hafner (2019), *Statistics of Financial Markets*, ed. a 5-a, cap. 11 (randamente, definiții)
  - exerciții rezolvate: Borak, Härdle & López-Cabrera (2013)
- Lecturi complementare: Tsay (2010), cap. 1 (randamentele activelor); Campbell, Lo & MacKinlay (1997), cap. 1
- Quantlet-urile Python ale capitolului: Quantlets/Ch_01
  - adaptate după Quantlet-urile SFE ale manualului (SFEReturns, SFEcomplogreturns, SFEportlogreturns, SFEfdescStat)
  - fiecare grafic din aceste slide-uri are un link către Quantlet-ul care îl generează
- Notebook-ul cursului: deschideți în Google Colab
- Curs video: Statistics of Financial Markets

Tipuri de date financiare

- **Date de piață**: informațiile produse de un loc de tranzacționare
  - prețuri de tranzacționare, cele mai bune cotații bid și ask, volum tranzacționat
  - niveluri de indici, cursuri de schimb, rate ale dobînzii, prețuri cripto
- **Evenimente corporative**: evenimente care schimbă prețul fără a schimba valoarea deținerilor
  - dividende, splituri, consolidări de acțiuni, acțiuni gratuite
- **Date fundamentale și macroeconomice**
  - bilanțuri și profituri; inflație, PIB, dobînzi de politică monetară
- **Date alternative**
  - știri, rețele sociale, imagini din satelit, tranzacții pe blockchain
- Cursul folosește în principal **date zilnice de piață**

Banda telegrafică și fluxurile electronice de date

- Din 1867, **telegrafele bursiere** (stock tickers) tipăreau pe o bandă de hîrtie prețurile tranzacțiilor, primite prin telegraf
  - primul flux de prețuri în timp real: prețul ultimei tranzacții, la cîteva minute după ce a avut loc
- Azi: fluxuri electronice de la burse, redistribuite de furnizorii de date
  - fiecare cotație și tranzacție, cu marcaj de timp la microsecundă
  - majoritatea studiilor folosesc rezumatele zilnice (OHLCV)
- Banda de cotații a rămas: prețuri și știri care defilează pe fațadele clădirilor; simbolurile bursiere (tickere), ca TLV, SNP, H2O la BVB



Bandă de știri pe clădirea Reuters, New York

Foto: bgilliard (2005); CC BY-SA 2.0; Wikimedia Commons

Frecvența datelor

- **Date tick**: fiecare tranzacție și cotație
  - milioane de înregistrări pe zi; zgomot cauzat de alternanța tranzacțiilor între bid și ask (bid–ask bounce)
- **Bare intrazilnice**: 1, 5 sau 60 de minute
  - folosite pentru volatilitatea realizată (Capitolul 8)
- **Date zilnice**: un rînd pe zi de tranzacționare
  - standardul acestui curs și al majorității studiilor academice
  - bursele funcționează aproximativ 250 de zile pe an; piețele cripto, 365 de zile pe an
- **Date săptămînale, lunare, anuale**
  - mai puține observații, mai aproape de orizontul investitorilor pe termen lung
- Regulă: anualizați cu numărul **real** de observații pe an al fiecărei serii

Surse de date

Sursa	Date disponibile	Acces
Burse (BVB, Deutsche Börse, NYSE)	prețuri oficiale, volume, niveluri ale indicilor	site, fluxuri contra cost
Furnizori de date (Bloomberg, LSEG, EODHD)	prețuri de pe multe piețe, prețuri ajustate, date fundamentale	abonament
Bănci centrale (BNR, BCE, Rezerva Federală)	cursuri de referință, rate ale dobînzii	gratuit
Baze de date statistice (FRED, Eurostat)	serii macroeconomice, randamente ale titlurilor de stat	gratuit
Baze de date academice (CRSP, biblioteca Kenneth French)	randamente curățate, inclusiv firme delistate; randamente ale factorilor	licență universitară sau gratuit
Burse și agregatoare cripto	prețuri 24/7, date on-chain	gratuit sau abonament

- Furnizorul colectează și curăță datele; **sursa primară** a cifrei (bursa, banca centrală) este referința atunci cînd două surse diferă

Datele acestui curs

- Date zilnice de piață de la **EODHD** (EOD Historical Data), salvate o singură dată în repository-ul cursului
  - 91 de serii: indici, ETF-uri, acțiuni (BVB, SUA, Europa), cripto, cursuri de schimb, randamente ale obligațiunilor
  - coloane: data, deschidere, maxim, minim, închidere, închidere ajustată, volum
- EUR/RON: cursul oficial de referință al **BNR**
  - seria EUR/RON de la EODHD are cotații eronate (Erori în date: două serii EUR/RON)
- Convenții
  - indici, cursuri de schimb, cripto: *close*; acțiuni și ETF-uri: *adjusted close*
  - cotațiile de weekend și închiderile repetate în zilele libere se elimină, cu excepția activelor cripto
  - fiecare serie se analizează pe propriul calendar
- Perioada de comparație a capitolului: 1 ianuarie 2015 – 18 septembrie 2026

Barele zilnice: OHLCV

- **OHLCV**: prețul de deschidere O_t, maxim H_t, minim L_t, de închidere C_t și volumul V_t din ziua t
- Prin construcție $L_t \le O_t, C_t \le H_t$; amplitudinea $H_t - L_t$ măsoară variația prețurilor în timpul zilei
- Banca Transilvania (TLV), ultimele cinci zile de tranzacționare din datele noastre:

Data	Open	High	Low	Close	Volum
14 sep. 2026	34,52	34,86	34,28	34,54	346 819
15 sep. 2026	34,40	34,68	33,10	34,00	676 653
16 sep. 2026	34,00	34,10	33,20	33,40	853 078
17 sep. 2026	33,40	33,40	33,40	33,40	1 084 380
18 sep. 2026	34,22	35,80	34,22	35,50	1 884 988

- Verificare: pe 17 septembrie 2026 deschidere = maxim = minim = închidere, cu 1 084 380 de acțiuni tranzacționate; o bară plată cu volum mare este suspectă

Evenimente corporative și prețul ajustat

- Prețul de închidere P_t face un salt în ziua unui eveniment corporativ; valoarea deținerii nu se modifică
- **Factorul de ajustare** F în ziua evenimentului
  - dividend în numerar D: $F = 1 - D/P_{t-1}$
  - split m-la-n (m acțiuni noi pentru n vechi): $F = n/m$; o consolidare este un split cu $m < n$
- **Prețul ajustat**: fiecare preț anterior se înmulțește cu factorii tuturor evenimentelor ulterioare, $P_t^{\text{adj}} = P_t \prod_{k:\, t_k > t} F_{t_k}$
  - t_k: ziua evenimentului k; produsul parcurge toate evenimentele de după ziua t, deci ultimele prețuri rămîn neschimbate
  - randamentele calculate din prețuri ajustate includ dividendul și nu au salturi artificiale
- Cu dividende, randamentul simplu este $R_t = (P_t + D_t)/P_{t-1} - 1$, unde D_t este dividendul plătit în ziua t Franke, Härdle & Hafner (2019)

Exemplu rezolvat: prețuri ajustate

- O acțiune cotează la 100; în ziua 5 plătește un dividend de 2; în ziua 10 are loc un split 2-la-1
- Factorii de ajustare
  - dividend: $F_5 = 1 - 2/100 = 0{,}98$; split: $F_{10} = 1/2 = 0{,}50$
- Prețurile ajustate
  - înainte de ziua 5: $P_t^{\text{adj}} = P_t \times 0{,}49$, deci 100 devine 49
  - între zilele 5 și 9: $P_t^{\text{adj}} = P_t \times 0{,}50$; din ziua 10: $P_t^{\text{adj}} = P_t$
- Seria ajustată se schimbă la fiecare eveniment nou: păstrați întotdeauna data descărcării
- Indici: un **indice de preț** ignoră dividendele; un **indice de randament total** le reinvestește

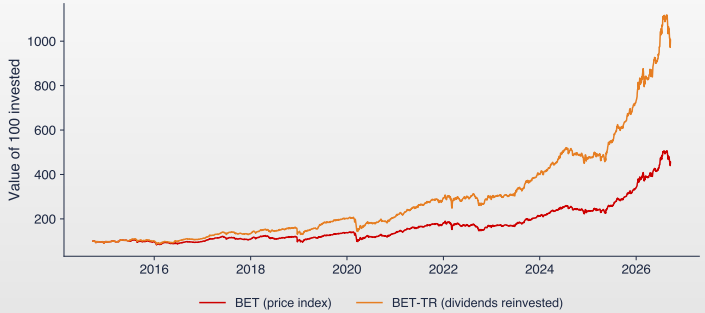
Date reale: Banca Transilvania, prețul de închidere și prețul ajustat



- Pe 16 august 2022 prețul crește brusc de la 2,12 la 21,20 lei: o consolidare a acțiunilor, 10 acțiuni vechi într-una
- Randamentul logaritmic din prețul de închidere: +230% într-o zi; din prețul ajustat: 0,0%
- Din 2015, randamentele zilnice ale celor două serii diferă cu peste 2 puncte procentuale în 24 de zile (dividende, acțiuni gratuite)

SFM_ch1_data_quality

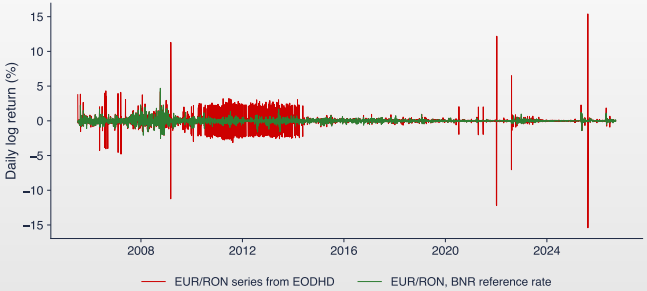
Efectul dividendelor: BET și BET-TR



- 100 de unități investite pe 23 septembrie 2014: 455 în BET (indice de preț), 1 005 în BET-TR (dividende reinvestite)
- Rata de creștere: 13,5% pe an pentru BET, față de 21,2% pentru BET-TR; dividendele adaugă 7,8 puncte procentuale pe an
- Comparați lucruri comparabile: DAX este un indice de randament total, S&P 500 și BET sînt indici de preț

SFM_ch1_data_quality

Erori în date: două serii EUR/RON



- Același curs, 5 154 de zile comune: variația zilnică a seriei EODHD diferă cu peste 1 punct procentual de cea a cursului BNR în 412 zile
- Volatilitatea anualizată: 13,4% (EODHD), față de 4,5% (BNR); cea mai mare variație zilnică a seriei EODHD, în valoare absolută: 15,4%, pe 14 august 2025, revenirea după o cotație greșită din ziua precedentă
- Soluția: folosiți sursa primară a cifrei, cursul de referință BNR

SFM_ch1_data_quality

Trei erori sistematice care supraestimează performanța

- **Survivorship bias**: în eșantion apar doar firmele sau fondurile care încă există
  - firmele cu rezultate slabe au dispărut, deci randamentul mediu este supraestimat Brown et al. (1992)
  - exemplu: componența actuală a BET, urmărită înapoi pînă în 2000
- **Look-ahead bias**: folosirea unei informații care nu era disponibilă la data deciziei
  - rapoarte anuale publicate în martie folosite pentru decizii din ianuarie
- **Data snooping**: încercarea multor reguli pe aceleași date și raportarea celei mai bune
  - sute de predictori publicați ai randamentelor sînt probabil descoperiri false Harvey, Liu & Zhu (2016)
- Fiecare dintre aceste erori supraestimează randamentele; niciuna nu se vede în statisticile descriptive

Lista de verificare a datelor

- Sursa: sursa primară a cifrei și coloana folosită (close sau adjusted close)
- Calendarul: weekenduri, zile libere, zile lipsă; aceleași date pentru toate activele într-o analiză comună
- Salturi: listați cele mai mari 10 randamente în valoare absolută și verificați-le pe fiecare în știri
- Bare plate, volum zero, prețuri repetate: zile libere completate cu ultimul preț sau prețuri stale (neactualizate)
- Unități și monedă: lei, euro, dolari; puncte de indice; prețuri în cenți
- Păstrați fișierul brut și data descărcării; nu modificați niciodată datele manual

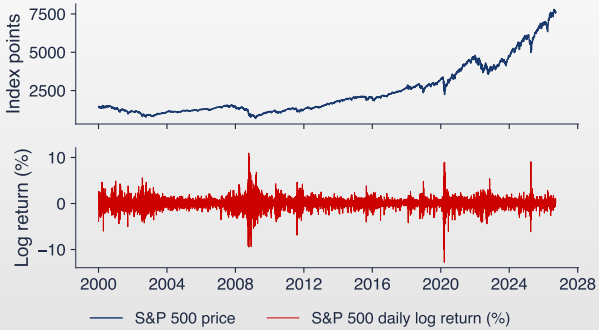
Recapitulare: date

- Prețurile provin de la burse, furnizori și bănci centrale; sursa primară a cifrei este referința
- Folosiți prețuri ajustate pentru acțiuni și indici de randament total cînd dividendele contează
- Datele reale conțin erori: verificați salturile, barele plate și o a doua sursă
- Survivorship bias, look-ahead bias și data snooping supraestimează rezultatele fără să lase urme vizibile

Randamentele în locul prețurilor

- Un preț are unitate de măsură și scară; un randament nu are
  - o variație de 1 leu este mare pentru o acțiune de 2 lei și mică pentru una de 500 de lei
- Randamentele unor active, monede și perioade diferite se pot compara direct
- Prețurile au trend și nu revin la un nivel fix; randamentele oscilează în jurul unui nivel stabil
  - serie **slab staționară**: medie constantă, variață constantă și o covariață care depinde doar de lag
  - majoritatea metodelor statistice ale cursului presupun (cel puțin) staționaritate slabă
- Investitorul este interesat de variația relativă a averii, adică exact de un randament

S&P 500: nivelul prețului și randamentele zilnice



- Sus: prețul are trend crescător, cu oscilații lungi; media și dispersia lui se schimbă în timp
- Jos: randamentele oscilează în jurul lui zero, în perioade calme și agitate (2008, 2020)

SFM_ch1_prices_returns

O verificare formală: testul ADF (1/2)

- **Testul ADF** (augmented Dickey–Fuller) Dickey & Fuller (1979): are seria o rădăcină unitară (un mers aleator)?
  - H_0: rădăcină unitară, seria este nestaționară; H_1: seria este staționară
- Testul regresează variația zilnică a seriei pe nivelul ei din ziua anterioară

$$\Delta y_t = \alpha + \beta t + \gamma\, y_{t-1} + \sum_{j=1}^{p} \delta_j\, \Delta y_{t-j} + \varepsilon_t$$

- Notațiile
  - y_t: seria testată (logaritmul prețului sau randamentul logaritmic); $\Delta y_t = y_t - y_{t-1}$: variația ei de la ziua $t-1$ la ziua t
  - α: constanta; βt: trendul liniar; δ_j: coeficienții celor p variații cu lag, care preiau autocorelația pe termen scurt; ε_t: eroarea
  - $\gamma = 0$: seria nu revine la un nivel (rădăcină unitară); $\gamma < 0$: seria este readusă spre nivelul ei (staționară)
- Respingem H_0 cînd statistica t a lui $\hat{\gamma}$ este sub valoarea critică (din tabelele Dickey–Fuller, nu Student)

O verificare formală: testul ADF (2/2)

- S&P 500, date zilnice din 2000
  - logaritmul prețului (cu trend): statistica $-2{,}08$, valoarea critică 5% $-3{,}41$, p-value 0,56: nu respingem rădăcina unitară
  - randamente logaritmice: statistica $-14{,}3$, valoarea critică 5% $-2{,}86$, p-value $< 0{,}001$: o respingem
- Interpretare
  - p-value: probabilitatea, în ipoteza H_0, de a obține o statistică cel puțin la fel de negativă ca cea observată; sub 0,05 respingem H_0
  - prețurile se comportă ca un mers aleator; randamentele sînt staționare, deci modelăm randamentele, nu prețurile
- Rădăcinile unitare și mersul aleator sînt subiectul Capitolului 7

SFM_ch1_prices_returns

Randamentul simplu

- **Randamentul simplu (net)** de la $t-1$ la t Franke, Härdle & Hafner (2019), Tsay (2010)
  - $R_t = \dfrac{P_t - P_{t-1}}{P_{t-1}} = \dfrac{P_t}{P_{t-1}} - 1$
  - **randamentul brut**: $1 + R_t = P_t / P_{t-1}$
- Mărginit inferior: $R_t \geq -1$ (răspundere limitată: nu puteți pierde mai mult decît ați investit)
- Exemplu: $P_{t-1} = 100$, $P_t = 110$: $R_t = 10\%$
- Este ceea ce vede investitorul în extrasul de cont: „fondul a cîștigat 10%"
- Notație: de aici încolo P_t este prețul de închidere al unui indice sau prețul ajustat al unei acțiuni

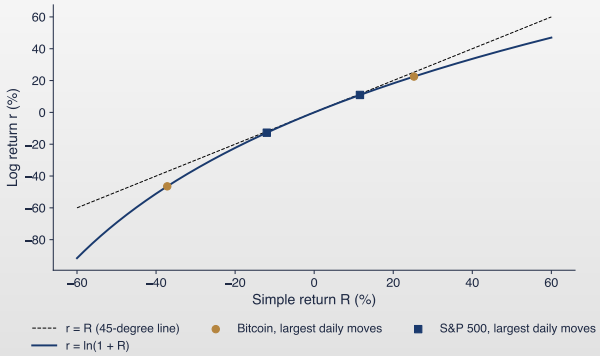
Randamentul logaritmic

- **Randamentul logaritmic** (randamentul compus continuu)
  - $r_t = \ln \dfrac{P_t}{P_{t-1}} = \ln P_t - \ln P_{t-1} = \ln(1 + R_t)$
  - înapoi la randamentul simplu: $R_t = e^{r_t} - 1$
- Originea denumirii: e^r este valoarea la care ajunge un leu la rata r, compusă continuu pe durata perioadei
- Nemărginit: $r_t \in (-\infty, +\infty)$; pierderea totală ($P_t = 0$) înseamnă $r_t = -\infty$
- Exemplu: $P_{t-1} = 100$, $P_t = 110$: $r_t = \ln 1{,}1 = 9{,}53\%$
- Randamentele logaritmice sînt alegerea standard în statistica piețelor financiare: se adună în timp (secțiunea următoare)

Diferența dintre r și R

- Dezvoltarea în serie Taylor a logaritmului în jurul lui 0
  - $r = \ln(1+R) = R - \frac{R^2}{2} + \frac{R^3}{3} - \dots$
  - $\ln$ este concavă, deci $r < R$ pentru orice $R \neq 0$
- Variație zilnică, $R = 1\%$
  - $r = 0{,}995\%$, iar $R - R^2/2 = 0{,}995\%$: diferența $R - r$ este de 0,5 puncte de bază
- Variație mare, $R = 10\%$
  - $r = 9{,}531\%$, iar $R - R^2/2 = 9{,}500\%$: diferența $R - r$ este de 46,9 puncte de bază
- Un **punct de bază** (bp) este 0,01 puncte procentuale
- Concluzie: pentru randamente zilnice $r \approx R$; pentru variații mari și orizonturi lungi diferă

Randamente simple și logaritmice în zilele extreme



- Bitcoin, 12 martie 2020: $R = -37{,}2\%$, dar $r = -46{,}5\%$
- Bitcoin, 7 decembrie 2017: $R = +25{,}2\%$, $r = +22{,}5\%$
- S&P 500, 16 martie 2020: $R = -12{,}0\%$, $r = -12{,}8\%$
- Curba este concavă: în randamente logaritmice, pierderile par mai mari, iar cîștigurile mai mici

SFM_ch1_simple_log_returns

Exemplu rezolvat: +10%, apoi –10%

- Traiectoria prețului $100 \to 110 \to 99$
  - randamente simple: $+10\%$ și -10%; suma lor, 0, este înșelătoare
  - compuse: $1{,}1 \times 0{,}9 - 1 = -1{,}0\%$
- Randamente logaritmice
  - $r_1 = \ln 1{,}1 = 9{,}53\%$, $r_2 = \ln 0{,}9 = -10{,}54\%$
  - suma: $-1{,}01\% = \ln(99/100)$, exact

Regula

- Randamentele simple se compun (se înmulțesc); cele logaritmice se adună

Întrebare pentru sală: +50%, apoi –50%

- O acțiune crește cu 50% într-un an și scade cu 50% în anul următor
- **Ce credeți?** La ce nivel ajunge prețul după doi ani, față de nivelul inițial?
- **Răspuns**
  - $1{,}5 \times 0{,}5 = 0{,}75$: averea s-a modificat cu -25%
  - randamente logaritmice: $\ln 1{,}5 + \ln 0{,}5 = \ln 0{,}75 = -28{,}8\%$
  - o pierdere x cere un cîștig de $x/(1-x)$ pentru revenire: 50% cere 100%

Alegerea tipului de randament

Situația	Randament	Motivul
Agregarea în timp	logaritmic	$r_t(k) = r_t + \cdots + r_{t-k+1}$, o sumă
Portofoliu din mai multe active	simplu	$R_p = \sum_i w_i R_i$, exact
Modelare statistică	logaritmic	nemărginit, mai aproape de o distribuție simetrică
Raportare către investitori	simplu	„fondul a cîștigat 10%" este un randament simplu
Date zilnice, variații mici	oricare	diferența este aproximativ $R^2/2$

- Precizați întotdeauna tipul de randament folosit: un „randament de 10%" este ambiguu

Recapitulare: randamente simple și logaritmice

- $R_t = P_t/P_{t-1} - 1$; $r_t = \ln(P_t/P_{t-1}) = \ln(1 + R_t)$
- $r < R$ pentru orice variație nenulă; diferența este aproximativ $R^2/2$, neglijabilă pentru date zilnice
- Randamentele logaritmice se adună în timp; cele simple se adună între active
- Câștigurile și pierderile de aceeași mărime nu se anulează

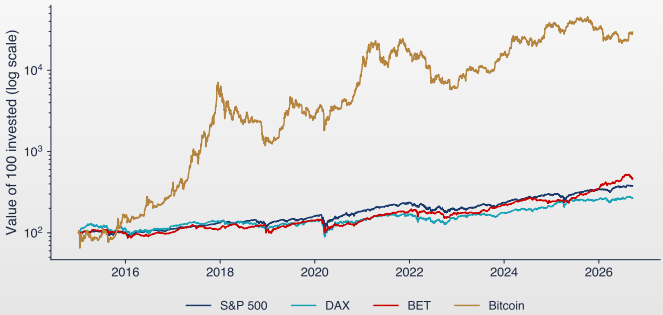
Randamente pe k perioade

- **Simple**: randamentele brute se înmulțesc
  - $1 + R_t(k) = \dfrac{P_t}{P_{t-k}} = (1 + R_t)(1 + R_{t-1}) \cdots (1 + R_{t-k+1})$
- **Logaritmice**: randamentele se adună
  - $r_t(k) = \ln \dfrac{P_t}{P_{t-k}} = r_t + r_{t-1} + \cdots + r_{t-k+1}$
- $R_t(k)$, $r_t(k)$: randamentul simplu și cel logaritmic pe cele k perioade care se încheie la t, de la P_{t-k} la P_t
- Consecință: media și varianța randamentului logaritmic pe k zile rezultă din cele ale randamentelor zilnice
- Randamentele săptămînale, lunare sau anuale se calculează din prețurile de la sfîrșitul fiecărei perioade, nu prin mediere

Randamentul cumulat și averea

- Averea după T perioade, pornind de la W_0: $W_T = W_0 \prod_{t=1}^{T}(1 + R_t) = W_0 \exp\left(\sum_{t=1}^{T} r_t\right)$
- $\prod_{t=1}^{T}$: produsul celor T randamente brute; $\exp(x) = e^x$ inversează logaritmul
- **Randamentul simplu cumulat**: $\mathrm{CSR}_T = W_T/W_0 - 1$
  - cîștigul investitorului, în procente
- **Randamentul logaritmic cumulat**: $\mathrm{CLR}_T = \ln(W_T/W_0) = \sum_t r_t$
  - legătura: $\mathrm{CSR}_T = e^{\mathrm{CLR}_T} - 1$
- Scara logaritmică pe grafice: distanțe verticale egale corespund unor variații procentuale egale

Evoluția a 100 de unități investite, 2015–2026



- 100 de unități investite pe 5 ianuarie 2015, prima zi cu prețuri pentru toate cele patru serii (tabelele CAGR pornesc fiecare serie din prima ei zi din 2015); valori finale: S&P 500 379, DAX 267, BET 470, Bitcoin 29 475
- Bitcoin: randament simplu cumulat 29 375%, randament logaritmic cumulat 5,69; BET: 370% și 1,55
- Fără scara logaritmică, toate seriile, cu excepția Bitcoin, ar părea plate

SFM_ch1_portfolio_returns

Randamentul unui portofoliu

- Portofoliu cu ponderile $w_1, \ldots, w_n$, $\sum_i w_i = 1$, fixate la începutul perioadei
  - R_i, r_i: randamentul simplu și cel logaritmic al activului $i = 1, \ldots, n$ pe perioadă; w_i: ponderea averii investite în activul i
  - **simplu**: $R_p = \sum_i w_i R_i$, exact
  - **logaritmic**: $r_p = \ln\left(\sum_i w_i e^{r_i}\right) \neq \sum_i w_i r_i$
- Exemplu: 60% în activul 1 ($R_1 = 10\%$), 40% în activul 2 ($R_2 = -5\%$)
  - $R_p = 0{,}6 \times 10\% + 0{,}4 \times (-5\%) = 4{,}0\%$; exact $r_p = \ln(1 + R_p) = 3{,}92\%$
  - randamente logaritmice ponderate: $0{,}6 \times 9{,}53\% + 0{,}4 \times (-5{,}13\%) = 3{,}67\%$, prea mic
- Ponderile se modifică odată cu prețurile: un portofoliu cu ponderi fixe trebuie **reechilibrat**

SFM_ch1_portfolio_returns

Date reale: un portofoliu TLV–SNP

- 50% Banca Transilvania, 50% OMV Petrom, reechilibrat zilnic, 2015–2026
- O zi: randamentul logaritmic exact față de suma ponderată a randamentelor logaritmice
  - diferența medie în valoare absolută 0,38 puncte de bază; cea mai mare 13,9 puncte de bază, pe 31 mai 2016
- Pe întreaga perioadă diferențele mici se acumulează
  - randamentul logaritmic cumulat exact 2,56, deci 1 leu devine 13,0 lei
  - randamentele logaritmice ponderate conduc la 2,47, deci doar la 11,8 lei
- Rata de creștere: portofoliul 24,5% pe an; TLV 27,3%; SNP 19,7%
- Regula: construiți portofoliile cu randamente simple, apoi treceți la logaritmi dacă este nevoie

SFM_ch1_portfolio_returns

Recapitulare: randamente pe mai multe perioade și ale portofoliilor

- În timp: înmulțiți randamentele brute sau adunați randamentele logaritmice
- Averea $W_T = W_0 e^{\mathrm{CLR}_T}$; $\mathrm{CSR}_T = e^{\mathrm{CLR}_T} - 1$
- Între active: $R_p = \sum_i w_i R_i$ exact; varianta logaritmică este doar o aproximare
- Erorile zilnice mici de aproximare devin vizibile pe un deceniu

Anualizarea mediei și a volatilității

- A = numărul de observații pe an: aproximativ 252 pentru burse, 365 pentru cripto, 52 pentru date săptămînale, 12 pentru date lunare
- μ, σ: media și abaterea standard a randamentelor la frecvența datelor (de exemplu, zilnică)
- **Media**: randamentele logaritmice se adună, deci $\mu_{\text{an}} = A\,\mu$
  - aceeași regulă se folosește pentru media randamentelor simple (anualizare aritmetică)
- **Volatilitatea** = abaterea standard a randamentelor: $\sigma_{\text{an}} = \sqrt{A}\,\sigma$, **regula rădăcinii pătrate a timpului**
  - valabilă cînd randamentele sînt necorelate și au variață constantă: atunci $\text{Var}(r_1 + \cdots + r_A) = A\,\sigma^2$, iar rădăcina pătrată dă $\sqrt{A}\,\sigma$
- S&P 500: σ zilnic $= 1{,}12\%$, deci $\sigma_{\text{an}} = 1{,}12\% \times 15{,}87 = 17{,}7\%$
  - Bitcoin: $\sqrt{365} = 19{,}10$ conduce la 66,7%; factorul bursier $\sqrt{252}$ ar conduce doar la 55,4%

Rata anuală compusă de creștere

- **CAGR** (compound annual growth rate, rata anuală compusă de creștere): rata anuală constantă care transformă P_0 în P_T în Y ani
  - $\text{CAGR} = \left(P_T/P_0\right)^{1/Y} - 1$, cu Y în ani calendaristici
- Legătura cu randamentele logaritmice: CAGR = exp(media anuală a randamentelor logaritmice) − 1
- S&P 500, 2015–2026: media randamentelor logaritmice 11,2% pe an, $e^{11,2\%} - 1 = 11,9\%$, iar CAGR este 11,9%
- CAGR folosește doar primul și ultimul preț: nu spune nimic despre traiectoria dintre ele
- Se mai numește randamentul mediu geometric

Precizia estimării mediei

- Eroarea standard a mediei anuale: $\text{SE}(\hat{\mu}_{\text{an}}) = \sigma_{\text{an}}/\sqrt{Y}$, cu Y ani de date
  - $\hat{\mu}_{\text{an}}$: media anuală estimată (notația $\hat{}$ indică o valoare estimată din eșantion)
  - depinde de numărul de **ani**, nu de numărul de observații
  - intervalul de încredere 95%: $\hat{\mu}_{\text{an}} \pm 1{,}96\,\text{SE}$, unde 1,96 este cuantila de 97,5% a distribuției Normale standard
- S&P 500, 11,7 ani: media anuală a randamentelor logaritmice 11,2%, SE 5,2%
  - intervalul de încredere 95%: [1,1%; 21,4%]
- Bitcoin: $47{,}4\% \pm 1{,}96 \times 19{,}5\%$
  - intervalul este [9,2%; 85,6%]: foarte larg
- Volatilitatea, în schimb, se estimează precis din date zilnice

SFM_ch1_descriptive_stats

Întrebare pentru sală: mai multe date, o medie mai bună?

- Înlocuiți 10 ani de date zilnice S&P 500 cu 10 ani de date la 5 minute: 78 de bare într-o zi de tranzacționare de 6,5 ore, deci de 78 de ori mai multe observații
- **Ce credeți?** Devine mai mică eroarea standard a randamentului mediu anual?
- **Răspuns**
  - Nu: media pe 10 ani este determinată de primul și ultimul preț, $\hat{\mu} = \ln(P_T/P_0)/Y$
  - doar mai mulți **ani** îi reduc eroarea standard, ca $1/\sqrt{Y}$
  - estimarea volatilității, în schimb, se îmbunătățește cînd crește frecvența

Recapitulare: anualizarea

- Media $\times A$, volatilitatea $\times\sqrt{A}$, cu A numărul real de observații pe an
- $\sqrt{A}$ presupune randamente necorelate, cu varianță constantă
- CAGR $= (P_T/P_0)^{1/Y} - 1$, exponențiala mediei anuale a randamentelor logaritmice minus 1
- Mediile randamentelor sînt imprecise: SE $= \sigma_{\mathbf{an}}/\sqrt{Y}$

Poziție și dispersie

- Eșantion $r_1, \ldots, r_n$ de randamente logaritmice zilnice (în %)
- **Media**: $\bar{r} = \frac{1}{n} \sum_{t=1}^{n} r_t$, randamentul zilnic mediu
- **Varianța**: $\hat{\sigma}^2 = \frac{1}{n-1} \sum_{t=1}^{n} (r_t - \bar{r})^2$; **abaterea standard** $\hat{\sigma} = \sqrt{\hat{\sigma}^2}$
  - în finanțe abaterea standard a randamentelor se numește **volatilitate**
- **Cuantila** q_α: valoarea sub care se află o fracție α din randamente
  - mediana $q_{0,5}$; cuantila de 1% $q_{0,01}$ stă la baza VaR 1% (Capitolul 10)
- **Minimul** și **maximul**: cea mai slabă și cea mai bună zi, întotdeauna cu datele lor

Forma: asimetria și boltirea

- **Asimetria** (skewness): $\hat{S} = \frac{1}{n}\sum_t (r_t - \bar{r})^3 / \hat{\sigma}^3$
  - puterea a treia păstrează semnul fiecărei abateri, deci pierderile mari împing $\hat{S}$ sub 0, iar cîștigurile mari peste 0
  - $S = 0$: simetrică; $S < 0$: coadă stîngă mai lungă, pierderile mari mai frecvente decît cîștigurile mari
- **Boltirea** (kurtosis): $\hat{K} = \frac{1}{n}\sum_t (r_t - \bar{r})^4 / \hat{\sigma}^4$; **excesul de boltire** (excess kurtosis) $\hat{K} - 3$
  - puterea a patra face ca zilele extreme să domine $\hat{K}$
  - distribuția Normală are $K = 3$, deci exces de boltire 0
  - exces de boltire > 0: **cozi groase**, zile extreme mai frecvente decît în cazul distribuției Normale
- Ambele sînt sensibile la zile extreme izolate: raportați-le împreună cu extremele
- Cozile groase și asimetria sînt **fapte stilizate** ale randamentelor Cont (2001); testele formale în Capitolul 2

Randamente logaritmice zilnice: indici și Bitcoin, 2015–2026

Seria	Media	Abaterea std.	Asim.	Exces bolt.	Minimul (data)	Maximul
S&P 500	0,045	1,12	−0,64	15,8	−12,8 (16 mar. 2020)	9,1
DAX	0,032	1,20	−0,56	10,0	−13,1 (12 mar. 2020)	10,4
BET	0,053	0,98	−1,41	17,8	−11,9 (19 dec. 2018)	6,8
BET-TR	0,080	0,98	−1,43	18,1	−11,9 (19 dec. 2018)	6,7
Bitcoin	0,130	3,49	−0,73	12,2	−46,5 (12 mar. 2020)	22,5

- Toate valorile în %, cu excepția asimetriei și a excesului de boltire; fiecare serie pe calendarul ei
- Bitcoin: abaterea standard zilnică de aproximativ trei ori mai mare decît a indicilor bursieri
- Toate seriile: asimetrie negativă și exces de boltire mare
- Cele mai slabe zile se concentrează în martie 2020 (COVID-19) și, pentru BET, pe 19 decembrie 2018

SFM_ch1_descriptive_stats

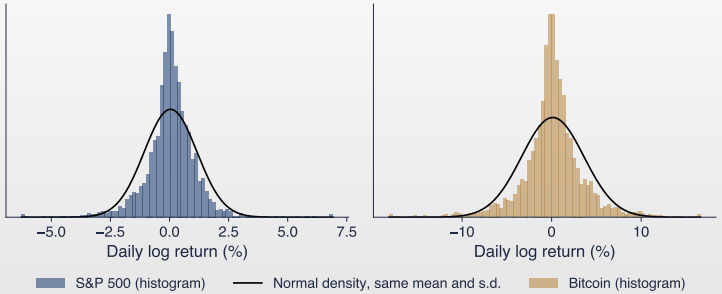
Randamente logaritmice zilnice: cinci acțiuni BVB, 2015–2026

Seria	Media	Abaterea std.	Asim.	Exces bolt.	Minimul (data)	Maximul
Banca Transilvania (TLV)	0,106	1,69	−0,55	25,2	−22,2 (19 dec. 2018)	20,5
OMV Petrom (SNP)	0,079	1,54	−0,66	9,2	−15,0 (9 mar. 2020)	10,0
BRD	0,075	1,61	−0,51	9,2	−18,2 (19 dec. 2018)	10,3
Nuclearelectrica (SNN)	0,118	1,53	0,29	5,2	−9,4 (9 mar. 2020)	11,3
Transgaz (TGN)	0,091	1,55	0,01	5,3	−10,1 (19 dec. 2018)	11,1

- Acțiunile individuale sînt mai volatile decît indicele BET: efectul diversificării
- Banca Transilvania și BRD au pierdut cel mai mult pe 19 decembrie 2018, după anunțarea taxei pe activele bancare, adoptată prin OUG 114/2018
- Excesul de boltire diferă mult între acțiuni: comparați TLV cu SNN
- Problemă de date: prețul ajustat EODHD al TLV din 30 mai 2016 este greșit (−15,0%, apoi +20,5% pe 31 mai); este păstrat aici, iar fără aceste două zile excesul de boltire este 18,1
- Crahurile de la BVB au o statistică proprie Pele & Mazurencu-Marinescu (2012)

SFM_ch1_descriptive_stats

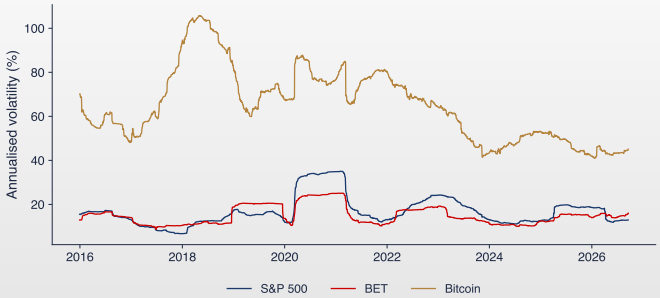
Histogramele față de distribuția Normală



- Linia neagră: densitatea Normală cu aceeași medie și abatere standard; axa orizontală este tăiată la cuantilele de 0,1% și 99,9%, iar zilele din afara lor sînt numărate în barele de la margine
- S&P 500: 79,6% din zile sînt la cel mult o abatere standard de medie (distribuția Normală: 68,3%)
- Dincolo de trei abateri standard: 1,39% din zile, față de 0,27% în cazul distribuției Normale, adică de 5,2 ori mai multe

SFM_ch1_descriptive_stats

Volatilitatea variabilă în timp



- Volatilitatea pe ferestre mobile de un an, anualizată: S&P 500 între 7% și 35%; Bitcoin între 41% și 106%
- O singură cifră pentru tot eșantionul ascunde perioadele calme și pe cele agitate
- Estimatori de volatilitate și modele GARCH: Capitolele 8 și 9

SFM_ch1_descriptive_stats

Recapitulare: statistici descriptive

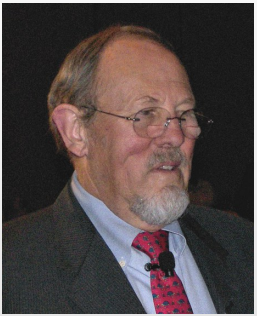
- Raportați media, abaterea standard, asimetria, excesul de boltire, cuantilele și extremele cu datele lor
- Randamente zilnice: medie apropiată de 0, volatilitate 1,0–1,7% pe zi (Bitcoin 3,5%), cozi groase, adesea asimetrie negativă
- Zilele extreme sînt de cîteva ori mai frecvente decît în cazul distribuției Normale
- Volatilitatea nu este constantă în timp

Limitele unui singur indicator

- Randamentul, luat izolat, favorizează asumarea de risc: o poziție cu efect de levier poate avea un randament mare ani la rînd
- Trei întrebări pe care le pune un investitor
  - cît a crescut? (CAGR)
  - cît de mari au fost fluctuațiile? (volatilitate, abaterea negativă)
  - cît de adîncă a fost cea mai mare scădere? (drawdown maxim)
- Indicatorii **ajustați la risc** împart un randament la o măsură a riscului: Sharpe, Sortino, Calmar
- Toate sînt estimări dintr-un singur eșantion: fiecare are o eroare de estimare

Raportul Sharpe

- **Raportul Sharpe** (Sharpe ratio) Sharpe (1966), Sharpe (1994): randamentul în exces pe unitatea de risc total
  - $\mathrm{SR} = \dfrac{\mu - r_f}{\sigma}$; μ randamentul mediu, r_f rata fără risc, σ volatilitatea
  - de la zilnic la anual: $\mathrm{SR}_{\mathrm{an}} = \sqrt{A}\,\mathrm{SR}_{\mathrm{zi}}$
- Aici: media randamentelor simple zilnice, $r_f = 0$
  - activele sînt în lei, euro și dolari, fiecare cu propria rată fără risc
  - S&P 500: SR zilnic 0,046, SR anual 0,72
- Același raport Sharpe înseamnă același randament pe unitatea de risc, oricare ar fi efectul de levier



William F. Sharpe, Premiul Nobel pentru economie 1990

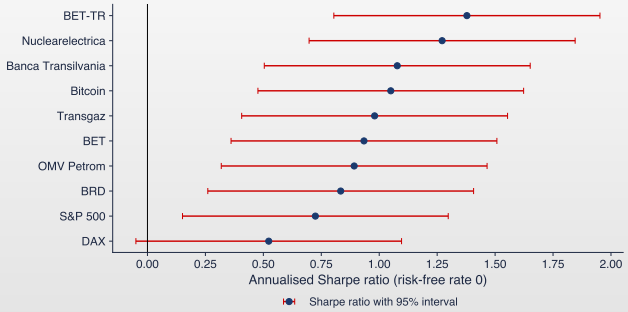
Foto: Larry D. Moore (2007); CC BY 4.0; Wikimedia Commons

Eroarea de estimare a raportului Sharpe

- Eroarea standard pentru randamente i.i.d. (independente și identic distribuite) Lo (2002)
  - $\mathrm{SE}(\widehat{\mathrm{SR}}) \approx \sqrt{(1 + \mathrm{SR}^2/2)/n}$ pe perioadă; anual: înmulțiți cu $\sqrt{A}$
  - $\widehat{\mathrm{SR}}$: raportul Sharpe estimat din eșantion; n: numărul de randamente; A: numărul de randamente pe an
  - aproximativ $1/\sqrt{Y}$ pentru raportul Sharpe anual: circa 0,3 cu 12 ani de date
- S&P 500, 2015–2026: SR = 0,72, SE 0,29
  - intervalul de încredere 95%: [0,15; 1,30]
- Cozile groase și volatility clustering fac eroarea reală mai mare decît cea dată de această formulă
- Alegerea celei mai bune dintre multe strategii supraestimează raportul Sharpe al strategiei alese Bailey & López de Prado (2014)

SFM_ch1_performance

Raportul Sharpe cu intervale de încredere 95%, 2015–2026



- Cea mai mare estimare punctuală: BET-TR
- Toate intervalele se suprapun: clasamentul este incert; suprapunerea nu este un test formal, care compară diferența a două rapoarte Sharpe cu propria ei eroare standard
- Intervalul DAX conține 0
- Lățimile intervalelor sînt aproape egale: depind mai ales de lungimea eșantionului

SFM_ch1_performance

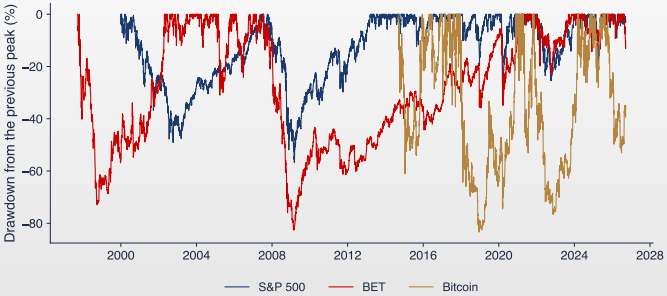
Raportul Sortino

- Investitorii se tem de pierderi, nu de cîștiguri; volatilitatea le tratează însă la fel
- **Abaterea negativă** (downside deviation) sub un prag τ (aici $\tau = 0$)
  - $\sigma_D = \sqrt{\frac{1}{n}\sum_t \min(R_t - \tau, 0)^2}$, contează doar zilele sub prag
  - $\min(R_t - \tau, 0)$: deficitul sub prag din ziua t, zero în zilele peste prag; n: toate zilele eșantionului
- **Raportul Sortino** (Sortino ratio) Sortino & van der Meer (1991): $\text{Sortino} = \frac{\mu - \tau}{\sigma_D}$, anualizat ca raportul Sharpe
- Pentru o distribuție simetrică $\sigma_D \approx \sigma/\sqrt{2}$, deci Sortino $\approx 1{,}4\times$ Sharpe
- S&P 500, 2015–2026: Sharpe 0,72, Sortino 1,02

Drawdown și drawdown maxim

- **Drawdown**: scăderea față de cel mai mare preț atins pînă atunci
  - $\mathrm{DD}_t = \dfrac{P_t}{\max_{s \le t} P_s} - 1 \le 0$; zero la un nou maxim
  - $\max_{s \le t} P_s$: cel mai mare preț din toate zilele s de pînă la ziua t, inclusiv
- **MDD** (maximum drawdown, drawdown maxim): $\mathrm{MDD} = \min_t \mathrm{DD}_t$, cea mai mare scădere de la un vîrf la minimul următor
  - raportați și datele: vîrful, minimul și **revenirea** la vechiul vîrf
- O pierdere x cere un cîștig $x/(1-x)$ pentru revenire
  - 10% cere 11%; 20% cere 25%; 50% cere 100%; 80% cere 400%
- MDD depinde de traiectorie și crește cu lungimea eșantionului Magdon-Ismail et al. (2004)

Drawdown-uri pe întreaga istorie



- S&P 500: MDD −56,8%, de la 9 oct. 2007 la 9 mar. 2009, cu revenire pe 28 mar. 2013
- BET (din 1997): MDD −82,5%, de la 24 iul. 2007 la 25 feb. 2009, cu revenire abia pe 16 mar. 2021
- Bitcoin: MDD −83,4%, de la 16 dec. 2017 la 15 dec. 2018

SFM_ch1_performance

Interpretarea graficului drawdown-urilor

- Adîncimea
  - după un drawdown de $-82{,}5\%$, BET a avut nevoie de $+473\%$ pentru revenire; Bitcoin, de $+502\%$
- Durata
  - BET a avut nevoie de 13,6 ani ca să revină la vîrful din 2007
  - ponderea zilelor cu peste 20% sub vîrf: S&P 500 27%, BET 56%, Bitcoin 66%
- Volatilitatea și raportul Sharpe ignoră ordinea randamentelor; drawdown-ul ține cont de ea
- În perioada de comparație 2015–2026, MDD-ul fiecărui indice bursier s-a produs în februarie–martie 2020; pentru Bitcoin și pentru mai multe acțiuni BVB, MDD-ul s-a produs la alte date

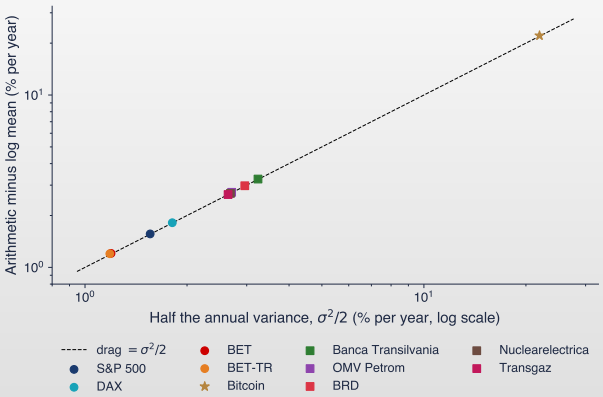
Raportul Calmar

- **Raportul Calmar** (Calmar ratio): creșterea pe unitatea celei mai mari scăderi, $\text{Calmar} = \dfrac{\text{CAGR}}{|\text{MDD}|}$
- Popular printre administratorii de fonduri: MDD este pierderea pe care investitorul o suportă efectiv
- S&P 500, 2015–2026: CAGR 11,9%, MDD $-33{,}9\%$
  - Calmar $= 11{,}9/33{,}9 = 0{,}35$
- Un singur eveniment determină numitorul: raportul Calmar este și mai instabil decît raportul Sharpe
- Comparați rapoartele Calmar doar pe aceeași perioadă

Volatility drag

- Pentru un randament logaritmic cu media m și varianța σ^2, randamentul simplu are media $\approx m + \sigma^2/2$
  - deci rata de creștere este $\mu_{\text{log}} \approx \mu_{\text{arith}} - \sigma^2/2$
  - μ_{arith}: media randamentelor simple; μ_{log}: media randamentelor logaritmice, adică rata de creștere a averii
  - diferența $\sigma^2/2$ este **volatility drag** Booth & Fama (1992); provine din concavitatea lui ln (inegalitatea lui Jensen)
- Două active cu aceeași medie aritmetică, 10% pe an
  - $\sigma = 5\%$: drag 0,12%, creștere 9,88% pe an; după 10 ani 1 leu devine 2,68
  - $\sigma = 30\%$: drag 4,50%, creștere 5,50% pe an; după 10 ani 1 leu devine 1,73
- Dublarea volatilității mărește de patru ori volatility drag

Volatility drag pe date reale



- Fiecare punct: un activ, 2015–2026
- Vertical: media anuală aritmetică minus cea logaritmică; orizontal: $\sigma^2/2$
- Toate punctele se află pe dreaptă: formula se verifică pe date reale
- Bitcoin: media aritmetică 69,5%, media logaritmică 47,4%, drag 22,1 puncte procentuale; $\sigma^2/2 = 21{,}9\%$

SFM_ch1_volatility_drag

Întrebare pentru sală: ce randament pentru Bitcoin?

- Bitcoin, 2015–2026: media aritmetică a randamentelor simple 69,5% pe an; CAGR 60,6%
- **Ce credeți?** Care dintre cele două cifre descrie creșterea averii unui investitor?
- **Răspuns**
  - CAGR: este rata care a transformat efectiv primul preț în ultimul
  - media aritmetică este randamentul așteptat al unui singur an, nu creșterea pe mulți ani
  - volatility drag de 22,1 puncte separă media aritmetică de media randamentelor logaritmice, 47,4%; CAGR este $e^{47,4\%} - 1 = 60,6\%$

Indicatori de performanță: indici și Bitcoin, 2015–2026

	CAGR %	Vol. %	Sharpe	Sortino	MDD %	Calmar	Drag %
S&P 500	11,9	17,6	0,72	1,02	−33,9	0,35	1,56
DAX	8,5	19,0	0,52	0,73	−38,8	0,22	1,82
BET	14,1	15,4	0,93	1,28	−31,1	0,45	1,21
BET-TR	22,1	15,4	1,38	1,92	−31,1	0,71	1,20
Bitcoin	60,6	66,2	1,05	1,55	−83,4	0,73	22,13

- Anualizate cu frecvența reală a fiecărei serii; Vol.: volatilitatea randamentelor simple zilnice; Sharpe și Sortino cu $r_f = 0$
- Bitcoin: de departe cel mai mare CAGR, cea mai mare volatilitate și cel mai adînc drawdown
- BET-TR față de BET: același risc, dar dividendele cresc toți indicatorii calculați pe baza randamentului

SFM_ch1_performance

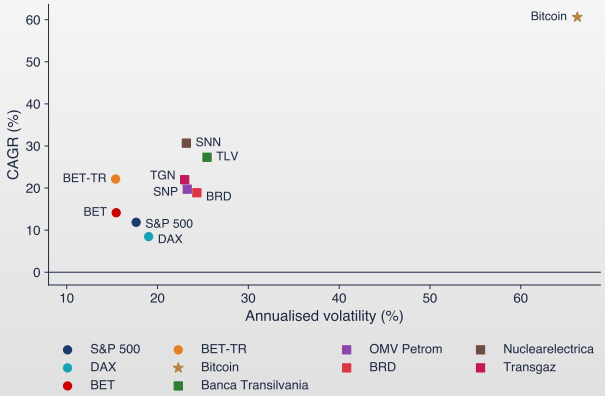
Indicatori de performanță: acțiuni BVB, 2015–2026

	CAGR %	Vol. %	Sharpe	Sortino	MDD %	Calmar	Drag %
Banca Transilvania (TLV)	27,3	25,5	1,08	1,59	−38,9	0,70	3,25
OMV Petrom (SNP)	19,7	23,3	0,89	1,29	−45,4	0,43	2,73
BRD	18,9	24,3	0,83	1,21	−35,6	0,53	2,98
Nuclearelectrica (SNN)	30,7	23,2	1,27	2,00	−36,8	0,83	2,69
Transgaz (TGN)	22,0	23,0	0,98	1,49	−44,8	0,49	2,65

- Prețuri ajustate: dividendele sînt incluse
- Acțiunile individuale: volatilitate mai mare și drawdown-uri mai adînci decît indicele BET
- Nu uitați eroarea standard, de circa 0,29 pentru fiecare raport Sharpe: diferențele mici nu sînt semnificative

SFM_ch1_performance

Risc și randament, 2015–2026



- Orizontal: volatilitatea anualizată; vertical: CAGR
- Cel mai bun după CAGR: Bitcoin; după Sharpe: BET-TR; după Sortino: SNN; după Calmar: SNN
- Cea mai mică volatilitate: BET-TR; cel mai mic MDD în valoare absolută: BET-TR ($-31{,}115\%$) și BET ($-31{,}124\%$), practic la egalitate
- Clasamentul depinde de indicator: precizați întotdeauna indicatorul folosit

SFM_ch1_performance

Recapitulare: indicatori de performanță

- CAGR (creștere), volatilitate (risc), Sharpe și Sortino (randament pe unitatea de risc)
- MDD și Calmar: cea mai mare scădere, pe care investitorul o suportă efectiv
- Volatility drag: $\mu_{\text{log}} \approx \mu_{\text{arith}} - \sigma^2/2$, creșterea este sub media aritmetică
- Fiecare indicator este o estimare: comparați diferențele cu erorile lor standard

O întrebare de cercetare deschisă

- **Persistă în timp rapoartele Sharpe?**
  - o acțiune BVB cu un raport Sharpe mare în 2015–2020 a avut un raport Sharpe mare și în 2021–2026?
  - dacă nu, clasamentul fondurilor sau al acțiunilor după raportul Sharpe din trecut îi este de puțin folos unui investitor
- De ce este deschisă: eșantioane scurte, puține acțiuni și o eroare standard de circa 0,4 pentru un raport Sharpe estimat pe jumătate din eșantionul nostru
- Rezultate conexe: multe regularități ale randamentelor, publicate în literatură, nu se confirmă pe date noi Harvey, Liu & Zhu (2016)
- Instrumentele AI pot accelera un astfel de studiu; nu înlocuiesc verificarea lui Wang et al. (2023)

Utilizări posibile ale instrumentelor AI

- **Literatura**: identificarea lucrărilor despre persistența performanței și rezumarea metodelor lor
- **Cod**: o primă versiune a unei funcții Python pentru raportul Sharpe pe ferestre mobile, pe datele cursului
- **Robustețe**: propunerea altor ferestre, altor indicatori (Sortino, Calmar), altor piețe
- **Explicație**: transformarea unui tabel de rezultate într-o primă versiune a interpretării
- Exemplu de prompt
  - `Write a Python function that computes the annualised Sharpe ratio of daily simple returns on non-overlapping 3-year windows, with the Lo (2002) standard error.`

Verificări necesare

- Datele: prețul ajustat pentru acțiuni, același calendar pentru toate activele, perioada corectă
- Formulele: factorul de anualizare ($\sqrt{252}$ sau $\sqrt{365}$), randamente simple sau logaritmice, n sau $n-1$
- Cifrele: recalculați două sau trei valori pe exemple mici sau cu o a doua metodă
- Referințele: fiecare lucrare citată trebuie să existe; verificați DOI-ul
- Selecția: dacă ați încercat zece ferestre, precizați acest lucru; un backtest repetat de multe ori duce la supraajustare (overfitting) Bailey et al. (2014)

Idee de proiect

- **Întrebarea**: persistă rapoartele Sharpe ale acțiunilor BVB de la o perioadă la alta?
  - date: prețurile ajustate a 10–15 acțiuni BVB, 2010–2026, din datele cursului
- Pași
  - împărțiți eșantionul în două jumătăți; calculați rapoartele Sharpe și erorile lor standard în fiecare jumătate
  - măsurați persistența cu corelația rangurilor dintre cele două jumătăți
  - repetați cu rapoartele Sortino și Calmar
- Livrabile: un tabel, un grafic și un paragraf despre ce pot și ce nu pot arăta datele
- Declarați orice utilizare a AI și enumerați erorile AI pe care le-ați corectat

Formule de reținut

Mărimea	**Formula**
Randament simplu	$R_t = P_t/P_{t-\mathbf{1}} - 1$
Randament logaritmic	$r_t = \ln(P_t/P_{t-\mathbf{1}}) = \ln(1+R_t)$
Randamente pe k perioade	$1 + R_t(k) = \prod_{j=\mathbf{0}}^{k-\mathbf{1}}(1+R_{t-j}), \quad r_t(k) = \sum_{j=\mathbf{0}}^{k-\mathbf{1}} r_{t-j}$
Portofoliu	$R_p = \sum_i w_i R_i$
Anualizare	$\mu_{\mathbf{an}} = A\,\mu, \quad \sigma_{\mathbf{an}} = \sqrt{A}\,\sigma$
CAGR	$(P_T/P_{\mathbf{0}})^{\mathbf{1}/Y} - 1$
Sharpe, Sortino	$(\mu - r_f)/\sigma, \quad (\mu - \tau)/\sigma_D$
MDD, Calmar	$\min_t \left(P_t/\max_{s\le t} P_s - 1\right), \quad \mathrm{CAGR}/\lvert\mathrm{MDD}\rvert$
Volatility drag	$\mu_{\mathbf{log}} \approx \mu_{\mathbf{arith}} - \sigma^{\mathbf{2}}/2$
SE a mediei și a raportului Sharpe	$\sigma_{\mathbf{an}}/\sqrt{Y}, \quad \sqrt{(1+\mathrm{SR}^{\mathbf{2}}/2)/n}$

Autoevaluare

- **Întrebare**: un fond raportează randamente logaritmice zilnice cu media 0,05% și abaterea standard 1%; care sînt valorile anuale ale mediei și volatilității?
  - **Răspuns**: $252 \times 0{,}05\% = 12{,}6\%$ și $\sqrt{252} \times 1\% = 15{,}9\%$
- **Întrebare**: de ce este media aritmetică a randamentelor Bitcoin mult peste CAGR?
  - **Răspuns**: volatility drag $\sigma^2/2$ este mare cînd σ este mare
- **Întrebare**: un indice scade cu 60%; ce cîștig este necesar pentru revenirea la nivelul anterior?
  - **Răspuns**: $0{,}6/0{,}4 = 150\%$
- Urmează: Capitolul 2, distribuții clasice și fapte stilizate ale randamentelor

Bibliografie (1/2)

- Bailey, D.H., Borwein, J.M., López de Prado, M., Zhu, Q.J. (2014). Pseudo-mathematics and financial charlatanism: the effects of backtest overfitting on out-of-sample performance. *Notices of the American Mathematical Society*, 61(5), 458–471.
- Bailey, D.H., López de Prado, M. (2014). The deflated Sharpe ratio: correcting for selection bias, backtest overfitting, and non-normality. *Journal of Portfolio Management*, 40(5), 94–107.
- Booth, D.G., Fama, E.F. (1992). Diversification returns and asset contributions. *Financial Analysts Journal*, 48(3), 26–32.
- Borak, S., Härdle, W.K., López-Cabrera, B. (2013). *Statistics of Financial Markets: Exercises and Solutions*, 2nd ed. Springer.
- Brown, S.J., Goetzmann, W., Ibbotson, R.G., Ross, S.A. (1992). Survivorship bias in performance studies. *Review of Financial Studies*, 5(4), 553–580.
- Campbell, J.Y., Lo, A.W., MacKinlay, A.C. (1997). *The Econometrics of Financial Markets*. Princeton University Press.
- Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2), 223–236.
- Dickey, D.A., Fuller, W.A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366), 427–431.
- Franke, J., Härdle, W.K., Hafner, C.M. (2019). *Statistics of Financial Markets: An Introduction*, 5th ed. Springer.

Bibliografie (2/2)

- Harvey, C.R., Liu, Y., Zhu, H. (2016). . . . and the cross-section of expected returns. *Review of Financial Studies*, 29(1), 5–68.
- Lo, A.W. (2002). The statistics of Sharpe ratios. *Financial Analysts Journal*, 58(4), 36–52.
- 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.
- Pele, D.T., Mazurencu-Marinescu, M. (2012). Modelling stock market crashes: the case of Bucharest Stock Exchange. *Procedia – Social and Behavioral Sciences*, 58, 533–542.
- Sharpe, W.F. (1966). Mutual fund performance. *Journal of Business*, 39(S1), 119–138.
- Sharpe, W.F. (1994). The Sharpe ratio. *Journal of Portfolio Management*, 21(1), 49–58.
- Sortino, F.A., van der Meer, R. (1991). Downside risk. *Journal of Portfolio Management*, 17(4), 27–31.
- Tsay, R.S. (2010). *Analysis of Financial Time Series*, 3rd ed. Wiley.
- Wang, H., Fu, T., Du, Y., Gao, W., et al. (2023). Scientific discovery in the age of artificial intelligence. *Nature*, 620, 47–60.